CHAPTER - II


**Review of literature**

In the preceding chapter, it has been pointed out that the study of conversation memory provides a means for examining the content and structure of social knowledge that play a role in interaction processes. Experimental studies of different types on conversation memory have amply justified this view. In this chapter, we shall review some of the important studies that bear direct or indirect relation with the present problem. The chapter is divided into three sections. The first section deals with the review of studies that demonstrates the influence of humorous material on conversation memory. The second section is devoted to the review of congruent-incongruent version on conversation memory. The third and final section of this chapter reviews those studies that demonstrate the relationship between emotions and memory. More specifically the final section review such studies which throw light on the impact of different shades of listener’s mood on conversation memory.

**SECTION- Ist**

Kaplan and Pascoe (1977) in a study entitled “humorous lecture and humorous examples: some effects upon comprehension and retention” explored the effect of two types of humor upon learning in a lecture situation. The two types of humor are humor related to the concepts presented in the lecture (or humorous examples) and humor unrelated to the lecture’s content. The primary concern was to determine how varying correspondence of humor with the topics of a classroom lecture moderates the comprehension and retention of lecture material.
The subjects were 508 undergraduate students at a large public university. These participants were enrolled in the 16 sections of introductory psychology and the experiment was part of their regular instruction. Thus, subjects were run in large groups ranging from 13 to 45 students. Close-circuit videotapes were customarily shown to all section as part of course instruction. Subject saw one of four version of a 20-min black and white videotape about Freudian personality theory. Factual material in the lecture came from the chapter on Freud presentation in Hall and Lindzey (1970). Points covered in the lecture included a biographical sketch of Freud; the concepts of id, ego and superego; anxiety and defense; stages of psychosexual development, neo-Freudians; and projective techniques of personality assessment. One serious version and three humorous versions of the lecture were recorded on videotape. The humorous version included humor directly related to some concepts in the lecture (concept humor), humor unrelated to any of the concepts (non-concept humor), and a combination of some non-concept and concept humor (mixed humor). Six main concepts were presented in the each version and they were spaced evenly throughout the lecture. In the concept humor version, all of the concepts were illustrated by way of humorous example. An assistant professor of psychology delivered all four versions of the lecture. All humorous lines were practiced with a pilot audience and delivered on videotape with appropriate inflection to maximize their impact. An example of concept humor centered on the difficulty in interpreting sentence completion assessments of personality. After explaining the typical clinical procedure, the lecture offered a variety of straight and humorous sentence completions. To the sentence root “Animals.....” the speaker responded with “scare me” for the nonhumorous version and used often try to fool me” as the humorous punch line. Similarly, the stem “I would like....” met with “to be a
doctor" for the serious condition and "to drink blood" for the humorous condition. An example of non-concept humor regarding psychosexual stages of development was the lecture's comment that "Freudians would probably have a good time deciding what stage Linda Lovelace was fixated at".

After the videotape was played, a two-part questionnaire and answer sheet were distributed to each student. The first section consisted of semantic differential scales describing separately the speaker and the lecture. Six word pairs taken from Smith (1959) were used to rate the speaker and the lecture. These word pairs were as follows: valuable-worthless, interesting-boring, serious-humorous, cold-hot, optimistic-pessimistic, and light-heavy. A numbered, five-choice response space separated each word pair. Subjects were instructed to mark on their answer sheet the number of the choice in each word pair that best described the lecture they had seen. The second part of the questionnaire was comprised of 11 multiple-choice questions on the content of the lecture. Of these items, 6 focused on the six critical examples used in the concept humor version. These items were dubbed humor items. The other 5 items were based upon other information in the lecture and are defined here as non-humor items. Subjects were instructed to record on their answer sheet which of four alternatives was the best response for each question. Two additional items on the questionnaire were included to code each subject's class section and grade on the previous unit exam. The same 11-item quiz was readministered to the subjects 6 weeks after initial tape viewing and testing. This retesting was conducted by the instructor of each section.

Result suggested that all of humorous lectures were perceived as significantly more humorous and light than the serious presentation.
Although there were some differences between humorous lecture groups (e.g., the concept humor lecture was slightly less effective than the other humorous presentation), differences between the humorous groups were trivial in comparison to the strong differences between each humor group and the group hearing the serious presentation. These distinctions were noticeable when audiotapes, which were recorded while subjects viewed the lecture, were compared for volume of laughter. Unlike students viewing the serious version, those viewing humorous version laughed during the lecture, especially in response to humorous punch lines. Thus there is compelling evidence that manipulation was effective. It was observed that any preexisting differences were statistically non-significant. It was also suggested that the groups differed on performance for non-humor items, but did not differ for humor items. It should be noted that these students performed non-significantly better on the items based on the humorous example. Analysis revealed no significant difference between the two groups on either of these test questions. However, performance on the postjoke item for the non-concept humor group was slightly better and in the direction predicted by the arousal hypothesis (p < .1). The analysis of variance results were as predicted—groups viewing lectures with more concept-related humor did significantly better on items testing recall of the humorous examples than did the serious lecture group. Although total test scores were not significantly different.

Keenan, MacWhinney, and Mayhew (1977) undertook a study entitled “Pragmatic in Memory: A Study of Natural Conversation”. They examined (a) the generalizability of laboratory experiments on sentence memory to real-world settings and (b) the possible existence of differences in memory for sentences that are personally meaningful.
versus memory for sentences that are not 19 members of the Psychology Department (Faculty and graduate students) at the University of Denver served as subjects. They were participating in research discussion and were unaware that they would later be given a memory test on the discussion. The procedure format for a research luncheon discussion at the University of Denver consisted of one participant presenting a set of ideas or experiments and others freely interrupting to ask question, expand ideas, give criticism, or make humorous remarks. This research discussion under study was held in the usual manner with one exception: it was tape-recorded. Participants were aware of the tape records but were not aware of the purpose for taping, nor did anyone inquire as to the purpose. The speaker for the day was one of the authors. He did not attempt to inject statements particularly suited to the experiment; he simply spoke in his usual manner. After the discussion they played the tape, selecting statements for the recognition test. Two classes of statements were defined: high interactional content statements and low interactional content statements. High interactional content statements typically conveyed wit, sarcasm, humor, or personal criticism. Through mutual decision, 15 such statements were selected for the recognition test. To match these 15 statements, 15 statements which were low in interactional content were selected from tape in a random fashion. Eight of the 15 high interactional content statements were uttered by the main speaker and the remaining 7 were uttered by other participants. The 15 low interactional content statements were matched with the high interactional content statements in terms of speakers.

The mean percentage of OLD (I'm sure the statement was said) responses to targets and paraphrases of targets, pooling over high and low interactional content items, was 28% contrasted with a mean percentage
of 5% for statements expressing new content. There were 37% OLD responses to high interactional content targets and their paraphrases, and only 18.5% OLD responses to low interactional content targets and their paraphrases. When the interactional content was high, there were 56% OLD responses to targets compared to 18% for paraphrases. On the other hand, when the interactional content was low, subjects responded OLD equally often to targets (19%) and paraphrases (18%). There was no effect of speaker on the memorability of utterances, nor was the interaction significant. The only significant effect was the effect of interactional content.

Control study 1 was conducted to answer the question: Are there differences in sentence memorability when the interactional component is removed? A typical sentence memory experiment was performed using the target statements of the first experiment. The sentences were presented in a list-learning fashion, and memory was tested by the same recognition memory test employed in the conversation experiment. 34 Introductory psychology students from the University of Denver volunteered to serve as subjects for course credit. Subject listened to a tape consisting of 100 conversational-tape sentences. The list of sentences included the 15 high and 15 low interactional content statements from the conversation experiment and 70 filler sentences. Fillers were included to prevent ceiling effects on the memory test. All of the sentences were recorded by the speaker who spoke in a conversational tone. The order of sentences presentation was randomized with constraint that the first 10 and last 10 sentences of the list were filler items to control for primacy and recency effects. Subjects were instructed to listen carefully to the sentences in order to do well on a subsequent memory test; the nature of the test was not specified. Immediately after listening to the 100
sentences, the subjects were given the same memory test used in the conversation experiment, except here the items were not identified by a speaker. As before, the subjects were told to rate each alternative using the 3 point scale of OLD (I am sure the statement was said), GUESS (I cannot remember anything about this statement), and NEW (I am sure this statement was not said). Collapsing over high and low interactional content statements, the percentage of OLD responses to targets and their paraphrases was 32%, while the percentage of OLD responses to new content destructors was only 7.5%. The subjects did not show differences in memory between high and low interactional content statements. The percentage of OLD responses to the target was exactly the same, 47%, for both high and low interactional content statements.

