

On Mood Variation and Memory: Reply to Isen (1985), Ellis (1985), and Mayer and Bower (1985)

Lynn Hasher
Temple University

Rose T. Zacks
Michigan State University

Karen C. Rose and Bonnie Doren
Temple University

In this reply we address a number of issues concerning two recently reported findings regarding the performance of mildly depressed college students compared with their nondepressed controls: (a) depressed students showed no overall deficit in recall performance and (b) depressed students failed to show selective recall for mood-congruent (negative) events in a story (Hasher, Rose, Zacks, Sanft, & Doren, 1985). These issues include areas of agreement among the commentaries and our article and interesting empirical questions raised in the commentaries. We address a number of methodological criticisms. We also report new evidence suggesting that the Beck Depression Inventory is not a mild mood-induction procedure.

Our article (Hasher, Rose, Zacks, Sanft, & Doren, 1985) represented the dovetailing of two long-term research interests: (a) the impact of individual differences in cognitive capacity on performance (Hasher & Zacks, 1979; Zacks, Hasher, & Sanft, 1982) and (b) the degree to which memory is selective about the information that is stored about an event (Alba & Hasher, 1983; Hasher & Griffin, 1978). In this recent article we asked whether normal downward deflections in mood (mild depression or sadness) would reduce the memory performance of college students and whether this mildly negative state would result in selective recall of information consistent with the mood (an effect we termed "selectivity" and Mayer & Bower, 1985, termed mood "congruency"). We organize our discussion of the commentaries (Ellis, 1985; Isen, 1985; Mayer & Bower, 1985) around our two research questions.

Deficits in Overall Performance

Hasher and Zacks (1979) as well as others (Cohen, Weingartner, Smallberg, Pickar, &

Murphy, 1982; Ellis, Thomas, & Rodriguez, 1984; Weingartner, Cohen, Murphy, Martello, & Gerdt, 1981) have hypothesized that depressed adults (and children; see Goldstein & Dundon, in press) would show performance deficits in cognition. Deficits would occur because depression may operate directly to reduce available capacity, or because depression may distract a person from attending to ongoing events by providing compelling and competing internal musings (e.g., Kovacs & Beck, 1978).

Hasher et al. (1985) failed to find any evidence of reduced performance for mild to moderately depressed subjects. As Ellis (1985) suggests, this failure may have been because the range or intensity of mood did not differ sufficiently between our groups of subjects. Although we ourselves gave credence to this view (Hasher et al., 1985, p. 111), we note that experimenters who have compared groups of subjects with extremely divergent moods have not always reported differences in overall recall performance (Bower, Gilligan, & Monteiro, 1981; Clark & Teasdale, 1982; Teasdale & Russell, 1983; but also see Breslow, Kocsis, & Belkin, 1981).¹

We thank David Goldstein and Philip Kendall for helpful discussions regarding issues raised in the commentaries.

Requests for reprints should be sent to Lynn Hasher, Department of Psychology, Weiss Hall, Temple University, Philadelphia, Pennsylvania 19122.

¹ Breslow et al. (1981) found that hospitalized depressed adults recalled fewer of the ideas in a story than did their controls. This overall difference in recall was produced almost entirely by the depressives' poorer recall of positive

We also agree with Ellis (1985) that our task (free recall) and materials (interesting and logically structured prose passages) may have been insufficiently demanding to show effects of reduced capacity.² The problem here, though, as Ellis (1985) and others (including ourselves) acknowledge, is that there is no widely agreed upon tool to assess task difficulty. Thus there is no way, before the fact, to predict with confidence when depression (or other possible sources of reductions of capacity) will disrupt performance (see Ellis & Lane, 1984). Consider, for example, the data from Experiment 2 (Hasher et al., 1985): subjects informed about the retention test recalled more than subjects not informed. One could interpret this result to mean that the intentional subjects boosted their expenditure of effort in the task (and so engaged in a number of difficult strategies). Notice though that depressed and nondepressed subjects benefited equally from the intentional instruction, when one might have easily predicted a greater benefit of instructions for the nondepressed than for the depressed subjects, who presumably had less capacity available.

With respect to prose memory, we suspect that mood effects might be found when unstructured materials are used (e.g., scrambled stories). It is also possible that the likelihood of generating inferences, and perhaps the content of those inferences, might be susceptible to mood related differences in capacity. Nonetheless, we agree with Ellis (1985) that a metric for measuring task difficulty would be highly desirable.

