1991

Recalling in a state of natural or experimental depression

Paula T. Hertel
Trinity University, phertel@trinity.edu

S. Rude

Repository Citation

Hertel, P. T., & Rude, S. (1991b). Recalling in a state of natural or experimental depression. *Cognitive Therapy and Research*, 15, 103-127.

Recalling in a State of Natural or Experimental Depression[1]

Paula T. Hertel[2]

Trinity University

Stephanie S. Rude

University of Texas, Austin

In three experiments we attempted to extend the cognitive-effort account of depressive deficits in memory to naturally depressed college students. This account maintains that depression reduces attentional resources, thereby impairing performance on demanding tasks, and has received support through experimental inductions of depressed moods. Nondepressed, naturally depressed, and (in Experiment 2) experimentally depressed college students performed unannounced tests of free recall following learning tasks with two levels of difficulty and (in Experiment 2) two degrees of structure. In Experiments 1 and 2 we measured cognitive effort on those tasks via latencies on a secondary task. Latencies and subsequent recall increased with the structure and difficulty of the learning task for nondepressed and naturally depressed subjects, but these effects were reduced or absent for experimentally depressed subjects. When the secondary task was omitted (Experiment 3), naturally depressed students still recalled without a deficit. We discuss possible differences associated with the two types of depression and implications for the cognitive-effort account.

KEY WORDS: depression; memory; recall; difficulty.

[2] Address all correspondence concerning this article to Paula Hertel, Department of Psychology, Trinity University, 715 Stadium Drive, San Antonio, Texas 78212.

Sad moods and depression are frequently associated with difficulties in remembering. The most common account of these difficulties relates them to corresponding deficiencies in the allocation of attentional resources (see Ellis & Ashbrook, 1988; Hasher & Zacks, 1979; Roy-Byrne, Weingartner, Bierer, Thompson, & Post, 1986). We refer to this general view as the cognitive-effort account of depressive deficits in memory.

The cognitive-effort account maintains that depression reduces or uses a portion of the limited capacity of conscious attention that would otherwise be allocated to a particular cognitive task. Depressive deficits should occur, therefore, in relatively attention-demanding tasks or in tasks that benefit from prior effortful procedures (such as free recall). The account implies that this drain or occupation of attentional resources by depression is pervasive and relatively constant. Other tasks are assumed to suffer to the extent that they require more attention than remains.

The cognitive-effort account has received support in experiments in which nondepressed students have been experimentally induced to feel depressed (see Ellis, Thomas, & Rodriguez, 1984), but only mixed support in experiments performed with naturally depressed students. For example, Potts, Camp, and Coyne (1989) attempted to replicate the results of Ellis et al. (1984, Experiment 1), but failed to find an overall depressive deficit in recalling words processed in base or elaborated sentences. The experiments described here were attempts to extend the results from experimentally induced depression in Ellis et al. (1984, Experiment 3) to the domain of naturally depressed students and thereby to provide support for the cognitive-effort account. We cannot assume the compatibility of naturally occurring depression with a laboratory procedure for inducing sad mood, and therefore this extension to samples of depressed subjects is an important step in theorizing about depressive memory (see Ingram, 1989). If "experimental depression" can effectively model natural depression in their relation to memory, then the results from experiments that manipulate mood might help us to make causal inferences about the relation of depression and memory (cf. Riskind, 1989). But if the correspondence between the types of depression is lacking, then conclusions from experiments employing experimental inductions of mood are confined to that domain.

EXPERIMENT

In a frequently cited series of experiments on depressive deficits in free recall, Ellis et al. (1984) used the Velten (1968) procedure to induce

either depressed or neutral moods in college students, then presented them with incidental learning tasks. In the learning phase of Experiment 3 subjects were exposed to a sentence with a missing noun and asked to choose one of two subsequently presented nouns to complete it. Across the 24 learning trials the sentences varied according to the difficulty of completion, as assessed by independent ratings. (A less difficult frame for the word "dream" was "The girl was awakened by her frightening _____ ," compared to the more difficult frame, "The man was alarmed by the frightening _____.") Subsequently, experimentally depressed subjects recalled fewer words from the difficult sentence frames than did nondepressed subjects. This finding was interpreted to mean that depression reduces task-relevant processing, especially in situations that require greater degrees of cognitive effort.

Although the materials used by Ellis et al. had been used in experiments demonstrating that greater degrees of cognitive effort were expended during processing in more difficult contexts (Tyler, Hertel, McCallum, & Ellis, 1979), their procedure did not include such measures. And so it is not clear if depressed subjects actually failed to expend the effort required by the difficult sentences (compared to the easy sentences). In Experiment 1 we investigated possible differences in the allocation of attentional resources by including attentional probes during the learning trials.

In brief, we selected subjects who had scored lower than 5 or higher than 9 on two administrations of the Beck Depression Inventory (BDI; Beck, Ward, Mendelson, Mock, & Erbaugh, 1961) 2–4 weeks apart. Our depressed subjects were interviewed to ensure that their elevated BDI scores reflected depression rather than other forms of psychopathology or illness. Most subjects met Research Diagnostic Criteria (RDC) for major or minor depression (Spitzer, Endicott, & Robins, 1978). The learning task conformed to the methods of Tyler et al. (1979, Experiment 4), but was computer-implemented. It consisted of a series of trials on which the presentation of a single word was followed by an incomplete sentence; during 80% of these trials a brief tone was presented over earphones. The subject's primary task was to decide if the word fit sensibly into the sentence, and the secondary task was to press a button in response to the tone. The secondary tone-detection task served as a probe for the amount of available resources (see Kahneman, 1973). Longer latencies to respond to the tone are offered as indications that fewer resources are available for tone detection; one assumes that they are either depleted by some extra-experimental factor, such as depression, or that they are allocated in greater proportion to the primary task.

Method

Materials and Apparatus

Learning Task. The learning task consisted of 40 trials of sentence completion. Each trial required the subject to view a word followed by a sentence with one word missing and to decide if the word fit sensibly into that sentence frame. Unknown to the subject at the outset, all words fit sensibly into the corresponding frames. The 40 words were nouns high in frequency, concreteness, and meaningfulness; their accompanying sentence frames had been pretested for difficulty of completion (see Hertel, 1989). Each word appeared with easy frames for approximately half of the subjects in each condition of depression and with difficult frames for the others. For example the word *artist* was sometimes followed by "The young man's portrait was painted by the _____." Other subjects saw *artist* followed by "The young man's physique was admired by the _____."