The findings of the previous experiments led Keenan et al. (1977) to assume that high recognition rates for high interactional content statements reflected subjects' memory for the surface forms of these statements. However, an alternative explanation may be that the paraphrase distractors for these statements were of poorer quality than paraphrases for low interactional content statements. Thus the third experiment was carried out to test this alternative explanation. In this experiment subjects were given the 15 high and 15 low interactional content statements, each paired with its paraphrase, and asked to choose which of the two statements would be most likely to occur in a conversational setting such as the research luncheon discussion. 21 students in an undergraduate learning course at the University of Denver agreed to serve as subjects during class time. None of the subjects knew the speakers from the conversation experiment. The subjects were first given an extensive description of the setting for the conversation experiment. This included a description of the people involved, the
purpose and procedural format for the discussion, the manner in which people interact in these discussions, and some information concerning the topic under discussion. They were also told certain statements had been selected from the discussion and had been constructed for them; their task was to distinguish paraphrases from the real statements. Following this description subjects were given test booklets containing 30 pair of statements. Each pair consisted of one of the 30 target statements from the conversation experiment and its paraphrase. The subjects were told that one of the statements in each pair had actually occurred in the conversation, and that their task was to try to determine which one it was. The paraphrases were selected more often than the targets for both higher interactional content statements, 59 versus 44%. This difference is highly significant and there was no interaction between target and paraphrase and high and low interactional content. These results clearly showed that there was nothing peculiar about the paraphrase distractors for high interactional content statements that could have led subjects in the conversation experiment to reject them in favour of the targets. The paraphrases for both high and low interactional content statements were equally good candidates for conversational statement, and, interestingly, were even thought to be better candidates than the targets themselves.

In the preceding experiment, subjects did not know the speakers. It was assumed that knowledge of the speaker’s stylistic habit may determine which statement occurred in the conversation experiment. The final experiment was conducted to test this assumption. Thus final experiment investigated whether knowledge of the speaker allows one to select the target rather than the paraphrase for high interactional content statement but not for low interactional content statements. If so, then the surface form effect in the conversation experiment could not be
interpreted as a result of memory; rather, it would reflect reconstruction on the basis of knowledge of the speaker’s stylistic habits. 15 University of Denver faculty and graduate students who rated themselves as familiar with the speakers and the setting of the conversation experiment, but who did not attend that particular discussion, served as subjects. The procedure was similar to that of preceding experiment. The only difference was that each pair of statements in this case was identified by the speaker. In making their judgments which statement occurred, the subjects were asked to consider, phrase structure, lexical items, and overall “sound” of the statement in terms of the particular speaker. It should be noted that all subjects were highly motivated and none felt that their selections were at all random. Subjects who knew the speakers were totally unable to distinguish the targets from the paraphrases, both in the case of the high interactive content statements and in the case of low interactive content statements. The analysis of variance confirmed that there was no main effect of interactive content, or target-paraphrase, and there was no interaction. These results clearly supported the contention that the participants in the discussion actually remembered the surface forms of high interactive content statements, and did not simply base their responses on their knowledge of the stylistic habits of a given speaker. The present experiment showed that knowledge of the speaker was not sufficient to distinguish the targets from the paraphrase. Thus, for the sentences used in these experiments, knowledge of the speaker was only sufficient to overcome a response bias for the paraphrase and allow responding in accordance with chance.

Kintsch and Bates (1977) carried out two experiments on memory for classroom lectures. They were interested in two important questions. First, will student remember only the meaning of a lecture or will they
remember the meaning plus the actual words used? Second, is there a difference in the amount of memory for various types of statements? In particular, are topic statements remembered better than mere illustrative material and is there preferential memory for extraneous statements (e.g., jokes, announcements) embedded in a lecture? Three types of statements were tested: topic statements, details, and such extraneous remarks as joke announcements. The findings revealed that two and five days after listening class lecture, students showed significance memory for meaning in that they were able to discriminate sentences actually spoken in the lecture or their paraphrases from control sentences that they had never heard before. Moreover extraneous sentences - jokes, comments to the audience, and announcements were recognized better than descriptive statements. Such a finding was explained by the researcher's term of von Restorff effect in which unique elements are remembered better than non-unique elements. Such an effect might be enhanced by the fact that jokes in particular provide relief from the heavy information load of the standard lecture.

Zillmann, Williams, Bryant, Boynton and Wolf (1980) investigated acquisition of information from educational television programs as a function of differently paced humorous inserts. The investigation was designed to determine the particular effects of the involvement of humor and, more specifically, of the pacing of humor in educational programs for children on information acquisition and on hedonic aspects of the learning experience. Nonhumorous educational programs were produced, and, in the humor conditions, identical humorous inserts were interspersed at a comparatively slow or fast pace. Since the humorous programs were thus necessarily of longer duration than the no-humor control program, giving the post-exposure
information-acquisition test, and since the particular schedule of humor interspersion can be seen as a unique disruption pattern that impairs information acquisition, two additional control conditions were employed. One control condition replicated the interspersion schedule of slow pacing, and the other control condition replicated that of fast pacing. Non-content (i.e. black screen) was inserted in these controls to assure minimal distraction from the educational message.

The humor employed in this investigation was unrelated to the specific contents of the educational message. As a consequence, any findings cannot readily be generalized to messages in which humor is a truly integral part of the educational information.

76 Kindergartens and first-grade children (ages 5 years 6 months through 7 years 2 month), approximately half whom were males and half of whom were females, participated as subjects in the experiment. All of the children were students at the Fort River Elementary School in Amherst, Massachusetts. The children's parents had been contacted by letter concerning the experiment. If they were willing to let their child or their children participate, the parents signed and returned an informed consent statement. Parental permission was received for 81% of the children enrolled in these two grades level. For the investigation, children were randomly selected from this subject pool. Four educational presentations were especially designed and produced for the experiment. Humorous episodes were taken from several televised cartoon and Muppet programs. In the humor conditions, humorous episodes were interspersed in the educational segments in such a way that the presentation of particular facts, circumstances, or views was not disrupted. In the humor conditions that controlled for the time of the interspersed humorous episodes, the screen was left blank (i.e. black
without sound) for exactly the periods of time taken up by the humorous episodes in the time-controlled condition. Each educational segment was composed of material that was considered to be novel to most children of the age level tested, yet the segments were written and produced so that the target audience could readily comprehend the bulk of the material presented. Each segment employed child actors and actresses; additionally, adult talent was employed in three of the four segments. All presentations used dialogue and were set in the home environment. Professional production facilities were employed in all instances, and the finished products were considered to be equivalent in quality to children's educational television stations. All segments were presented in color. The following four educational presentations were employed. (1) The "Philatelist", a segment about the hobby of stamp collecting, which presented two 8 years-old boys who were examining a stamp album while discussing stamp of other countries, the prices of rare stamps, and some of the procedures used by philatelists. This segment was 2 min 7 sec long. (2) "The Angkor of the Ennuit", which featured a little girl who was sick with a cold and whose mother consoled her while comparing the procedures of the typical American physician with the functions and method of the Ennuit medicine man. The segment was 4 min 2 sec long. (3) "Epiphytes", which showed a mother potting root-bound plants and discussing with her daughter plants that grow in environments other than soil. The segment was 2 min 53 sec long. (4) "Marriage in the Gisu Tribe", which presented a mother who was trying to wake her sleepy daughter so that they could attend a wedding. The mother described the customs and bonding ritual of the Gisu tribe of Uganda. This segment was 4 min 30 sec in length. The humorous bits were eight segments from children's cartoon and Muppet programs. They were each approximately 30 sec long and were composed primarily of slapstick routines and short
“one-liners” delivered by the Muppets. The episodes were self-Contained and were selected on intuitive grounds as at least moderately funny and moderately appealing to children at the age level tested. The contents of humorous episodes were totally unrelated to the educational messages in which they were embedded educational segments, episodes humorous and time controlling blanks were edited as to create the fine experimental communication. The no-humor control condition was composed of the four educational segments described previously (1-4, in that sequence). Its duration, accordingly, was 13 min 32 sec. All other conditions were about 4 min longer (8 times 30 sec, the average time of a humorous episode). In the slow-paced humor condition, two consecutive humorous episodes were placed after each of the four educational segments. The average time the educational message without humor was thus about 200 sec. In the fast-paced humor condition, in contrast, one humorous episode was placed approximately in the middle, and one was placed at the conclusion of each of the four educational segments. The order of the humorous episodes (1-6) was not changed, however. In this condition, the average time of the educational message without humor was about 100 sec. The two remaining conditions were controls for the two different distributions of humorous episodes. With both location and time of the episodes being controlled by blanks, these conditions provided a slow-paced and a fast-paced no-humor control.