Selectivity Effects in Memory

Selectivity effects are of interest beyond the identification of their existence. For one thing, they inform us of strategies subjects may use in dealing with complex or supraspan events. Selectivity on a memory test, especially on a recall test, does not, however, provide unambiguous information about what was actually encoded into memory. In making inferences from selectivity effects seen in recall protocols to what was encoded (or to the contents of memory), there is a potential for

serious error in the direction of considerably underestimating the true contents of memory. This error in turn would lead to underestimating the potential accuracy of memory (Alba & Hasher, 1983).

Our concern about the possibility of underestimating the contents of memory, coupled with widespread reports of selectivity effects shown by depressives, prompted the selectivity portion of this project. The high frequency of recall of negative events by adult depressives in therapy is the *sine qua non* of their diagnosis (see e.g., Beck, 1967). Furthermore, Breslow et al. (1981) report that hospitalized depressives and controls exposed to the same information show different recall patterns, with depressives recalling far less of the positive information than controls. The interpretation of these observations is critical. Are these differences the result, as Beck (1967) has suggested, of depressive schemata that guide the selective encoding of information into memory? Or are the differences the result of postencoding processes such as selective rehearsal and retrieval plans (see e.g., Teasdale & Russell, 1983, p. 170-171)? If the latter, depressives may actually encode the same information as nondepressives. And if encoding is equivalent, then a retrieval manipulation introduced by Hasher and Griffin (1978) might induce depressives to report more of the positive information stored in memory. Such a finding would be of interest to memory theorists because it addresses the issue of the fundamental accuracy of memory. Such a finding would also be of interest to clinical theorists and practitioners because Beck's theory of depression and its treatment (e.g., 1967; Kovacs & Beck, 1978) suggests that recalling more of the positive events in one's life will make a depressed person feel better.

Hence our initial step of a series of studies designed to evaluate these issues was intended to demonstrate that mildly depressed college students would show the selectivity that had been reported with profoundly depressed people (Breslow et al., 1981; Clark & Teasdale,

ideas; recall of the negative and neutral ideas did not differ between the two groups of subjects.

² Bower et al. (1981) also used interesting, well-structured passages. They too failed to find overall differences in recall between subjects hypnotized to be either depressed or elated.

1982) and with hypnotized or Velten-induced people (Bower, Gilligan, & Monteiro, 1981; Teasdale & Russell, 1983). Because we initially accepted the assumption that mood falls along a continuum from mildly through moderately to severely negative, an assumption we now see reason to doubt (Hasher et al., 1985, p. 117), we anticipated that the effect among mildly depressed people would be smaller than that reported for profoundly depressed people. We prepared to detect the small effect by running large numbers of subjects (356 in Experiment 1). Nevertheless, we were unsuccessful in three attempts to find selectivity effects. We eventually concluded that selectivity may not exist among mildly depressed persons and our reading of the literature, now substantially buttressed by Isen's (1984, 1985), made it seem likely that our null effects were meaningful on their own merits and were not the product of insensitive measurement procedures.

Unlike Isen (1985), to some extent Ellis (1985), and certainly ourselves, Mayer and Bower (1985) are less impressed by the other negative findings in the literature. They argue that selectivity effects probably do exist among mildly depressed college students, but because the effect size is small, an extremely sensitive study is required to detect it. Thus they turn their attention to the methodological problems in detecting small effects.

Thematic Intertwining of Ideas

A central criticism is directed at two of the three story versions used in Experiments 1 and 3. In particular, Mayer and Bower argue that in these two versions, the negative and positive evaluative ideas were thematically intertwined and as a result would tend to cue each other at recall. This cuing, in turn, would dissipate the otherwise small selectivity effect. The role of thematic intertwining in recall certainly warrants research. Note, though, that our neutral story version was not criticized for extensive intertwining, and a glance at the data in Tables 3 and 5 of Hasher et al. (1985) reveals no more of a tendency for selectivity effects in the less intertwined neutral story than in the more intertwined positive and negative versions. In addition, selectivity effects with a story com-

parable to our neutral one have been reported, but these are based on a comparison of profoundly depressed versus nondepressed adult subjects (Breslow et al., 1981). Bower et al. (1981) used three different narrative passages which varied in (among other things) the degree of intermingling of good and bad events. Selectivity (for hypnotized subjects) was shown for each story.