The learning task was implemented on an TRS-80 computer. The program began with the presentation of four practice trials that represented each level of difficulty and tone delay. All trials began with the presentation of a word at the top of the screen for 1 s. Its offset occurred simultaneously with the onset of the sentence frame, which was centered on the screen and remained exposed for 8 s. On 80% of the trials a weak but detectable tone occurred at delays of 1, 2, 3, or 4 s after the onset of the frame. With the offset of the sentence a question mark appeared and the trial was terminated by a key press. A blank screen lasting 1 s separated trials.

The 40 words were grouped into blocks of 10. Each block contained four 5-letter words, four 6-letter words, one 7-letter word, and one 8-letter word. Across subjects, words were maintained in these blocks as they rotated through the 10 conditions obtained by crossing sentence difficulty (easy vs. difficult) with tone delay (no tone, or a delay of 1, 2, 3, or 4 s after the onset of the sentence). The order of tone delays was constant for all subjects.

Depression Inventory. Selection of subjects was based on their scores on the BDI. The BDI consists of 21 items describing cognitive, affective, and somatic symptoms of depression. Higher scores indicate depression.

Subjects and Design of Mood Conditions

Sixty subjects were recruited from a pool of approximately 200 volunteers who filled out the BDI in their lower-division psychology classes at Trinity University. Those volunteers who scored lower than 5 and higher

than 9 were contacted again and asked to volunteer for an experiment on word processing. All volunteers received extra credit in their courses in exchange for participation.

Upon completion of the experimental tasks (2--4 weeks later), subjects filled out the BDI a second time. Those whose scores had shifted out of the range of the group in which they had started were excluded from data analyses. Subjects who remained in the high-scoring group were interviewed in order to determine whether they met RDC for depression. Approximately one-third of these interviews were conducted by a graduate student in psychology who was trained and supervised by a Ph.D.-level psychologist. The supervising psychologist conducted the remaining interviews and determined the diagnostic category from notes taken on the subjects' responses.

Of the 40 subjects who began the experiment with high (depressed) scores on the BDI, 17 remained for data analysis. Nineteen did not score in the depressed range on the second administration, and four more subjects were determined not to be depressed on the basis of the interview. Twelve of the remaining 17 subjects met full RDC for either major or minor depression. Five met all but one of the criteria[3] clinician was satisfied that the subjects' BDI scores reflected depression rather than illness or another form of psychopathology. Of the 20 subjects who began the experiment with low scores on the BDI, three shifted out of that range and were therefore excluded.

Procedure

The experimenter began by reading the instructions for the learning task. Subjects were informed that two different tasks would be performed simultaneously: a primary sentence-completion task and a secondary task of detecting an auditory signal. In the primary task subjects were instructed to decide if the word presented at the top of the screen fit sensibly into the incomplete sentence presented below it. They were told to wait until the question mark appeared at the end of each trial and then repeat the word and the decision (by saying *yes* or *no*). They were also informed that most trials would require a *yes* response, but that some subjects would experience a few negative instances; therefore, they should consider all decisions carefully. For the secondary task they were asked

[3]Most of these subjects did not clearly meet the inclusion criterion of having either sought help for their depression or suffered marked impairment in work or social functioning. We counted this criterion as being met only when the subject either had sought formal counseling or had experienced severe and unmistakable problems in functioning, such as being in danger of failing school or of being fired from a job due to depressive impairments.

Table 1. Mean Latencies (in Seconds) and Proportion Recalled (Experiment 1)[a]

| | Response measure | | | |
| | Latency | | Recall | |
Mood	Easy	Difficult	Easy	Difficult
Nondepressed	0.54	0.62	.20	.28
Depressed	0.64	0.68	.30	.37

[a]Note. n = 17. Median latencies were averaged across subjects to produce mean latencies. Recall was measured as the proportion of words repeated during the learning task that were subsequently recalled.

to respond to the tone by pressing the button on a switch held in the nondominant hand. Instructions emphasized pressing as quickly as possible in response to the tone, but without sacrificing attention to the primary task.

The experimenter answered any questions, fit the subjects with earphones (for tone presentation), and began the program. The four practice trials preceding the learning trials provided an opportunity for further clarification.

After the last learning trial, subjects worked multiplication problems for 5 min, when the experimenter announced (for the first time) the test of free recall. Subjects were asked to write down all the words they could remember from the learning phase. When the subjects indicated that they were finished they were asked to try harder to recall and to write at least 20 words, guessing if necessary.

After the experiment was completed, we obtained subjects' verbal Scholastic Aptitude Test (SAT) scores. The scores were available for 12 of the 17 subjects in each condition of mood. They were obtained from university records in a manner than protected subjects' confidentiality.

Results and Discussion

BDI and SAT Scores

Mean scores on the BDI administered at the end of the session were 1.24 for nondepressed subjects and 16.47 for depressed subjects. SAT verbal scores did not reliably differ according to mood group; the mean was 558 for nondepressed subjects and 539 for depressed subjects.

Latencies

The median latency (in centiseconds, or cs) to respond to the tone was submitted to an analysis of variance, with a between-subjects factor for mood group (depressed vs. nondepressed) and a within-subjects factor for the difficulty of the sentence frame (easy vs. difficult). (For this and subsequent analyses the significance level was set at .05.) The mean for each condition is shown in the left half of Table I.

Median latencies were reliably longer when the tone occurred during the more difficult sentence frames, $F(1, 32) = 10.66$, $MSe = 107.050$. This reliable main effect replicated previous results with these and similar materials (Hertel, 1989; Tyler et al., 1979). Median latencies were longer in the depressed group than in the nondepressed group, $F(1, 32) = 4.44$, $MSe = 519.596$. Finally, the interaction of mood group and difficulty was not reliable ($p > .20$).

The association of longer latencies with depression could be interpreted in at least three ways. First, depressed subjects might experience motor retardation. Second, while processing the materials in the primary task, they might be engaged in depressive ruminations which would occupy further attention and delay their responses to the tone. Third, they might be processing the materials in the primary task in a more effortfully elaborative or distinctive fashion, with no time for depressive thoughts. If the last account is viable, then we should expect performance on the subsequent test of free recall to profit from the more effortful processing.