Analyses of variance (ANOVAS) performed on the various dependent measures failed to reveal any appreciable differences between male and female respondents. The sex-of-respondent variable was consequently collapsed, and all findings reported apply to both male and female Kindergarten and first-grade children. The findings strongly support the view that humor in educational programs for children, even if
the humorous stimuli are rather arbitrarily interspersed, fosters increased attentiveness and, ultimately, superior information acquisition. The data regarding visual attention leave no doubt about the fact that exposure to humorous stimuli created high levels of attentiveness and that these elevated levels of attentiveness extended into the exposure to educational materials. At the attention level then, the data unequivocally disconfirm and appose the view that the involvement of humor in education only distracts from the essential message and, hence, impairs information acquisition. The finding that the presentation of humorous materials in a fast pace produced the enhancement effect on information acquisition more rapidly than their presentation in a slow paced is entirely in accord with the reasoning on vigilance and alertness. In practical terms, the educator who deals with an audience whose attentiveness is below the level necessary for effective communication should indeed benefit from employing humor early and in frequent short burst. The findings of the present investigation fail to support the view that education suffers from humor involvement in that the serious segments conveying the educational message will be perceived as particularly joyless, uninteresting, and dull, in contrast to the humorous portions. Rather, the data lend limited support to the opposite claim: the positive affective reaction to humor appears to generalize to the educational message. In the present investigation, humor was entirely secondary to education (in terms of time). If this emphasis is reversed and humor becomes dominant, it must, of course, be expected that sooner or later the disruption of the educational message will produce learning impairments. Needless to say, the facilitating effect of humor on information acquisition can also be neutralized and overpowered by interference from disruption. The excessive use of humor is certainly not advisable. However, it remains to be determined empirically at exactly what point and under exactly what
circumstances the use of humor in educational produces undesirable effects and, hence, warrants the verdict of excessive.

Finally, it should also be clear that the findings of the present investigation can't readily be generalized, in fact, they should not be generalized to all conceivable motivational circumstances in the educational setting. The findings accord with the proposal that humor-instigated vigilance mediates greater attentiveness to nonhumorous portions of the surrounding program. The proposal presupposes that attentiveness to those nonhumorous portions is not at extreme levels.

Stafford and Daly (1984) conducted an experimental study entitled "conversation memory: The effort of recall mode and memory expectancies on remembrance of natural conversation". They have pointed out that conversations though, are part of our every day life and central to the social life of an individual, unfortunately very little attention has been paid to the remembrance people have of them. Conversations represent stimuli that are distinctly different from most other stimuli that people are required to remember. Conversations are jointly created, constantly update and necessarily are readily accessible by participants. Thus Stafford and Daly (1984) explored memory for the natural conversation in 128 undergraduates who participated in dyadic social exchange. Even only after 5 minutes, subjects were able to recollect only about 10 percent of what was said. The modes of the recall (written verses oral) and the presence and type of memory goal affected what and how much, interaction was recalled. Subjects also remembered more of their partner's comments in the interaction than their own. It was concluded that personality, recall mode, and expectancy concerning the upcoming memory task all determine how much and what was remembered.
Gibbs (1987) examined the hypothesis that people best remember requests that do not specify the main projected reason for addresses not complying with the request. Result of a recognition memory using eighty-four undergraduates supported this idea when subjects heard requests stated in conversational content. Without content, however, there were differences in subject memory for the different type of requests. It was suggested that people memory for requests in conversation is dependent on the assessment made by speakers and listeners of the particular plans and goals that each has in various social situations.

Benoit and Benoit (1987) carried out an experimental study to investigate conversational memory using cued and free recall. This study focused on a comparison of cued and free recall, testing two hypotheses: (a) verbal reports of Conversational interaction employing cued recall are significantly more accurate than those employing free recall (b) conversational interactants remember more of their interactional partner’s utterances than their own. Results showed that subjects recalled significantly more conversation under the cued recall than under free recall condition. There was no difference between recall of own and other comments overall under the free recall condition. However, subjects in the cued recall condition, subjects remembered more of their interactional partner’s utterances. Finally, they concluded this study tested the accuracy of verbal reports elicited from memory by free and cued recall items. Cued recall produce significantly higher amounts of memory than free recall, demonstrating that humans are able to store and retrieve substantial amounts of conversational information, when properly elicited. Subjects did not recall more of their interactional partner’s remarks than their own. In one condition (cued recall), subject remembered more (82% verses 76%) of their own than the others.
utterances, but this is quite possibly an artifact of the accuracy criterion employed (gist). Further work needs to be conducted into the relative efficacy of various recall cues, and data here supported the notion that semantic cues are stronger than temporal one (89% versus 69%). Research employing a verbatim criterion could determine whether the greater memory for own comment in the cued recall condition was an artifact of coding criterion, or whether subjects actually remember more of their own than other’s utterances under cued recall.

Hjelmquist (1989) in an experimental study tested Recognition memory for utterances in conversation. In his study forty four university student listened to a dyadic conversation and were tested immediately afterwards or four days later.

Neither two people in tape recorded conversation know each other nor did the audience know the two speaker-listeners on the tape. This design was used to further clarify the role of familiarity and personal participation for memory of conversational discourse. Gist memory was high in both conditions with lower performance after four days. Surface memory could be verified, and the level of variables used to index surface memory was about the same in both the immediate and delayed condition.

Long (1994) explored two components of natural discourse (i.e. pragmatics and discourse style) in the context of conversation embedded in to narratives. Result from three experiments with a total of 180 undergraduates suggested that both components play a role in memory for the surface of the sentences. Recognition memory increased as a function of information about the speaker’s positive and negative attitudes. Recognition memory was enhanced to the extent that such evaluative
information was conveyed by formulaic expressions (i.e. idioms, common expressions) that are characteristic of oral style discourse. Results were confirmed using a recall task. Subject exhibited substantial verbatim memory for evaluative and formulaic expressions.

Schmidt (1994) in an experiment compared the memory for humorous and nonhumorous version of sentences. In order to investigate the memory of humorous and nonhumorous version of sentences, Schmidt (1994) carried out six experiments.

The purpose of experiment 1 was to develop a set of materials which was used to study the effects of humor on memory. The example of humorous sentence was “The only way to keep your good health is to eat what you do not want, drink what you would rather not”. The example of nonhumorous sentence was “The only way to keep your good health is to eat good food, drink healthy drinks, and do healthy activities”. Thus the researcher used 20 sentences for which the humorous and nonhumorous versions differed in their rated humor. Lists containing half humorous and half nonhumorous sentences were presented to 38 subjects in an intentional memory task. Then subjects were asked to recall the sentences. The findings revealed that subjects recalled greater proportion of humorous sentences than nonhumorous sentences. Similarly it was found that larger proportion of words was recalled from humorous sentences than from nonhumorous sentences. These results demonstrated that sentences humor is an effective mnemonic variable.

Experiment 2 was designed to replicate the results of experiment 1 and to determine whether the effects reported in the experiments 1 generalized to a between-subjects design. The impact of humorous sentences and nonhumorous sentences on memory sentences in the same
list was determined in experiment 2. In the between-subjects design, 40 subjects studied an all-humorous list and 40 subjects studied an all nonhumorous list. He used the same sentences as in experiment 1 and constructed these sentences into four lists. Two mixed lists were constructed that contained half humorous and half nonhumorous sentences. The results from these findings were somewhat consistent with the hypothesis that humorous sentences received increased attention when presented in the context of nonhumorous sentences. However, if increased attention is an automatic process, why did the presence of humorous sentences have a negative effect on the recall of the nonhumorous sentences? This result is most consistent with the rehearsal hypothesis in which increased rehearsal of humorous sentences was at the expense of rehearsal on nonhumorous sentences.

Experiment 3 had two main purposes: to determine whether the effects of humor generalized to a cued-recall test and to test the hypothesis that the effects of humor on memory resulted from “surprise” at encountering humorous sentences in the context of a nonhumorous experiment. The comparison between cued and free recall provides another test of retrieval explanations for enhanced memory of humorous sentences. 80 introductory psychology students participated in partial fulfillment of a course requirement. Ten subjects served in each cell of a 2 (labeled vs. not labeled) X 2 (free vs. cued recall) X 2 (list 1 vs. list 2) between-subject factorial. Type of sentence (humorous vs. nonhumorous) was manipulated within subjects. He used the same sentence sets as those used in experiment 1 and 2. One or two words that were central to the meaning of each sentence were selected as retrieval cues and cue words were underlined in all sentence sets. The measures of recall reported in experiments 1 and 2 were again used. He calculated separate ANOVAs
on each measure with label, type of test, list, type of sentence, and sentence serving as variables. The effect of humor on sentence memory occurred independently of labeling and type of memory test. Telling subjects that the sentence they are about to read is humorous does not diminish the effect of humor on memory. Thus, the effects of humor are not due to the surprise resulting from unexpectedly encountering a humorous sentence. In addition, the effects of humor occurred in both free and cued recall. In the cued recall test, recall order was controlled. As a result, the average recall position of humorous material was not less than the recall position of nonhumorous material. The positive effect of humor in cued recall thus eliminated output interference as an explanation of those effects. To further investigate the relationship between rated humor and memory, experiment 4 was carried in which he asked subjects to rate the sentences on humor before taking the memory test. He was also interested in investigating the effects of humor on incidental memory. There were 48 subjects in this experiment selected from the same pool as the subject in experiments 1-3. A within-subjects manipulation of humor was used. There were two sentence lists, each half humorous and nonhumorous. 24 subjects were tested on each list. He used the same 18 sentences as those in experiment 3. Unlike the previous experiments, the sentences in this experiment were videotaped. The sentences were projected in the middle of the screen, with one sentence displayed at a time. Each sentence was displayed for 10 sec. followed by a brief tone. The results of experiment 4 provided some additional support for the role of intentional rehearsal in producing the positive effect of humor on memory. That is, under incidental learning conditions the effects of experimentally defined humor were attenuated. However, intentional rehearsal did not provide a complete explanation of the effect of humor. With the strict scoring criteria, more sentences, defined by the
experimenter as humorous, were recalled than nonhumorous sentences. To the extent that rated humor reflects attentive-getting properties of the sentences, these results supported the role attention and the incongruity hypothesis.