Mayer and Bower (1985) may be correct in arguing that intertwining of positive and negative story elements reduced selectivity effects. However, materials which clearly segregate positive and negative events may be invalid on other grounds. Our working assumption in preparing materials was that in daily life good and bad events are intermingled, not segregated. Although one might recall a day as basically a bad or good one, events that were discrepant with the larger conclusion were no doubt interspersed with the ones that led to the summary assessment. Further, many days are neither overwhelmingly good or bad. We were trying to simulate the characteristics of normal experiences because we wanted our results to be relevant to theories of depression such as Beck's (1967).

According to Beck, it is not the case that the experiences of seriously depressed people consist of almost entirely negative events. The events in the lives of many depressives are not, to an objective observer, different from the events in the lives of others who are not seriously depressed. What is critical is that the negative schema of depressives and of depression prone individuals induces them to selectively emphasize the negative events.

Measurement Instruments: The State Versus Trait Issue

A second concern raised by Mayer and Bower (1985) centers on the paper and pencil instruments we used to measure mood. Two points are made: (a) these instruments may measure enduring traits rather than temporary mood states and (b) our use of them in Experiment 2 may have been reactive in that they altered the performance of subjects in such a way as to obscure the existing selectivity effects. With respect to the state versus trait issue, we note that there is controversy in the clinical literature. Our goal was to

measure momentary mood, and to this end we followed procedures in wide use in the experimental clinical literature. These procedures of course may be seriously flawed (see e.g., Depue & Monroe, 1978).

There is at least one major reason to believe that our paper and pencil instruments were measuring state aspects of mood. Both the Multiple Affect Adjective Checklist (MAACL) and the Beck Depression Inventory (BDI) have been shown to produce unstable scores over intervals of several days to weeks (e.g., for the MAACL, Pankrantz, Glaudin, & Goodmonson, 1972; for the BDI, Bumberry, Oliver, & McClure, 1978; Hammen, 1980; Hatzenbuehler, Parpal, & Mathews, 1983), suggesting that they measure states not traits. Similar findings were reported in Experiment 1 of Hasher et al. (1985). There, 146 subjects returned for a recall test after a 48-hr interval. Of these, 37 to 81 (depending on the BDI criterion) had been determined to be depressed at the initial session, two days earlier. Upon retesting with the BDI, 41% to 59% of these depressed subjects were no longer depressed by the initial measure.³

It is entirely possible of course, as Mayer and Bower (1985) and others (see e.g., Howarth & Schokman-Gates, 1981) suggest, that the paper and pencil instruments in widest use assess both state and trait aspects of depressed mood. However, if some of our subjects were in fact "trait" depressed, our failure to find selectivity is, given Beck's schematic view and the Breslow et al. (1981) findings, all the more surprising.

One additional mood measure was available in Experiment 2, a 7-point self-rating scale in which subjects were asked to rate how they were feeling at that moment. This was a measure used by Bower et al. (1981) and is akin to the 100-point mood thermometer used by Teasdale (e.g., Teasdale & Russell, 1983). A reanalysis of the data from Experiment 2 using subjects who rated themselves at points 1, 2, and 3 (*extremely sad*, *very sad*, and *slightly sad*, respectively) versus those who rated themselves at points 5, 6, and 7 (*slightly happy*, *very happy*, *extremely happy*, respectively) did not alter the findings of our study. Subjects who were sad versus happy recalled roughly the same overall amount of information, $F < 1$. There was

also no evidence of mood-congruent selectivity ($F < 1$, for the Mood \times Item Type interaction). Thus even with a measure that might be a cleaner measure of momentary mood, we find no evidence of different patterns of recall.

Measurement Instruments: A Source of Reactivity

The issue of reactivity was raised with respect to the filled retention interval between the reading of the story and its recall in Experiment 2. There, subjects worked on a variety of problems and then filled in two mood inventories (in addition to the BDI which was given at the beginning of the session). Both inventories contained negative material. Mayer and Bower argue that the negative content of the inventories could have interfered with the ability of depressed subjects to remember selectively the original negative material. Although Mayer and Bower offered no particular mechanism to explain the predicted interference, we could propose that the additional negative materials cause depressed subjects to reflect on their own personal negative thoughts, resulting in reduced rehearsals of the original story. The additional negative materials might have reminded nondepressed subjects of the negative material in the original story, resulting in increased rehearsals of this material. These complementary effects would then obscure a latent small selectivity effect.