Recall

The proportion of words recalled out of those that were correctly repeated during the learning trials served as the dependent variable. (The very low number of errors in repeating the words did not reliably differ according to mood.) Proportions were submitted to an analysis of variance, with factors for mood group and sentence difficulty.

As expected, words were more likely to be recalled if they were processed in more difficult sentence frames, $F(1, 32) = 10.04$, $MSe = .021$. And, consistent with the latency data, depressed subjects recalled more words than did nondepressed subjects, $F(1, 32) = 4.46$, $MSe = .061$. This unexpected advantage for depressive recall did not interact with sentence difficulty; in contrast to the findings of Ellis et al. (1984), the effect of sentence difficulty obtained in both groups.

The depressive advantage in free recall probably cannot be attributed to higher verbal abilities among our depressed subjects, because they

tended to score slightly lower than the nondepressed subjects on the SAT. We have not, in fact, developed a *post hoc* explanation for the depressive advantage, believing that it requires replication before such speculations are worthwhile (see Hasher, Rose, Zacks, Sanft, & Doren, 1985, for a similar finding and interpretation). Instead, it is important to note that effort effects in free recall and response latencies were obtained for nondepressed and depressed subjects alike — that a deficit in the recall of words from the more difficult contexts was clearly not experienced by naturally depressed students as it had been by students induced to feel depressed in the experiment by Ellis et al. (1984).

EXPERIMENT 2

The results of Experiment 1 showed that, regardless of depression, processing words in the more difficult sentences required more capacity and produced better memory. In short, we found no evidence of memory impairment for depressed individuals and failed to replicate the results of Ellis et al. (1984). What might account for the differences in findings? Our methods differed from those of Ellis et al. in a number of ways. In Experiment 2, we attempted to bring the major differences into the context of one study so that they could be examined systematically. We now describe each type of difference and how its investigation was incorporated into the design of Experiment 2.

In the first place, we employed diagnostic criteria and self-reported indices to identify subjects with naturally occurring depression, whereas Ellis et al. induced depressive moods. When they work, experimental inductions clearly produce a variety of effects on the individual, some of which seem to mimic clinical depression. But as several reviewers (e.g., Blaney, 1986; Hasher et al., 1985) have noted, the use of induced moods to model naturally occurring depression carries with it a number of potential problems. Our intent was to address these issues by directly comparing the levels of recall associated with experimental and natural depression. To that end, a major aspect of our design was the concurrent manipulation and selection of depressed moods. Three groups of subjects participated: nondepressed subjects in a neutral mood, nondepressed subjects induced to feel depressed, and naturally depressed subjects.

In the second place, the experimental procedures of our first experiment and of Ellis et al. (1984) differed in the extent to which attention to the words was required by the learning task. Ellis et al. presented sentence frames and word pairs successively for 4 s each and allowed subjects to choose the correct alternative at any point during presentation of the word

pair. In contrast, we presented a word for 1 s only and followed it with the sentence frame; after 8 s the frame disappeared, and the subjects were required to report the word and the decision about its fit. Thus, our procedure implicitly required that all subjects rehearse the word for the duration of the trial, whereas the subjects in Ellis et al.'s experiment could engage in further processing after the choice between words was made or allow their minds to wander.

In the design of Experiment 2, we addressed the role of these differences by manipulating the structure of the learning task. We repeated the methods of Experiment 1 in the structured condition and loosened the demands in the unstructured condition. Subjects in the unstructured condition were not required to maintain the target word in memory for the 8 s of the trial; the word remained on the screen and the subjects reported their decisions at any time during the display of the word and sentence frame.

Finally, we modified our use of the tone-detection task. In addition to presenting probes on 80% of the learning trials, we included 15 baseline trials of tone presentations prior to the learning trials, as well as another 15 trials afterward. The baseline trials were used to evaluate preexisting differences in response latencies (as might be incurred by task-irrelevant thinking or by motor retardation on the part of depressed subjects). Moreover, when baseline latencies are used as covariates, latencies to respond to secondary-task tones during the learning trials should be more clearly indicative of task-related efforts.

In summary, the design of Experiment 2 allowed us to compare findings obtained through the experimental induction of a depressed mood with those obtained through the selection of natural depression. The design also included a manipulation of task structure as a way of determining the boundary conditions of depressive deficits. Finally, additional measures of cognitive effort were taken in further tests of the cognitive-effort account.

Method

Materials and Apparatus

Group Definition and Mood-Induction Procedures. As in Experiment 1, we used the BDI to select subjects and categorize them as nondepressed or depressed. Neutral, depressed, and elated moods were produced with the aid of the Velten mood-induction procedure (Velten, 1968), a series of statements that are read silently and aloud by the subject in isolation. The subject is asked to try to adopt the mood that is represented by the statements. In the case of the neutral induction, the 60 statements are

(somewhat boring) facts. In the case of the depression induction, the 60 statements, many of which are written in the first person, become increasingly negative. The elation induction, administered to experimentally depressed subjects at the end of the session, contains statements that are increasingly positive about the self.

In previous experiments (Ellis, Thomas, McFarland, & Lane, 1985; Hertel & Hardin, 1990), depressed moods have been successfully established by 30 statements (from the original 60 employed by Velten). We employed the shorter version in this study and eliminated statements that included any of our experimental words or referred to difficulties in remembering (to reduce demands for compliance).

The success of the experimental induction was assessed by asking subjects to complete the Depression Adjective Check List (DACL; Lubin, 1965), a list of positive- and negative-mood adjectives that are checked for self description. A depression score is obtained by summing the endorsed negative adjectives and the nonendorsed positive adjectives. Also, at the very end of the session we administered a questionnaire on which we asked subjects to speculate about the nature of their mood during the experiment, how long it had lasted, and whether they had described their mood accurately on the DACL or had told us what they thought we wanted to find out.

Learning Task. The materials for the learning task were those used in Experiment 1. However, we made several changes in the task itself: (a) The task was programmed for an IBM-XT that presented tones without the use of headphones. The tones occurred again on 80% of the trials, at delays of 1, 2, 4, or 6 s following the onset of the sentence frame. (b) The program inserted 15 trials of tones presented alone, prior to the 40 learning trials, and another 15 baseline trials after the learning trials. (c) In the structured condition the word disappeared after 1 s (when the sentence appeared), but in the unstructured condition the word remained exposed for 9 s and offset with the sentence frame.