Experiment 5 had two main purposes: to further explore the relation between rated humor and memory and to further explore the effects of humor on the conditional probability of word recall given sentence recall. Another purpose of this experiment was to provide a test of the hypotheses developed by Dixon, Willingham, Strano, and Chandler (1989). Subject were asked to rate their sense of humor. They were then subdivided to determine whether self-reported sense of humor was related to recall. 65 students were sampled from various psychology classes during the summer term in exchange for extra course credit. The subjects were tested in groups ranging in size from 15 to 20. One subject was dropped for failure to complete the rating scale, and two subjects were dropped to balance the number of subjects receiving each list. Thus, he used data from two groups of 31 subjects. Each subject read one of the two lists of sentences in a within-list manipulation of sentence humor. The 18 sentences used in experiments 3 and 4 served as the to-be-remembered items. The 2 sentences from the original list that were not used as sample sentences in the instructions. For all conditions, sentence 1 was presented in its nonhumorous form, and sentence 12 was presented in its humorous form.

This experiment provided a replication and extension of the effect reported in experiment 4. The effect of experimentally defined humor were attenuated in an incidental tasks in which subjects were asked to rate the humor of sentences. This result was obtained when open-ended ratings of humor were used (experiment 4) and when subjects were asked
to confine their attention to the relative humor of sentences occurring in the experiment. Thus, even when humorous and nonhumorous sentences clearly received different ratings of humor, the experiment manipulation of humor had a relatively small effect on retention. However, subject's definition of humor was strongly related to memory, with memory increasing as related humor increased. The analysis of recall as a function of rated humor provided evidence against the hypothesis that subject's rehearsal anchors used in the rating task, and that this rehearsal was responsible for increased retention of humorous material. Instead, the results supported the hypothesis that subject-defined humor affects the amount of attention an individual devotes to a sentence.

The main purpose of experiment 6 was to explore the effects of humor on incidental memory in a between-subject design. Experiment 2 demonstrated that, in intentional learning, the effects of experimentally defined humor were only obtained in a within-subjects design. On the basis of the attention hypothesis developed, one would expect that the same pattern of results would be found in incidental learning. A more interesting prediction was that relative humor in a list should predict relative recall of sentences from the list. From the experimenter's perspective, both the humorous and nonhumorous lists in a between-subjects design were homogeneous. However, from the subject's perspectives each list contained sentences that vary in humor. The amount of attention a sentence received may still vary with perceived humor, and rated humor should still predict recall. To test these predictions, he used a rating task similar to that of experiment 4, and subjects were given an unexpected recall test. The open-ended rating scale was used so that differences between rating of humorous sentence from one group of subjects and ratings of nonhumorous sentences by the
other group of subjects could be detected. 52 subjects participated for extra course credit in a between-subject design with 26 subjects in each group. The same 18 sentences as those from experiments 3-5 were again used. Two lists were constructed and recorded on video tape. One list contained the humorous versions of the sentences, whereas the other list contained the nonhumorous versions. The results provided some additional support for the role of intentional rehearsal in memory for humorous materials. In intentional learning, increased recall of humorous material occurred at the expense of recall of nonhumorous material (experiment 2). This trade-off did not occur in incidental learning. Recall of nonhumorous material with incidental learning in within subjects design (experiment 4 and 5) did not differ from recall of the same material in the between-subjects design (experiment 6). For example, compare the strict sentence recall in experiment 4 with the strict sentence recall in experiment 6. In addition, a significant effect of humor was observed in the experiment 4 with the strict measure of sentence recall. Mean recall of humorous and nonhumorous sentences was .21 and .14 in experiment 4, and .15 and .13 in experiment 6, respectively. The enhanced recall of humorous sentences in experiment 4 was not at the expense of nonhumorous sentences.

Wyer and others (1995) tested the hypothesis that (1) reaction to the statements made during an informal conversation use is likely to be influenced by perceptions of their pragmatic implications as well as literal implications for the persons and object to which they refer; and (2) recipients sensitivity to a statement pragmatic in meaning may depend on their information processing objectives when they encounter the statement and the extent to which the statement violates the recipients normative contextual and stylistic expectations of the messages. The
hypotheses were supported by findings obtained by Wyre et al. (1995) and two studies in which 284 university students were asked to form impression of persons on the basis of information exchanged during conversations in which they did and did not actively participate.

Berk (2000) investigated "does humor in course tests reduces anxiety and improves performance"? Given the paucity of evidence on the effects of humor in tests as well as the limited focus and conditions of previous studies (i.e., primarily multiple-choice items with content-irrelevant humor in undergraduate psychology classes), he designed a different study of humor to target (a) constructed-response item formats, but also included multiple-choice and matching formats, (b) content-relevant humor that is an integral part of what an item is measuring, (c) undergraduates and graduates course in statistics, (d) day and evening graduate courses, (e) students perceptions of the extent to which humor in reducing their test anxiety and improving their performance, and (f) the reliability and validity of these perceptions with multiple cross-validation samples of different students over six years. Seventeen convenience samples of students (N = 695) enrolled in six undergraduate and eleven graduate introductory (level one) statistics courses at the Johns Hopkins University's (JHU) School of Nursing were selected for this study. These two-and three-credit courses met for 15 weeks during each regular semester from fall 1994 through fall 1999. The accreditation board required all students to satisfy these course requirements in the bachelor and master's degree programs. He administered the Humor Effectiveness Evaluation, developed six years ago for a study of humor strategies (Berk 1996; Berk and Nanda 1998), to students at the end of the last class (before the final exam). They completed it anonymously. The confidentiality was preserved until final grades were submitted to the
registrar's office. The same questionnaire and procedure were used for both undergraduates and graduates courses. Only one item on two of the three subscales focused on "humorous material on exams". The students were asked to rate the effectiveness of the strategy in producing two outcomes, "to reduce your anxiety on the exam" to perform your best on the exam". The student perceptions of effectiveness were graduated on a scale of ineffective, somewhat effective, moderately effective, very effective, and extremely effective. These were holistic ratings of effectiveness. Student were not requested to rate the effectiveness of the humor in the different item formats. Their rating reflected the overall impact of the humor on their anxiety and performance. The student ratings of effectiveness for each class were analyzed by computing median ratings for the anxiety and performance outcomes or criteria.

First, the findings indicated that most students feel that humor makes a difference in their test performance; second, this effect was demonstrated with three item formats measuring "real-life" higher-order thinking skills, third, the effect was not idiosyncratic to just one class or level of students; fourth, the effect was consistent over time (six years) and accurate for different cross-validation samples; and finally, the effect was general across undergraduates courses, day and evening graduates courses, and a range of class sizes, students, and exams.

SECTION IIInd

As stated earlier, this section exclusively reviews those studies that highlight the impact of congruent-incongruent version of conversation on conversation memory.
Hunt and Elliott (1980) investigated the role of non semantic information in memory: orthographic distinctiveness effects on retention. Two general issues around which this series of studies was designed involved the theoretical mechanisms mediating non semantic factors in memory and the empirical issue of the source of distinctiveness in the physical stimuli.