Of course, another scenario is also possible; the negative material in the intervening rating tasks might trigger increased rehearsal of the story by both depressed and nondepressed subjects. If depressives did indeed have more of the negative material in memory after reading the story than nondepressives, selectivity should still have been seen. Thus the equivalent recall of the positive and negative information by mildly depressed and nondepressed subjects could simply reflect the confluence of equivalent encodings and equivalent rehearsals. In any event the issue of the

³ The stability versus instability of mood did not reliably influence either total recall or selectivity of recall (see Hasher et al., p. 107). These findings are of course in keeping with the overall findings of the study.

impact of interpolated materials warrants research.

The implicit assumption that underlies the argument made by Mayer and Bower (1985) on the interpolated materials may be seen as suggesting that the mood inventories are themselves mild mood inducers. This view had occurred to one of us: Rose (1984) argued that the BDI might itself be a mood induction procedure because as much as 75% of its material is negative. This possibility seemed particularly plausible when we compared the BDI with the MAACL; the MAACL has a far smaller proportion of blatantly negative items which are, in any event, interspersed among some positive adjectives. Rose gave both the BDI and the MAACL to 240 college students, half of whom received the BDI first, half the MAACL first. She expected to find that more subjects tested as depressed on the MAACL when it followed rather than preceded the BDI. Using an MAACL score of 14 as the measure of depression, approximately 48% (of 120) were depressed when the MAACL was first, and 53% (of 120) were depressed when it was second. This difference is not reliable by a χ^2 test.

Emotionality of Subjects

Mayer and Bower (1985, p. 399) also proposed that our depressed subjects may have been generally more emotionally reactive than the subjects who were not detected as depressed. This argument might lead one to expect our depressed subjects to recall more of both the positive and the negative items than nondepressed subjects, at least as compared with neutral items. But no such evidence was found. Nor is there such evidence when the data from Experiment 2 are reanalyzed using subjects who rated themselves as happy versus sad. Although our experiment was not designed to test this hypothesis, one certainly could be.

Conclusion

Isen's (1985) commentary makes clear that our negative results on mood selectivity are not divergent with the preponderance of findings on the effects of sadness on memory, including findings from studies using mood-

induction procedures. Because induction procedures allow control of mood intensity, and because they are clearly state (rather than trait) manipulations, the similarity of our findings to findings with induced sadness temper concern about the methodological limitations of our research.

We concur with Isen that the data as a whole indicate qualitative differences between positive and negative affects on the one hand, and between mild and clinical levels of depression on the other. For both of these dimensions, Isen provides intriguing hypotheses about the nature of the difference, hypotheses that raise interesting empirical questions. Take, for example, her idea that the depressive schemata of normal and mildly depressed individuals are less well elaborated and integrated than those of severely depressed individuals. This idea suggests that measures of schematic structure (such as those obtained from free association tasks) will show differences between normal (or mildly depressed) and clinically depressed individuals, even when mood intensity has been equated by subjecting the nonclinical group to sadness induction procedures. In short, future research that follows up on Isen's theoretical analyses should greatly enrich our understanding of the relation between affect and memory.

In closing, we believe that the commentaries taken together suggest a number of interesting empirical questions. Until new evidence is in, we believe that our findings must be taken seriously by those interested in interactions between mood and memory. Our research also suggests that the use of mild to moderately depressed college students as analogues of clinical depressives requires careful exploration of the similarities and differences between behavior in normal moods and behavior in depression.