Subjects and Design of Mood Conditions

Selection. A total of 146 Trinity University students completed the experiment; 117 volunteered in exchange for extra credit in their lower-division psychology classes. When we recruited subjects in this way, the BDI was administered to large groups of volunteers in classroom settings. Those volunteers who scored 5 or below and 9 or above were contacted again and asked to volunteer for an experiment on word processing. At the start of the individual sessions we administered the BDI a second time to subjects whose first scores were 5 or below. If their second scores fell

in the same range, we randomly assigned them to receive a neutral- or depressed-mood induction (Velten, 1968). We dismissed the subjects who were preselected as nondepressed but whose second scores fell beyond the cutoff.

We obtained the second BDI scores of the subjects who initially scored 9 or above somewhat differently, because we suspect that naturally depressed subjects deflate their scores when they report to experimenters in individual sessions. Subjects who initially scored 9 or above filled out the second BDI at the end of the experimental session, together with an interview form soliciting further information regarding the context of their depressed mood. Having been told that these forms were being collected for a different study, the naturally depressed subjects delivered them in a sealed envelope to the department secretary. The data from these subjects were not used if their second BDI scores or responses to the form did not meet the specifications. These procedures have been used by Hertel and Hardin (1990) and conform to Deardorff and Funabiki's (1985) specifications for identifying naturally depressed subjects.

In the second procedure for recruiting subjects, an additional 29 subjects were paid $5.00 each for participating. They were obtained through summer-school classes in disciplines other than psychology and were therefore not preselected on the basis of a first BDI. These subjects all filled out the BDI at the beginning of the experimental session. Those who scored 5 or below were randomly assigned to a mood-induction procedure; those who scored 9 or above received neutral-mood inductions and were given the interview form at the end of the session. These subjects were evenly distributed across the experimental groups.

Attrition. We discarded the data from 18 subjects whom we initially thought to be naturally depressed, either because their second score on the BDI was lower than 9 or because they indicated on the interview form that they had not been depressed for much of the previous 2 weeks. Some subjects, for example, revealed that they had not slept well for the past few nights due to class assignments; lack of sleep could elevate scores on the BDI but not produce other symptoms of depression.

We also eliminated the data from three nondepressed subjects who received the neutral-mood induction but produced high DACL scores (in the 20s), indicating that in spite of their low BDI scores they were experiencing a very negative mood. And finally the data from five experimentally depressed subjects were discarded on the basis of very low DACL scores *and* a subsequent phone interview in which they denied having been induced to feel depressed or lethargic.

The Final Sample. Data from 120 subjects remained to be analyzed; 34 subjects were male (the number in each condition ranged from 5 to 7).

The 40 naturally depressed subjects all scored 9 or above on the in-session BDI and reported on the interview form having been depressed for the past 2 weeks. The 40 experimentally depressed subjects each reported in a telephone interview that the induction procedure produced either depression or lethargy (or their DACL scores were in the range of others' scores).

The assignment of subjects to learning tasks and mood inductions was random within blocks determined by initial BDI cutoff and gender (with the exception that data from ineligible subjects were replaced when ineligibility was determined). Within each between-subjects condition, two subject were assigned to each of the 10 counterbalancing conditions of the learning task.

Procedure

Subjects who had scored 5 or below on the first administration of the BDI (and all subjects who had not been pretested) filled out the BDI at the start of the session, sealed it in an envelope, and gave it to the experimenter, who left the room ostensibly to prepare the next task but instead scored the BDI. If the score was 5 or below she assigned the subject to a treatment condition (neutral or depressed induction and unstructured or structured orienting task.) If the BDI score fell between 5 and 9 the subject was told that he or she was part of a control group and was dismissed. If the score on the first administration was 9 or above the subject was assigned to receive a neutral induction (so that these subjects would be treated like the nondepressed subjects). The induction procedure lasted approximately 15 min and was followed by the DACL, to assess the subject's mood.

The instructions and procedure for the learning and recall tasks conformed to those used in Experiment 1, with the following exceptions: (a) Some instructions were presented on the computer screen. (b) Fifteen baseline trials of tone detection occurred before and after the 40 learning trials. (c) Subjects were asked to press any key (with a finger on the nondominant hand) in response to the tone. (d) The subjects assigned to the unstructured condition of the learning task were told to report their decision about the fit of the word as soon as they made it, whereas subjects in the structured condition were required to wait until the question mark appeared to report the decision and repeat the word. Word repetition was not required in the unstructured task.

A second administration of the DACL followed the recall test. Then experimentally depressed subjects underwent an induction of elation. All subjects completed the final questionnaire as part of their debriefing. At

Table II. Mean Scores on Mood Indices (Experiment 2)[a]

Measure	Mood		
	Nondepressed	Experimental	Naturally depressed
BDI[b]	2.2	2.1	15.2
DACL[b]			
First	6.6	15.1	11.8
Second	8.0	9.9	15.0
Final questionnaire			
Sadness	2.0	3.6	3.2
Lethargy	3.5	4.6	4.8
Distraction			
Unstructured	3.5	4.4	5.2
Structured	4.1	4.0	4.1

[a]Note. Means were computed on scores from 40 subjects, with the following exceptions. DACL second administration for nondepressed: $n = 34$. Distraction means: $n = 20$.
[b]BDI = Beck Depression Inventory; DACL = Depression Adjective Check List.

the end of the session, naturally depressed subjects were given the BDI (if they had not taken it at the start of the session) and interview forms to fill out and turn in to the department secretary. Experimentally depressed subjects were carefully quizzed about their moods before dismissal; they were also phoned on a subsequent day and asked if the depressed induction had worked and if it had lasted beyond the session. (All subjects reported that they left the session in a reasonably good mood.)

Results and Discussion

Mood Indices and SAT

BDI and DACL Scores. The naturally depressed students showed moderate levels of depression, on the average (mean BDI = 15.2, vs 2.2 for the nondepressed students). Naturally depressed male students tended to report lower levels of depression than did female students ($M = 11.2$ vs 16.8, respectively).