The initial experiments were devoted to delineating the circumstances under which the OD effect may be obtained. The rationale for these studies was to introduce manipulations that would reduce or eliminate the effect. The natures of these manipulations then provide a reasonable basis for inferring the source of distinctiveness in the physical stimulus as well as the functional locus of distinctiveness effect in memory. The first experiment, therefore, involved presentation of OD (orthographic distinctiveness) and OC (orthographic common) words in mixed verses unmixed lists. All previous studies reporting OD effects have manipulated OD in mixed lists. It is possible that the immediate list content of OD encodings of the OD words. If such is the case, no OD effect would be expected following unmixed presentation of OD and OC words. An interaction of this form, an OD effect in mixed lists but no effect in unmixed lists, would suggest a very simple differential attention interpretation of OD effect. That is, OD words in mixed lists may be so different as to “surprise” the subject and attract additional processing time, perhaps in the form of additional rehearsal, at the expense of OC words. In unmixed lists, the competition for processing time would not exist, and no OD effect would be obtained. The subjects were 60 undergraduate volunteers randomly assigned to one of four conditions, each containing an equal number of subjects. The four conditions were based on an orthogonal combination of two variables, list type and word
type. The two levels of word type were OD and OC words. List type was represented by mixed and unmixed presentation of OD and OC words. In order to directly compare performance as a function of list type, Grice’s (1966) design for direct comparison of between and within-subject variables was used. According to this design, the unmixed-list conditions involved lists containing only OD or OC words. The mixed-list groups were given lists containing an equal number of OD and OC words; however, for one group, performance was analyzed only on OD items and for the other group, performance was analyzed only for OC items. Thus, the design was a 2(list type) X (word type) factorial with four groups: unmixed OD (UOD), unmixed OC (UOC), mixed OD (MOD), and mixed OC (MOC). The stimuli were obtained by selecting 50 words that appeared to be orthographically distinctive. These words were matched with OC words on the basis of Thorndike-Lorge frequency, associative meaningfulness, number of letters, and first letter. The meaningfulness data on these words were obtained by asking 50 subjects to write as many associations as possible to each word in a 30-sec period. The words also were scaled for distinctiveness by a second group of 50 subjects who rated the words for visual “weirdness” on a 5-point scale where 5 = very weird. As the result of these scalings, 20 OD-OC words pairs were selected for which the mean meaningfulness values were 5.11 and 5.12 for OC and OD words, respectively. The average distinctiveness ratings were 2.43 for OC and 4.06 for OD words. The following OC-OD examples were included in the lists: leaky-lymph, kennel-khaki, airway-afghan. The 20 OD words made up the study list for the UOD condition, and the 20 OC words were presented in the UOC condition. For the MOD and MOC conditions, the 20 OD-OC pairs were arranged randomly into two lists of 20 OD and OC words. Each of these two lists was presented to half of the MOD and MOC subjects. In each condition, two different
presentation orders of the lists were used. With each order, the words were arranged randomly, with the restrictions on the mixed lists that no more than two OD and OC words would occur contiguously and that an equal number of OD and OC words would occupy the last six serial positions.

The effect of list type on recall was not reliable. The total recall from mixed as compared to unmixed lists was numerically equal. Word type had a significant effect on performance. However, inspection of the means suggests that the beneficial effect of OD words was limited to the mixed-lists comparison. A Duncan's Range Test revealed a significant difference between MOD and MOC recall but no reliable difference between UOD and UOC. Thus, the data clearly demonstrated that OD words facilitate performance only when presented in a list containing OC words. They also found the superior recall of OD words in mixed list without a reliable difference between OD and OC words in unmixed lists. This interaction implicated the immediate test context as an important condition for OD effects on memory.

Experiment 2 was designed to test directly the time-sharing explanation of OD effects. In order to test this hypothesis, subjects were allowed to view each word in a mixed OD-OC list for as long as they thought necessary to remember the word. The amount of time spent on each word was then considered in relation to recall performance. In addition, the manipulation designed to provide initial information concerning the aspect of a word that make it distinctive for memory purposes and second purpose of experiment 2 was to provide information on the source of physical distinctiveness by comparing recall of capitalized and un-capitalized words. 30 undergraduate volunteers were randomly assigned to one of two groups containing 15 subjects each. The
two conditions differed only in that all of the words were capitalized in one (UC condition), and were in lowercase for the other (LC condition). In each condition, the same lists of words, half OD and half OC, were presented. Thus, the design was 2 (type font) X 2 (word type) mixed factorial. The words were the 20 OC-OD pairs used in the mixed-list conditions of experiment 1. Two lists of 10 matched pairs each were used with half of the subjects in each of the type font conditions receiving one of the lists. Each of the two 20-word lists was presented in two different random orders.

Neither the list variable nor the order of presentation variable had any effect on recall, and both were ignored in subsequent analyses. The amount of time spent viewing a word was not reliably affected by either type font or orthographic distinctiveness. Both variables, however, had a significant effect on recall performance. The pattern of these conditional probabilities is the same as in experiment 1, but in the present experiment, the differences were not reliable. The results showed an OD effect in the LC condition in the absence of a reliable difference in the amount of time devoted to OD and OC words. The pattern of results from contingent recall scores and study time suggests that enhanced recall of OD words is not satisfactorily explained by a time-sharing hypothesis.

In the third experiment, the major variables of the first two experiments were combined in a single design. Subjects received OD and OC words in either mixed or unmixed lists; the words were either all capitalized or all in lowercase. Based on the results of the first two experiments, one would expect the OD effect to occur only in the case of mixed-list presentation of lowercase words. The absence of contrast with OC words in the unmixed lists should preclude enhanced recall of OD words. The lack of visual irregularity in the mixed lists with uppercase
words also should eliminate the OD effect. There is little reason to believe that these latter manipulations would differentially affect performance. The subjects were 88 introductory psychology students who were randomly assigned to one of eight different conditions. The conditions were formed by an orthogonal combination of list type (mixed vs unixed presentation of OD), type font (upper vs lower case), and word type (OD vs OC). The list type and word type variables were combined in Grice’s (1966) design in the same fashion as described for experiment 1. Type font was manipulated entirely between subjects. The eight groups so formed were mixed-list-lowercase-OD, mixed-list-lowercase-OC, unmixed-list-lowercase-OD, unmixed-list-lower-OC, unmixed-list-uppercase-OD, and unmixed-list-uppercase-OC. Each of the four unmixed-list conditions contained 15 subjects, and 7 subjects were assigned to each of the four mixed-list condition. Procedural details, including stimulus materials, were the same as in experiment 1 with the exception of the capitalization of all words in two of the mixed and unmixed lists.

Results suggested that neither word type, list type, nor type font had a main effect on performance and word type did interact with list type. The effect of word type also interacted marginally with type font. Most important, however, was the reliable interaction between list type, type font, and word type. This triple interaction was the primary prediction of the experiment. The findings suggested that the effect is due to enhanced recall of lowercase OC words in mixed lists.

Subsequent analyses confirmed this interpretation of the triple interaction. In separate analyses of the mixed and unmixed list data, the mixed list yielded no effect of type font or, most importantly, word type, nor did the two variables interact. The mixed-list analyses were
considerably different, revealing a main effect for word type and a reliable word Type X Type Font interaction.

The purpose of the fourth experiment was to examine the relative contributions of phonemic and orthographic attributions to the OD effect by comparing the OD effect following aural and visual presentation of the same lists. The comparison was based on the assumption that auditory presentation encourages phonemic encoding and visual presentation encourages orthographic encoding. The subjects were 30 undergraduate volunteers who were randomly assigned to one of two groups of 15 subjects each. The two groups differed only in that list presentation was visual for one and aural for the other. In each condition, the word lists were the same as those of experiment 2 and 3. Although it was not feasible to equate the auditory and visual exposure rate for each word, the total presentation time in the two conditions remained constant 10 word/min. In addition, the amount of time allotted to each word was 6 sec. In the visual condition, each word exposed on slides for 3 sec followed by 3 sec of a blank slide prior to presentation of the next word. In the auditory condition, the words were presented via a tape recorder at the rate of 1 word every 6 sec. All subjects were given the same standard free-recall test with a maximum time limit of 3 min was administered.

Recall of OD words exceeded that of OC words but the OD effect was not consistent across presentation mode as evidenced by a reliable presentation Mode X Word Type interaction. The source of interaction is clear from inspection of the means. The OD effect was present with visual but not with auditory presentation. The main effect of presentation mode was not reliable. These results are quite straightforward in implicating visual presentation in the OD effect. Although it's highly likely that visual presentation is accompanied by covert pronunciation, it
does not seem reasonable to argue that visual presentation sensitizes the
subject to phonemic attributes to a greater extent that does auditory
presentation, consistent with the finding in the second study that OD
effects are eliminated by capitalizing the words, distinctive orthography
appears to arise from the visual pattern presented by the letter
combinations.

The experiment fifth addressed the questions concerning the
mechanisms mediating OD effects through an examination of the
interaction of semantic processing and non-semantic structure. The
subjects were 80 undergraduate volunteers who were randomly assigned
to one of five equally sized groups. Each group performed different
orienting tasks as follows: In seen group the subjects simply viewed the
words, in the write group subjects wrote each word as it appeared, in the
association group subjects wrote the first word that came to mind when
each list word appeared, in the pleasantness group subjects rated each
word for pleasantness on a 5-point scale, and in the sentence group
subjects wrote a sentence containing each list word. The word lists were
those used in experiment 2 and 3, and each word was exposed for 5 sec.
The words were printed in lowercase and all subjects were given
intentional memory instructions. All other details of procedure were the
same as experiment 1. Orienting Task X Word Type interaction was not
reliable indicating that the OD effect was consistent across the orienting
tasks. The OD effect itself was reliable. Thus, within the range of these
orienting tasks, OD words had a consistent positive effect regardless of
the particular orienting tasks. Overall performance was highest in the
sentence condition, and the same Newman-Keuls analysis revealed no
reliable differences among the other orienting tasks. Inspite of the
differential effect of orienting instructions, the relative effect of OD
remained essentially constant over conditions. The size of the OD effect is virtually the same in the four orienting task conditions. It is equally interesting to note that the reduction in the magnitude of the OD effect relative to the seen condition is due in all cases to an increase in the percentage of OC words recalled and a corresponding decrease in the percentage of OD words recalled. Thus, this analysis indicates that, although qualitatively different orienting tasks produce differential recall, the size of the OD effect is independent of the type of orienting task.