References

- Alba, J. W., & Hasher, L. (1983). Is memory schematic? *Psychological Bulletin*, 93, 203-231.
- Beck, A. T. (1967). *Depression: Clinical, experimental and theoretical aspects*. New York: Harper & Row.
- Bower, G. H., Gilligan, S. G., & Monteiro, K. P. (1981). Selectivity of learning caused by affective states. *Journal of Experimental Psychology: General*, 110, 451-473.
- Breslow, R., Kocsis, J., & Belkin, B. (1981). Contribution of the depressive perspective to memory function in

- depression. *American Journal of Psychiatry*, 138, 227-230.
- Bumberry, W., Oliver, J. M., & McClure, J. N. (1978). Validation of the Beck Depression Inventory in a university population using psychiatric estimate as the criterion. *Journal of Consulting and Clinical Psychology*, 46, 150-155.
- Clark, D. M., & Teasdale, J. D. (1982). Diurnal variations in clinical depression and accessibility of memories of positive and negative experiences. *Journal of Abnormal Psychology*, 91, 87-95.
- Cohen, R. M., Weingartner, H., Smallberg, S. A., Pickar, D., & Murphy, D. L. (1982). Effort and cognition in depression. *Archives of General Psychiatry*, 39, 593-597.
- Depue, R. A., & Monroe, S. M. (1978). Learned helplessness in the perspective of the depressive disorders: Conceptual and definitional issues. *Journal of Abnormal Psychology*, 87, 3-20.
- Ellis, H. C. (1985). On the importance of mood intensity and encoding demands in memory: Commentary on Hasher, Rose, Zacks, Sanft, & Doren. *Journal of Experimental Psychology, General*, 114, 392-395.
- Ellis, H. C., & Lane, J. W. (1984, November). *Mood states and memory for generated words*. Paper presented at the meeting of the Psychonomic Society, San Antonio, TX.
- 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.
- Goldstein, D., & Dundon, W. D. (in press). Affect and cognition in learning disabilities. In S. J. Ceci (Ed.), *Handbook of cognitive, social, and neuropsychological aspects of learning disabilities*. Hillsdale, NJ: Erlbaum.
- Hammen, C. L. (1980). Depression in college students: Beyond the Beck Depression Inventory. *Journal of Consulting and Clinical Psychology*, 48, 126-128.
- Hasher, L., & Griffin, M. (1978). Reconstructive and reproductive processes in memory. *Journal of Experimental Psychology: Human Learning and Memory*, 4, 318-330.
- Hasher, L., Rose, K. C., Zacks, R. T., Sanft, S., & Doren, B. (1985). Mood, recall, and selectivity effects in normal college students. *Journal of Experimental Psychology: General*, 114, 106-120.
- Hasher, L., & Zacks, R. T. (1979). Automatic and effortful processes in memory. *Journal of Experimental Psychology: General*, 109, 356-388.
- Hatzenbuehler, L. C., Parpal, M., & Mathews, L. (1983). Classifying college students as depressed or nondepressed using the Beck Depression Inventory: An empirical analysis. *Journal of Consulting and Clinical Psychology*, 51, 360-366.
- Howarth, E., & Schokman-Gates, K. (1981). Self-report multiple mood instruments. *British Journal of Psychology*, 72, 421-441.
- Isen, A. M. (1984). Toward understanding the role of affect in cognition. In R. Wyer and T. Srull (Eds.), *Handbook of social cognition*. Hillsdale, NJ: Erlbaum.
- Isen, A. M. (1985). The asymmetry of happiness and sadness in effects on memory in normal college students. *Journal of Experimental Psychology: General*, 114, 388-391.
- Kovacs, M., & Beck, A. T. (1978). Maladaptive cognitive structures in depression. *American Journal of Psychiatry*, 135, 525-533.
- Mayer, J. D., & Bower, G. H. (1985). Naturally occurring mood and learning: Comment on Hasher, Rose, Zacks, Sanft, and Doren. *Journal of Experimental Psychology: General*, 114, 396-403.
- Pankrantz, L., Glaudin, V., & Goodmonson, C. (1972). Reliability of the Multiple Affect Adjective Checklist. *Journal of Personality Assessment*, 36, 371-373.
- Rose, K. C. (1984). *Is the Beck Depression Inventory a mood induction procedure?* Unpublished manuscript.
- Teasdale, J. D., & Russell, M. L. (1983). Differential effects of induced mood on the recall of positive, negative, and neutral words. *British Journal of Clinical Psychology*, 22, 163-171.
- Weingartner, H., Cohen, R. M., Murphy, D. L., Martello, J., & Gerdt, C. (1981). Cognitive processes in depression. *Archives of General Psychiatry*, 38, 42-47.
- Zacks, R. T., Hasher, L., & Sanft, H. (1982). Automatic encoding of event frequency: Further findings. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 8, 106-116.

Received March 20, 1985 ■