Table II presents the mean DACL scores. As can be seen in the table, the naturally depressed subjects began the learning task in a slightly better mood than did the experimentally depressed subjects. Yet by the time of recall the difference was reversed, the naturally depressed subjects reported more depressed feelings than did the other two groups, who reported similar feeling states. An analysis of variance performed on the DACL scores revealed a reliable interaction of mood group with the time of ad-

ministration (before the learning task versus after recall), F (2, 108) = 31.71, MSe = 12.253. (Note that the second DACL was not administered to 6 nondepressed subjects, due to experimenter error.)

We also found a reliable three-way interaction of gender, mood group, and time of administration, F (2, 108) = 3.35, MSe = 11.994. The interaction of depression with time of administration described above was exaggerated for male students, but still obtained for female students.

Final Questionnaire. A section of this form asked subjects to characterize their mood during the computer task and presented a list of adjectives, each of which was followed by a rating continuum anchored by *not at all* (1) and *extremely* (7). Ratings for *happy* (after reversing the direction), *sad*, and *blue* were averaged to constitute the dimension of sadness. Ratings for *tired, lethargic*, and *sleepy* were averaged to constitute a measure of lethargy. And distraction was made up of ratings for *distracted* and *Did your mind wander?* Means on these three dimensions are presented in Table II.

Mood-dimension scores were submitted to analyses of variance, with factors for mood group and learning task. Scores on the sadness dimension differed according to depression, F (2, 112) = 14.39, MSe = 1.822; the two depressed groups did not differ reliably and both indicated more sadness than did the nondepressed group. Similarly, a reliable main effect of mood group on lethargy scores indicated that the two depressed groups reported feeling more tired than did the nondepressed group, F (2, 113) = 7.77, MSe = 2.286.

In Table II, mean distraction scores are broken down according to learning task (unstructured vs. structured) in order to reveal the nature of the reliable interaction of task structure with mood group, F (2, 114) = 3.27, MSe = 1.971. The two depressed groups reported feeling more distracted than did nondepressed subjects, but only if they had received the unstructured learning task.

The final questionnaire also inquired about the duration of the mood that was characterized on the adjective dimensions. The rating scale ranged from *stopped immediately* (1) to *still feel this way* (7). A main effect of mood group revealed that the moods of experimentally depressed subjects (M = 3.4) changed more than the moods of the nondepressed subjects (M = 4.9) and the naturally depressed subjects (M = 4.8), F (2, 114) = 10.66, MSe = 2.589.

Finally, we attempted to assess demand characteristics of the DACL by asking subjects to rate the extent to which they had felt the same as they had indicated on the DACL (*felt the same* = 1; *felt very different* = 7), rather than complied with what they thought the experimenter wanted. No reliable differences between groups were revealed; M = 1.5.

Table III. Mean Latencies (in Seconds Adjusted by Baseline Latencies (Experiment 2)[a]

Mood	Learning task			
	Unstructured		Structured	
	Easy	Difficult	Easy	Difficult
Nondepressed	.47	.52	.51	.59
Experimentally depressed	.48	.52	.46	.54
Naturally depressed	.46	.50	.55	.61

[a]Note. n = 20.

SAT Scores. We procured the scores on the verbal component of the SAT for 110 of the subjects (18 or 19 in each experimental condition). SAT scores did not differ according to experimental condition, nor were they reliably related to BDI scores, total recall, or the difference in recall between difficult and easy trials.

Latencies

Baseline Trials. The median time in milliseconds (ms) to respond to the tone during the baseline trials was submitted to an analysis of variance, with between-subjects factors for the structure of the learning task (unstructured vs. structured) and mood (nondepressed vs. experimental vs. natural), and a within-subjects factor for when the trials occurred (before vs. after the learning task). The only reliable effect indicated that all subjects responded more quickly on the 15 trials at the end of the learning task than they did on the 15 trials at the beginning ($M = 343$ ms vs. 358), $F (1, 114) = 5.71$, $MSe = 2377.94$. Evidently, these experimentally and naturally depressed college students did not experience motor retardation or other limitations that would delay responses to the tones.

Learning Trials. The median time in milliseconds to respond during the learning trials was submitted to an analysis of covariance, with factors for task structure, mood (between-subjects), and sentence difficulty (within-subjects). The average of the baseline medians (pre- and postlearning trials) served as the covariate for analyses of between-subjects effects. As expected, baseline latencies were reliably correlated with learning-trial latencies, $R = .67$, $F (1, 113) = 93.38$, $MSe = 33755.87$. Second, longer latencies were observed in the structured task, $F (1, 113) = 5.31$, $MSe = 33755.87$. Third, longer latencies occurred during the difficult sentences, $F (1, 114) = 32.43$, $MSe = 5301.48$. No reliable differences were associated with mood. For each condition the mean of the median latencies, adjusted by baseline latencies, is presented in Table III.

Table IV. Mean Number of Words Recalled (Experiment 2)[a]

	Learning task			
	Unstructured		Structured	
Mood	Easy	Difficult	Easy	Difficult
Nondepressed	3.0	5.5	4.6	5.6
Experimentally depressed	3.3	4.4	3.8	4.7
Naturally depressed	3.3	4.8	3.2	5.8

[a]Note. n = 20.

Additional analyses of the latency data included factors for gender and for the position of the tone during the trial (1 or 2 s vs. 4 or 6 s after the onset of the sentence). These factors did not reliably interact with other factors in the design. (The factor for gender also failed to produce reliable effects in the analysis of recall data and therefore was excluded in the results reported below.)

Recall

The number of words correctly recalled was submitted to an analysis of variance with between-subjects factors for mood and the structure of the learning task, and a within-subjects factor for the difficulty of the sentence frames. A reliable main effect of task structure showed that requiring rehearsal of the words improved overall recall, $F(1, 114) = 5.01$, $MSe = 3.958$. Also, a reliable main effect of difficulty replicated the usual finding that words evaluated in the more difficult sentences were better recalled, $F(1, 114) = 59.09$, $MSe = 2.519$.

These main effects, however, were attenuated by the reliable three-way interaction of difficulty, structure, and mood group, $F(2, 114) = 3.35$, $MSe = 2.519$. Table IV presents the mean number of words recalled in each condition. We evaluated this interaction through a series of planned comparisons among the mood groups, in order to reveal the conditions under which the findings of Ellis et al. (1984, Experiment 3) and those of Experiment 1 in this report were replicated.