The rationale of the last experiment was to see if distinctive orthography has the same relative effect on recognition and recall as does meaningfulness. The subjects were 58 undergraduate volunteers, an equal number of whom were assigned to recognition and recall test conditions. For both test conditions, meaningfulness (m) and OD were manipulated within subjects. Each list contained an equal number of OD and OC words, and within levels of OD, half of the words were high m and half were low m. The lists were constructed by first selecting the 20 OD and 20 OC words from the previous studies as the high m words. An additional group of 20 OD and 20 OC words matched for length and Thorndike-Lorge frequency were selected from Zechmeister's (1969) norms. These additional words were selected on the basis of their low m value. Since the study was concerned with interactions between OD and m, it was necessary to provide additional scaling information on the two variables. First, the words selected from Zechmeister's norms were rated for m by 30 students who wrote as many associations as possible to each word in a 30-sec period. The second scaling procedure involved examining distinctiveness within levels of meaning and was necessitated by the fact that low m words were selected from Zechmeister and high m words from our own scaling. Thus, all 8 words were rated by an
additional 30 subjects on a 5-point distinctiveness scale where 5 = highly
distinctive. The 80 words then were divided into separate lists. In recall,
four lists of 20 words each were constructed. Each list contained an equal
number of high m-OD, high m-OC, low m-OD, and low m-OC words. In
recognition, two 40-word lists were constructed, containing an equal
number of each word type. For the recognition test, all 80 words were
presented, and consequently, an equal number of each type of word was
present on the test.

The recognition and recall results were analyzed separately because
of traditional concerns about unequal scale values. The relative effects of
the primary variables, however, were sufficiently different to encourage
confident conclusion based on relative comparisons. It was found that
meaningfulness was directly related to recall and the main effect of OD
was not reliable. The m X OD interaction was significant. This
interaction was due to superior recall of OD words at high m levels and
no reliable differences in recall of OD and OC words at low levels of m.
Newman-Keuls analysis led to this interpretation of the interaction. The
only reliable influence on recognition performance was exerted by OD.
Recognition was higher for OD than OC words. The OD effect in
recognition was consistent across levels of m, as evidenced by the
unreliable OD X m interaction. Finally recognition was not affected by
m. Thus, the recognition and recall data differed in two aspects, the most
important of which was the reliable OD effect at both m levels in
recognition, but only with high m words in recall. In addition, level of
meaning had a direct effect on recall but no reliable influence on
recognition performance.

Kemper and Thissen (1981) conducted a study on memory for the
dimensions of requests. The present research was undertaken to scale
requests dimensions of politeness and directness and to determine whether requests that violate conversational conventions governing the form of requests are more memorable than requests that conform to the rules. Different forms of requests were placed on interval scales of politeness and directness by applying nonmetric multidimensional scaling (MDS) to similarity judgments of the effectiveness of the requests. In this study, the stimuli were 19 different forms of requests and the subjects were asked to judge the relative effectiveness of those forms. This procedure did not require the subjects articulate the bases for their judgments nor did it restrict them to predefined dimensions of similarity. 80 naive native speakers of English were recruited from introductory psychology courses. Ten directive request forms were selected from among those listed by Ervin-Tripp (1976a, 1976b). Initially, a pair of sentences consisting of an action imperative and a related need-assertion was prepared. Form the imperative, five new forms were constructed by (1) combining the imperative with please, (2) embedding it under the two hedges I think you should and why don’t you and (3) forming an affirmative or negative interrogative by embedding the imperative under Do you think should or Don’t you think you should. Three new forms were created from the need-assertion by (1) embedding it under I think and (2) embedding it under either Do you think or Don’t you think. These transformations of the imperative and need-assertion resulted in a set of request that included both affirmative and negative questions, hedges, and direct and indirect requests. Two general themes and four examples of each theme were used to create eight sets of requests. The two themes and the specific topics were: (1) requests for simple actions that are, raking the leaves, watering the plants, making coffee, and turning off the television, and (2) requests were requiring the exchange of modest sums of money, that is, lending five dollars, buying shampoo or milk or a new
typewriter ribbon. For each set of requests, 45 pairs of sentences were constructed so that each of the 10 directive forms was paired with each of the others. Eight sets of response booklets were constructed by randomly ordering the lists of 45 pair of requests was accompanied by a 19-point rating scale of relative effectiveness marked from "-9" to "+5". The subjects were instructed to rate the relative effectiveness of the two sentences labeled A and B. Negative values were to be assigned to these pairs in which sentence B was judged to be more effective than sentence A in securing the intended result; positive values were to be assigned when sentence A was judged to be more effective than sentence B. In addition, each subject was asked to rate a second set of 10 requests on six bipolar adjective scales. They were to rate these requests on 9-point scale of politeness, assertiveness, intimacy, naturalness, clarity, and importance. These ratings were obtained in order to provide convergent evidence for the interpretation of the MDS results. The obtained two-dimensional solution for the request forms was essentially identical to the nonmetric MDS solution, confirming that this configuration is not specific to a single method of data analysis. The correlation between the dissimilarity data and the obtained distances in this solution is .56; approximately 20% of the total variance was accounted for by dimension 1 and 11% by dimension 2. Dimension 2 was not related to any of six bipolar scales; it was interpreted as a dimension of directness. Need-assertive and questions are located at the negative end of this dimension while the imperative forms are located at the positive. Examination of the subject weight space revealed that individual subjects differed in their reliance on these two dimensions in a manner related to the theme of the requests placed less weight of politeness (dimension 1) than did subjects who judged money oriented requests. The opposite was true for
directness (dimension 2): the significantly greater than that for money-oriented requests.

The second experiment was undertaken to investigate the recall performance under different forms of request. To achieve this, the scale values of the 10 directive forms were used to predict recall of the different forms requests. It was hypothesized that requests that violate conversational conventions would be more salient and more memorable than requests that conform to the rules. Eighty naïve native speakers of English were recruited from introductory psychology classes. The 10 directive forms from Experiment I were again used. These forms were combined with 20 different topics to create 200 different requests. The topics involved either simple actions or the exchange of modest amounts of money. These requests were then used as the captions of cartoons. For each set of 10 requests about a particular topic, two cartoons were selected. These cartoons portrayed two (or more) individuals in a social context that clearly specified the relative status of speaker and addressee. The two cartoons differed in whether the speaker was of higher or lower social status relative to the addressee. Some of these speaker-addressee pairs were: waiter-diner, attorney-judge, clerk-customer, commoner-king, secretary-employer, monk-abbot, sinner-devil. Twenty additional cartoons and captions were selected as filler items. Each subject received a booklet of 30 cartoons and captions. They were asked to rate, on a 7-point scale, the funniness of each caption. The general theme of the target requests and the status of the speaker varied between groups of subjects. There were four groups of subjects: those who rated (1) action-oriented requests by high-status speakers, (2) action-oriented requests by low-status speakers, (3) money-oriented requests by high-status speaker and (4) money-oriented requests by low-status speakers. Each subject rated
the humor of one request of each directive type (e.g., one imperative, one need-assertion) and each of the 200 different requests was rated by two subjects. After a 20-minute interval during which the subjects carried out an unrelated task, they were asked to recall the captions for the 10 targets cartoons, which were provided as recall prompts. Verbatim recall of the original captions was stressed. The funniness ratings were not significantly related to any design feature. There were no substantial differences between action-based and memory-based requests and this distinction was ignored for the remaining analyses of the requests forms. Nor were there differences between requests by low and high-status speakers for overall recall. However, further analysis revealed that verbatim recall was a function of speaker status, politeness, and directness. When the data were pooled and an attempt was made to use the same quadratic function for the easiness-of-recall for both sets of subjects (those with high and low-status speakers), the resulting likelihood ratio test for equality of the quadratic functions for the two groups was extremely large, indicating a highly significant interaction between speaker status and the effects of politeness and directness. Part of that interaction was specifiable in terms of the complexity of the model required for each group. The single-degree-of-freedom likelihood ratio $X^2$ tests indicated that the contribution of the quadratic politeness term was significant for the low-status speakers but not for the high-status speakers and the quadratic directness term was significant for the high-status speakers but not for low-status speakers. Subjects who studied cartoons with high-status speaker, 66 and 63% of the substitutions of old and new forms, respectively, involved such deletions; for the other subjects, 69 and 68% of the old and new forms resulted from such operations. The new directive forms were no more polite than the originals (high-status and low-status speakers) but they were more direct than the originals.
Substitutions of old or new directive forms did not alter the politeness of the original requests but shifted the request to a more direct form. The investigators concluded that a speaker must select form of a request that is appropriate to the situation. As Ervin-Tripp and Lakoff have observed, this selection depends on conversational conventions governing politeness and directness. Although the precise rules have not been formulated, speakers must consider such factors as: setting (at home vs. at work), magnitude of the requested service (a loan of five dollars vs. one of five thousand dollars), asymmetries in age or rank, familiarity, and expected role responsibilities. A speaker's request may fail if a superior is offended by a direct or an impolite request, if the addressee does not interpret an indirect request as such, or if the addressee chooses not to comply with an indirect request. The benefits and risks of being polite must be balanced with those of being direct: different forms of requests enable speakers to offset directness with politeness. Scale values of politeness and directness obtained from judgments of the probable effectiveness of requests are highly related to the memorability of those requests. Thus a speaker who violates the conversational conventions governing the form of requests, who does not appropriately balance politeness and directness, will pay two penalties: not only may the request fail but the inappropriate request will be accurately remembered.

Crocker and Vitkus (1983) conducted a study entitled "Attributions, Attention, and Person Memory: Processing Congruent and Incongruent Information". They tried to resolve an apparent paradox of person perception. The paradox is this: Impression of people (and social beliefs in general) is resistant to change. Information that contradicts an initial impression of a person has relatively little impact on the
impression (Schneider, Hastorf, and Ellsworth, 1979). At the same time, research on memory for social information indicates that information that contradicts an impression of another person is particularly likely to be recalled (Hastie and Kumar, 1979; Hastie, 1980; Srull, 1981). If information that violates our impression of another is more likely to be recalled than information that confirms an impression.

They conducted an experiment to examine more directly the amount of attention congruent and incongruent information receives as a function of the causal attribution given to the item. The experiment replicated the procedure of Crocker et al. with a few modifications. Subjects received information about the behavior of a target person which was congruent with their initial impression of him with the exception of one item. The target item was either congruent or incongruent with the initial impression and was attributed to a dispositional or a situational cause. In a departure from the procedure of Crocker et al., the behavioral information was presented on a CRT screen controlled by a microcomputer, and subjects controlled the length of time each item appeared on the screen. The viewing time, which was the operationalization of amount of attention, was recorded by the computer. In addition to viewing time, we measured recall for the information, and subjects' impression of the target person.

They found that when the target item was incongruent, subjects remembered more non-target congruent items than when it was congruent, regardless of how the target item was explained. The relationship between attention and impression, which showed an unexpected pattern for the in-congruent, dispositionaly attributed item. This pattern of results suggested that the relationship between attention and weight in impression is not as straight forward as Fiske (1980)
previously suggested. The result of analysis of variance showed very similar patterns for looking time, recall, and impressions, but correlational analyses indicated that within conditions longer looking items are not always related to an item having more impact on impressions. These results suggested that the information that one attends to will not always be over represented in judgments (C.F. McArthur, 1981; Taylor and Fiske, 1978). Under some circumstances, greater attention to information can lead to less impact on impressions.

The correlations between recall and impressions also failed to show the expected pattern. For congruent items, the relationship between recall and impression was stronger for the situational attributed item than for the dispositionally attributed item. For incongruent items, both these attributed to the situation and those attributed to dispositional causes, recalling the item was related to less impression change (i.e. impression more like the initial impression). The impressions are not simply derived from subject recall of the target item.

Rajaram (1998) investigated the effects of conceptual salience and perceptual distinctive on conscious recollection. This study examined the hypothesis that processing the salient or distinctive attributes of stimuli leads to the experience of remembering. The constructs of salience and distinctiveness have often been used interchangeably in the memory literature (Schmidt, 1996) and the extant studies do not explicitly differentiate between them. Distinctiveness has been variously defined as the property that (1) separates items or events that share few rather than many features with other items in memory (Nelson, 1979), (2) emerges due to differences rather than similarities among items (Hunt & McDaniel, 1993; Schmidt, 1996), (3) arises from presentation of isolated items in the content of background items (Hunt & Mitchell, 1982), or (4)
characterizes “events…… that are incongruent with active conceptual framework, or that contain salient features not present in active memory” (Schmidt, 1991). Thus, a common thread among these definitions is the differences among items those presumably uniquely specify some item. or the salience of items that make them stand out from among the background items.

In this study, the role of salience or distinctiveness in producing the experience of remembering was tested with variables that can a priori be designated as generating conceptual salience (experiment 1) or perceptual distinctiveness (experiment 2). The main objective of experiment 1 was to examine the effects of semantic or conceptual saliency on remembering. As noted, facilitatory effects of conceptual manipulations on remembering have been documented in a number of studies (e.g. Gardiner, 1998; Rajaram, 1993). Note that in most of the previous experiments, semantic or conceptual encoding was contrasted with non-semantic encoding. For example, encoding of meaning was compared with encoding of phonemic features (Gardiner, 1988; Rajaram, 1993). To understand the nature of conceptual attributes that would specifically influence remembering, the present experiment contrasted two types of conceptual encoding. This manipulation was achieved by using homograph (e.g. chest); subjects focused on the salient conceptual attribute of the stimulus in one condition (the dominant meaning, e.g., body part-CHEST) and on the non-salient conceptual attribute in another condition (the non-dominant meaning, e.g. cabinet-CHEST). 36 undergraduates from the state University of New at Stony Brook participated for credit in partial fulfillment of course requirement. At study, subject encoded the dominant or the non-dominant meanings of homographs in a within-subjects design. Accuracy data for explicit
memory and retrieval experience of remembering were measured using a recognition memory paradigm. Altogether, 60 homographs served as critical items in the experiment. This set of 60 homographs was selected from the Nelson, McEvoy, Walling and Wheeler (1980) norms. According to the Nelson et al. norms, in this item set the average frequency with which the dominant meaning was produced by subjects (35) differed reliably from frequency for the non-dominant meaning. The dominant and non-dominant meanings of the homographs were biased with phrases constructed and reported by Rajaram et al. (1998). In addition to the 60 homographs, one additional set of 40 words was used as fillers in the study lists and a second additional set of 40 words was used as fillers in the construction of the recognition test. Three study lists were constructed for purposes of counterbalancing in such a way that, across lists, each homograph served as a studied dominant target, a studied non-dominant target, or a non-studied item. Within each study list, 20 homographs were preceded by phrases that biased their dominant meaning (e.g. yellow-name of color), 20 homograph were preceded by phrases that biased their non-dominant meaning (e.g., yellow-cowardly or chicken). In addition, 40 phrase-word fillers (e.g. sadness: an emotion) were included to conceal the experimental manipulation from subjects. The 80 phrase-word pairs were arranged randomly with respect to the study conditions in the study booklets. In addition, a scale displaying numbers from 1 to 5 was placed in front of each phrase-word pair to enable subjects to rate the pair. The test list consisted of 40 studied homographs (20 studied for their dominant meaning and 20 for non-dominant meaning), 20 non-studied homographs, and 20 filler words to maintain 1:1 study-test ratio from the subject’s point of view. Two test lists were constructed to create two random orders of stimuli with respect to the conditions. In the test booklet, each word was followed by two
blank lines in that raw. The first line was provided to indicate whether or not the word had been studied earlier. The second line was provided to indicate whether the subject remembered or knew the item from the study list. In addition to the study and test lists, a cover sheet was constructed with a window that exposed one row at a time. Study exposure was controlled through a tape on which a female voice prompted with "next" every 5 sec.

For the overall recognition memory measured, there was a significant advantage for homographs encoded for their dominant meaning compared with homograph encoded for their non-dominant meaning. The critical question posed in this experiment was whether this advantage for the dominant meaning would also appear in the remember judgments. The results showed that subjects gave a significantly higher proportion of remember judgments to homographs encoded for their dominant meaning than to homographs encoded for their non-dominant meaning. For know responses, the reversed difference was not statistically significant. These findings clearly revealed that processing of salient conceptual attributes leads to a higher proportion of remember responses than processing of non-salient conceptual attributes.

The aim of experiment 2 was to examine the hypothesis that remember judgments increase as a function of processing the distinctive attributes of the stimuli even if the processed dimension is perceptual in nature (Rajaram, 1996). A new group of 32 undergraduates from the State University of New York at Stony Brook participated for credit in partial fulfillment of course requirements. Orthography of words was manipulated at two levels, distinctive and common, in a within-subjects design. Accuracy of recognition memory and remember-know response frequencies were measured in a recognition memory test. A total of 64
blank lines in that raw. The first line was provided to indicate whether or not the word had been studied earlier. The second line was provided to indicate whether the subject remembered or knew the item from the study list. In addition to the study and test lists, a cover sheet was constructed with a window that exposed one row at a time. Study exposure was controlled through a tape on which a female voice prompted with "next" every 5 sec.