First, we examined comparisons between nondepressed and experimentally depressed subjects. [In contrast to neutral-mood subjects, Ellis et al.'s (1984) experimentally depressed subjects recalled fewer words and did not show the effect of sentence difficulty.] Our analyses revealed that an induction of depression caused an overall deficit in recall, $F(1, 114) = 3.95$, $MSe = 3.958$. Within the unstructured condition (the condition that was roughly comparable with Ellis et al.), the effect of sentence difficulty

tended to be greater for nondepressed subjects than for experimentally depressed subjects, as indicated by a marginally reliable simple interaction, F (1, 114) = 3.62, MSe = 2.519, $p < .06$. Finally, experimentally depressed subjects in our structured task tended to recall fewer words overall than did the nondepressed subjects, F (1, 114) = 3.87, MSe = 3.958, $p < .06$. No other effects involving the comparison of these two groups came close to reaching significance. We conclude that the results of Ellis et al. (1984) were essentially replicated with experimentally depressed subjects.

Second, we examined the difference in recall between nondepressed and naturally depressed subjects. An overall deficit was not reliably obtained, but the difference between the two groups did reliably interact with sentence difficulty and task structure, F (1, 114) = 6.71, MSe = 2.519. Within the unstructured task, no reliable main effect or interaction was associated with natural depression. However, within the structured task, naturally depressed subjects recalled fewer words from *easy* sentences but similar numbers of words from difficult sentences, as shown by the reliable interaction of sentence difficulty with the comparison of nondepressed vs. natural depression, F (1, 114) = 4.77, MSe = 2.519. The mechanisms that produced this interaction are unclear to us, but the important outcome was the lack of depressive deficit in recalling words from difficult sentences. Thus, although the results of Experiment 2 failed to replicate the depressive advantage found in Experiment 1, they also failed to produce evidence of a reliable deficit associated with natural depression.

Third, in order to examine the basis for concluding that experimentally and naturally depressed subjects might perform differently, comparisons between those two conditions were made. They revealed a marginally reliable interaction of this comparison with sentence difficulty, F (1, 114) = 3.58, MSe = 2.519, $p < .06$. The effect of difficulty was larger for the naturally depressed subjects than for the induced subjects. Furthermore, this effect clearly reached significance in the structured task, F (1, 114) = 5.08, MSe = 2.519.

EXPERIMENT 3

In the unstructured condition of Experiment 2, naturally depressed subjects appeared to show a tendency toward lower recall of words from difficult sentences, compared to their nondepressed counterparts. This tendency did not approach statistical significance; nevertheless we were dissatisfied with a conclusion that emphasized the failure to find a naturally depressed deficit while performance in that condition seemed relatively low. We also realized that our method for the unstructured condition yet departed from the methods of Ellis et al. (1984) in that ours included a secondary task and theirs did not. The secondary task might mobilize the

attention that depressed subjects might otherwise allocate in a task-ir-relevant manner (see Krames & McDonald, 1985; Williams, Watts, Mac-Leod, & Mathews, 1988). In order to address these concerns, Experiment 3 was conducted.

In Experiment 3 nondepressed and naturally depressed subjects participated in a learning phase that was identical to the unstructured task in Experiment 2, with the exception that the secondary task was eliminated. Other departures from the method of Experiment 2 included the elimination of the mood-induction procedure and a modification of the recall task. When subjects indicated they were finished recalling, the experimenter drew a line below the last word written and requested additional effort in recall until the subjects had written at least 20 words. The change from Experiment 2 was the drawing of the line. This change was implemented so that we could evaluate the level of recall before and after the request to continue was made. Ellis et al. (1984, Experiment 3) had encouraged guessing but not required it. Perhaps naturally depressed subjects recall at higher levels only when they are required to guess. (See Dunbar & Lish-man, 1984, for evidence of a conservative bias in depression.)

Method

Subjects

As in Experiments 1 and 2, 25 subjects were selected on the basis of their scores from a first administration of the BDI, obtained approximately 10 days before the experimental session. Those who scored 9 or above were assigned to the depressed group and those who scored 5 or below to the nondepressed group. The data from five subjects were discarded on the basis of the second BDI score and interview form, both of which were obtained at the end of the session. Two initially nondepressed subjects had moved beyond the cutoff for the BDI and produced DACL scores above 11. Three initially depressed subjects produced second BDI scores that fell below 9 or claimed not to be depressed on the interview form. Of the remaining 10 subjects in each condition of mood, two were men.

Materials and Procedure

Immediately after informed consent, subjects participated in the learning task. One subject in each condition of mood was assigned to each of the 10 counterbalancing conditions (related to word order and to the assignment of words to level of sentence difficulty). The materials and proce-

dure for the learning task conformed to those for the unstructured task of Experiment 2, but the secondary task was eliminated and no tones occurred.

After the 5-min distraction task (arithmetic problems), the experimenter requested free recall of the words. When the subjects reported that they were finished, she drew a line below the last word and requested that they try to recall at least 20 words, by guessing if necessary. She informed them that guessing often pays off in a correct response.

The recall task was followed by an administration of the DACL and a shortened version of the final questionnaire from Experiment 2. Last, subjects were asked to fill out the BDI and the interview form, seal them in an envelope, and deliver them to the departmental office. The envelope was marked with a subject-identification number, but the forms were anonymous.

Results and Discussion

Mood Indices

The mean BDI score on second administration was 1.0 for nondepressed subjects and 17.0 for depressed subjects. The corresponding mean DACL scores were 5.6 and 10.8, respectively, and significantly differed, F (1, 18) = 5.22, MSe = 25.889. On the final questionnaire, depressed subjects reported feeling more distracted than nondepressed subjects (M = 3.8 and 1.6, respectively), F (1, 17) = 8.60, MSe = 2.385. Depressed subjects also felt more lethargic (M = 3.2) than did nondepressed subjects (M = 2.4), F (1, 17) = 6.44, MSe = .652. (No apparent gender differences were obtained for measures of mood or recall.)

Recall

The number of words recalled was submitted to an analysis of variance with a between-subjects factor for mood and within-subjects factors for sentence difficulty and type of recall (free vs. forced). Table V shows the mean for each cell in the design. The analysis revealed an advantage of words processed in difficult sentences, F (1, 18) = 5.40, MSe = 3.003, but this effect of sentence difficulty depended on the type of recall, F (1, 18) = 8.39, MSe = 1.525. The effect of difficulty was greater under free recall conditions (i.e., before guessing was required), perhaps because few numbers of words were produced by the forcing procedure. Also, in a few cases free recall levels neared the cutoff of 20 words, such that few words had to be forced in order for these subjects to reach the final criterion. Obviously, subjects recalled more words under conditions of free recall, F (1, 18) = 19.93, MSe = 6.525.

Table V. Mean Number of Words Recalled (Experiment 3)[a]

Mood	Type of recall					
	Free		Forced		Total	
	Easy	Difficult	Easy	Difficult	Easy	Difficult
Nondepressed	3.5	4.7	1.5	1.6	5.0	6.3
Depressed	2.8	5.0	1.3	1.4	4.1	6.4

[a]Note. $n = 10$.

No reliable differences according to mood were obtained. Even the apparent interaction of mood with difficulty did not approach significance ($p > .50$). These results confirmed that under conditions very similar to those employed by Ellis et al. (1984, Experiment 3) no depressive deficit obtained. Although the secondary task may have mobilized attention and prevented depressive musings in the first two experiments, removing that task did not produce a depressive deficit in subsequent recall.

GENERAL DISCUSSION

The results from these experiments are among the first to validate concerns about the adequacy of experimental inductions as models of naturally occurring depression in the realm of memory research (see Ingram, 1989; Riskind, 1989). Specifically, the data do not support an extension of the cognitive-effort account to the domain of recall by naturally depressed students. In the ensuing discussion we first review the effects of task difficulty. We then propose possible explanations for the lack of correspondence between the effects of experimental depression and the differences associated with natural depression. Last, we summarize the relations of our findings to the cognitive-effort account of depressive deficits in remembering.

Task Difficulty

Performance on the secondary tone-detection task in Experiments 1 and 2 showed that attentional resources of depressed and nondepressed subjects alike were allocated in greater proportion to the words evaluated in the difficult sentences, compared to the easy sentences. And, as has been the case in other experiments that used these materials and tasks (Hertel, 1989; Tyler et al., 1979), words from the difficult sentences were better

recalled by all mood groups. Compared to the easy sentence frames, the difficult sentence frames seemed to have invited more effortful procedures which, in turn, produced better recall (see McDaniel, Einstein, & Lollis, 1988; Mitchell & Hunt, 1989).

In Experiment 3 we took an additional step in bringing the conditions of the learning task in line with those of Ellis et al. (1984) by removing the secondary task, which might have mobilized attention and stimulated more task-related effort from naturally depressed subjects. Yet effects of task difficulty on recall were obtained for both mood groups, without evidence of a depressive deficit. We have not taken the last step, which is to alter the learning task so that subjects choose between two words, as did Ellis et al., but we cannot imagine why deficits would obtain following choice procedures and not verification.

Effects of Experimental Depression

One constraint on the general finding of effects of task difficulty on recall was provided by experimentally depressed subjects in Experiment 2. These subjects did not produce as large an effect as the ones obtained for nondepressed and naturally depressed subjects. Why would subjects who experience a transitory depressed mood perform differently from naturally depressed subjects?

One possible reason why lower levels of performance might be caused by inductions is related to the severity of the depressed mood (see Johnson & Magaro, 1987). It might be argued that the induction technique produces a more extreme depressive state than what is experienced by naturally depressed students (see Ellis & Ashbrook, 1988). The scores on the DACL in Experiment 2 do suggest that depression-induced subjects had more negative feelings immediately following the induction than did naturally depressed subjects. But because that difference was not statistically reliable and because the moods of the naturally depressed subjects were *more* negative at the time of recall, the poor performance of experimentally depressed subjects probably cannot be attributed to a more severe mood impairment.

Another explanation for differences between the two types of depressive groups is related to possible differences in prior experience in mood management. As argued by Rude and Hertel (1987), mood inductions are very recent experiences, whereas a depressed mood characterized at least 2 weeks of experience for our naturally depressed subjects. Naturally depressed college students may develop coping strategies to aid their performance in cognitive tasks, whereas experimentally depressed subjects are less likely to benefit from prior experience in managing their mood, espe-

cially if the mood is particularly acute. Experimental techniques perhaps induce moods that are more acute (if not more severe) than moods that are experienced naturally (Ingram; personal communication, July 13, 1990).

Next, following Rude and Hertel (1987), we considered motivational factors that might have influenced the performance of experimentally depressed subjects. A focus on possible demand characteristics of the induction procedure suggests that experimentally depressed subjects might believe that we expect poor performance from them. Why else would we ask them to feel depressed and then administer a learning task? This argument is pertinent even when the induction procedure seems to have established an alteration in mood. The point is that when a subject is faced with the choice between continuing serious attempts to recall vs. writing down enough words to total 20, experimentally depressed subjects can opt for the latter without any loss of self-esteem. They can blame their poor performance on our induction procedure. Naturally depressed subjects, in contrast, have no excuse for poor performance.

In short, whenever differences in attributions, differences in prior experience in mood management, or both are likely to affect performance, induction procedures may provide poor models of natural depression. Ingram (1989) has pointed to additional factors that could encourage faulty inferences about the modelling of depressive memory through laboratory manipulations. They include the possibility that experimental techniques produce moods that lack associated features of depression. We suggest that the associated features could, in some contexts, ameliorate depressive deficits.

A few final comments about our methods of assigning subjects to mood groups are in order. First, we used a relatively low cutoff point on the BDI for selecting subjects for the nondepressed control groups. Kendall, Hollon, Beck, Hammen, and Ingram (1987) advised that such a method does not permit generalization to "normal" populations, given that very low scores on the BDI might indicate abnormal cheerfulness or even some forms of psychopathology. Because we did not interview nondepressed subjects we cannot rule out the presence of hypomania, psychopathy, or extreme cheerfulness. A related limitation is that we did not formally assess the reliability of our RDC decisions concerning the depressed subjects. However, we took a relatively conservative approach to counting subjects' responses as meeting criteria for depression. This approach, together with the procedure of repeated BDI assessments, enhanced our confidence that our naturally depressed subjects were indeed depressed.

The Cognitive-Effort Account

The cognitive-effort account of depressive deficits in recall implies a pervasive allocation of cognitive resources to depression. How does this interpretation fare in accounting for the present results on naturally occurring depression? In Experiment 1, depressed subjects took longer to respond to the tone *and* recalled higher percentages of words than did nondepressed subjects. Although we did not replicate this depressive advantage in the subsequent experiments, it was an important finding. The similar patterns in latencies and recall suggested that depressives' attention was allocated to the primary task, rather than being limited by depression.

In Experiment 2 baseline and secondary-task latencies did not differ according to naturally occurring mood. Subsequently, the two groups recalled similar numbers of words, with the exception of one anomalous finding. The structured learning task produced a (naturally) depressive deficit in the recall of words from the *easy* sentences, but not from the difficult sentences. It appears that naturally depressed subjects had sufficient resources available for processing and recalling words from difficult contexts. It is therefore hard to understand how they would lack sufficient resources to process the words in easy contexts.

Our research suggests greater flexibility in resource allocation than what has been previously believed about depressive cognition. Rather than focusing on the *availability* of resources, accounts of depressive memory should emphasize the *allocation* of resources (Ellis & Ashbrook, 1988), which is influenced by the requirements of cognitive tasks and the extra-experimental context for performing them. For example, college students are accustomed to performing tasks of the sort that we have used. Moreover, all of our learning tasks were somewhat structured from the perspective of people who are still functional in a college setting. They do not invite a significant degree of initiative in resource allocation, the lack of which might lead depressed students to perform poorly in other arenas (see Hertel & Hardin, 1990). But tasks such these might indeed be affected by reduced initiative in resource allocation if they were performed by clinically depressed patients.

REFERENCES

Beck, A. T., Ward, C., Mendelson, M., Mock, J., L Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry, 4,* 561-571.

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

Deardorff, W. W., & Funabiki, D. (1985). A diagnostic caution in screening for depressed college students. *Cognitive Therapy and Research, 9,* 277-284.

Dunbar, G. C., & Lishman, W. A. (1984). Depression, recognition-memory and hedonic tone: A signal detection analysis. *British Journal of Psychiatry, 144,* 376-382.

Ellis, H. C., & Ashbrook, P. W. (1988). Resource allocation model of the effects of depressed mood states on memory. In Fiedler, K., & Forgas, J. (Eds.) *Affect, cognition, and social behavior* (pp. 25-43). Toronto: Hogrefe.

Ellis, H. C., Thomas, R. L., McFarland, A. D., & Lane, W. (1985). Emotional mood states and retrieval in episodic memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11,* 363-370.

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

Hasher, L., Rose, K. C., Zacks, R. T., Sanft, H., & Doren, B. (1985). Mood, recall, and selectivity effects in normal college students. *Journal of Experimental Psychology: General, 114,* 104-118.

Hasher, L., & Zacks, R. T. (1979). Automatic and effortful processes in memory. *Journal of Experimental Psychology: General, 108,* 356-388.

Hertel, P. T. (1989). The generation effect: A reflection of cognitive effort? *Bulletin of the Psychonomic Society, 27,* 541-544.

Hertel, P. T., & Hardin, T. S. (1990). Remembering with and without awareness in a depressed mood: Evidence of deficits in initiative. *Journal of Experimental Psychology: General, 119,* 45-59.

Ingram, R. E. (1989). External validity issues in mood and memory research. In D. Kuiken (Ed.), Mood and memory: Theory, research, and applications. [Special Issue]. *Journal of Social Behavior and Personality, 4,* 57-62.

Johnson, M. H., & Magaro, P. A. (1987). Effects of mood and severity on memory processes in depression and mania. *Psychological Bulletin, 101,* 28-40.

Kahneman, D. (1973). *Attention and effort.* Englewood Cliffs, NJ: Prentice-Hall.

Kendall, P. C., Hollon, S. D., Beck, A. T., Hammen, C. L., & Ingram, R. E. (1987). Issues and recommendations regarding use of the Beck Depression Inventory. *Cognitive Therapy and Research, 11,* 289-299.

Krames, L., & MacDonald, M. R. (1985). Distraction and depressive cognitions. *Cognitive Therapy and Research, 9,* 561-573.

Lubin, B. (1965). Adjective checklists for measurements of depression. *Archives of General Psychiatry, 12,* 57-62.

McDaniel, M. A., Einstein, G. ●., & Lollis, T. (1988). ●ualitative and quantitative considerations in encoding difficulty effects. *Memory & Cognition, 16,* 8-14.

Mitchell, D. B., & Hunt, R. R. (1989). How much "effort" should be devoted to memory? *Memory & Cognition, 17,* 337-348.

Potts, R., Camp, C., & Coyne, C. (1989). The relationship between naturally occurring dysphoric moods, elaborative encoding, and recall performance. *Cognition & Emotion, 3,* 197-205.

Riskind, J. H. (1989). Will the field ultimately need a more detailed analysis of mood-memory? In D. Kuiken (Ed.), Mood and memory: Theory, research, and applications. [Special Issue]. *Journal of Social Behavior and Personality, 4,* 39-44.

Roy-Byrne, P. P., Weingartner, H., Bierer, L. M., Thompson, K., & Post, R. M. (1986). Effortful and automatic cognitive processes in depression. *Archives of General Psychiatry, 43,* 255-267.

Rude, S. S., & Hertel, P. T. (1987, November). *Remembering as a consequence of cognitive effort and depression.* Paper presented at the meeting of the Association for the Advancement of Behavior Therapy, Boston, MA.

Spitzer, R. L., Endicott, J., & Robins, E. (1978). Research Diagnostic Criteria: Rationale and reliability. *Archives of General Psychiatry, 36,* 773-782.

Tyler, S. W., Hertel, P. T., McCallum, M. C., & Ellis, H. C. (1979). Cognitive effort and memory. *Journal of Experimental Psychology: Human Learning and Memory, 5,* 607-617.

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

Williams, J. M. G., Watts, F. N., MacLeod, C., Mathews, A. (1988). *Cognitive psychology and emotional disorders* (Chapter 3). New York, Wiley.