For the overall recognition memory measured, there was a significant advantage for homographs encoded for their dominant meaning compared with homograph encoded for their non-dominant meaning. The critical question posed in this experiment was whether this advantage for the dominant meaning would also appear in the remember judgments. The results showed that subjects gave a significantly higher proportion of remember judgments to homographs encoded for their dominant meaning than to homographs encoded for their non-dominant meaning. For know responses, the reversed difference was not statistically significant. These findings clearly revealed that processing of salient conceptual attributes leads to a higher proportion of remember responses than processing of non-salient conceptual attributes.

The aim of experiment 2 was to examine the hypothesis that remember judgments increase as a function of processing the distinctive attributes of the stimuli even if the processed dimension is perceptual in nature (Rajaram, 1996). A new group of 32 undergraduates from the State University of New York at Stony Brook participated for credit in partial fulfillment of course requirements. Orthography of words was manipulated at two levels, distinctive and common, in a within-subjects design. Accuracy of recognition memory and remember-know response frequencies were measured in a recognition memory test. A total of 64
critical words were used in this experiment. 32 words in this set were orthographically distinctive (e.g. subpoena, calypso, graw, lymph) and the remaining 32 words were orthographically common (e.g. sailboat, cookie, grit, loser). 16 orthographically distinctive and 16 orthography common words were taken from the materials published by Hunt and Toth (1990). In order to increase the list length, a norming study was conducted to collect an additional set of 16 orthographically common words. In this norming study, 110 words were printed in a booklet. The orthographically distinctive and common words reported in the Hunt and Toth (1990) article were included in this pool in order to validate the ratings given for the remaining words. A new group of 75 subjects who did not participate in the two experiments reported in this article took part in this norming study. Following Hunt and Toth’s procedure, subject in this norming study were instructed to rate on a scale from 1 to 5 the visual “weirdness” of each word (1 = not weird, 5 = very weird). The set of critical words was selected in this norming study in a way that the mean weirdness rating given by subjects for the new set of words differed significantly for orthographically common words (1.67) and orthographically distinctive words [3.37; t (30) = 11.84, SE = .14]. Furthermore, the orthographically distinctive and common words in the new set of materials (as in Hunt and Toth) were matched for frequency. In order to ensure that each of the 64 words served as studied and non-studied items across subjects, two study lists were prepared. Each study list, presented in booklet form, consisted of 16 orthographically distinctive and 16 orthographically common words. Care was taken to ensure that half of the words in each orthography type were drawn from the Hunt and Toth (1990) material and the other half from the new set of materials created for this experiment. Study words within each booklet were arranged randomly with respect to orthography. Furthermore, two
random orders of each study booklet were prepared and presented equally often across subjects. The test booklet consisted of all 64 words, 16 studied orthographically distinctive words, 16 non-studied orthographically common words, and 16 non-studied orthographically common words. Two random orders of the test booklet (with respect to orthography as well as study status of materials) were prepared and presented equally often across subjects. In the test booklets, two blank lines were provided in front of each word. The first line was provided for recognition (Y/N) responses and the second line was provided for remember-know responses. Finally, a cover sheet with a window and a tape with 5 sec prompts in a female voice were also used in this experiment.

The proportions of false alarms for orthographically distinctive and orthographically common words were quite low and within the range reported in the literature. In overall recognition, performance was significantly better for orthographically distinctive than for orthographically common. More critically, the effect of this perceptual manipulation was evident only in remember responses: significantly more remember responses were made for orthographically distinctive than for orthographically common words. The orthography manipulation had no effect on know judgments. The results from this experiment confirmed that remember responses are influenced by perceptual variable as well. Furthermore, these results demonstrated that remember responses increase as a function of processing the distinctive attributes of the encoded stimuli.

Hunt (2003) conducted a study to investigate the two contribution of distinctive processing to accurate memory. This research was provoked by levels of processing and led to the evolution of distinctive processing
as the progeny of depth of processing (e.g. Eysenck, 1979; Humphreys, 1976; Hunt and Einstein, 1981; Jacoby and Craik, 1979). Why would a manipulation that has such a dramatic effect on correct rejection have no effect on correct acceptance? The specific goal of the present study is to address this question by manipulating both forms of distinctive processing.

The target list items shared an obvious dimension of similarity, distinctive processing at the item level should facilitate recognition of the targets themselves. Although the items in the second list of Dobbins et al. (1998) shared the obvious spatial/temporal similarity of the list, there was no manipulation designed to encourage processing of differences aligned with that similarity and thus no item-based distinctive processing of the second list. To implement item-based distinctive processing, categorized lists were used in this experiment. This first experiment was a replication of Dobbins et al. (1998) in that participants saw a pre-exposure list and study list on which either the same or a different orientating task was performed. A recognition test for the second list items followed and included target items, pre-exposed distracters, and new distracters. The participants were 96 volunteers from introductory psychology. The design called for an equal number of participants in each of four conditions defined by the combination of orienting tasks on the pre-exposure and study lists. With category judgment (c) and pleasantness rating (p) as the two orienting tasks, the four condition were CC, CP, PC and PP. Due to an error in assignment, 25 participants were in the CC condition and 23 were in the CP condition. Conditions PC and PP consisted of 24 participants each. Three types of items appeared on the recognition test, targets from the study list, pre-exposed lures, and new lures. Target items and pre-exposed lure occurred either one or three
times prior to test. Item repetition was a within participant manipulation. A pool of 90 words, 9 instances from each of 10 categories, comprised the materials for the experiment. All words were of intermediate frequency in the Battig and Montague (1969) category production norms. The 9 instances from each category were arbitrarily assigned to groups of three and each of these word-triads, served as the targets, the pre-exposed lures, or the new lures. Each word-triad appeared equally often in each of those roles across participants. Pre-exposed lures and targets occurred either one or three time prior to test. Half of the pre-exposed lures were shown three times in the pre-exposure list. Half of the target items were shown twice in the pre-exposed list and half of the targets were absent from the pre-exposed list. Regarding of pre-exposure, targets appeared once in the study list. Thus, the repetition factor describes the number of times an item was experienced prior to the memory test. This repetition factor was between category such that all of the items from a given category occurred one or three times. The items were counter-balanced such that across participants the word-triads from each category appeared equally in the two repetition conditions. Both the pre-exposure and study list were presented blocked by category, and the order of categories with lists was randomized for each participant. The recognition test list contained 90 words, 9 instances from each of the 10 categories. Among these 90 words were 30 targets from the study list, 30 pre-exposure lures, and 30 new lures. The test lists were randomly ordered with the restriction that no more than two items from a category occur contiguously.

The first measure indexes discrimination of study items from pre-exposed distracters in that hits are corrected by false alarms to distracters shown once were subtracted from hits on study items shown one and false
alarms to distracters shown three times were subtracted from hits to study items that had been shown three times. The second measure of accuracy, hits corrected by false alarms to novel distracters, was computed by subtracting false alarm rates to novel distracters from hit rates to novel distracters from hit rate collapsed over frequency of presentation. This measure of accuracy was affected reliably by orienting condition. Subsequent analysis showed that the PP condition was more accurate than CC, PC and CP. The CP condition was marginally more accurate than CC in both cases. No other differences were reliable. Thus, the accuracy of discriminating targets from novel distracters showed a very different pattern than accuracy of discriminating targets from pre-exposed distracters. More detailed examination of the overall hit and false alarm rates to pre-exposed distracters was provided by an analysis of hit and false alarm rates as a function of experimental condition and frequency of presentation. Analysis of the hit rate and false alarm rate to pre-exposed distracters revealed that the main effects of orienting condition and item type were qualified by an interaction between the two variables. Simple effects analysis showed a main effect of orienting condition on both hit rate and false alarm rate. In addition to the effects of orienting condition and item type, the overall analysis revealed an effect of frequency of presentation, which interacted with orienting condition. Simple effects analyses via independent group t-tests showed the source of the interaction to be an increase in “old” responses from 1 to 3 presentations only in the CC group. Frequency of presentation did not interact reliably with item type nor was the triple interaction significant.

The design of second experiment was identical to that experiment 1. A new sample of 98 participants was drawn from the pool of volunteers in introductory psychology. Condition CC and PP contained
24 participants each while condition CP and PC contained 25 participants. The materials were the same as those used in the first experiment. The procedure also was the same as the first experiment except that the final test was free recall. Participants were instructed to write all of the words they could remember from the second list they saw. The instructions included the warning that some of the words on the second list had appeared on the first list, but they were only to recall a word if it had appeared in the second list. Five minutes were allotted to the recall test. The analysis of recall parallels that of the recognition analysis from experiment 1. Accuracy will be presented first in the form of correct recall of list items minus intrusions of pre-exposed distracters. Unlike the recognition data, accuracy cannot be corrected for recall of novel distracters because we do not know the population of such distracters in recall. The accuracy analysis was followed by analysis of correct recall from the study list (hits) and incorrectly intruded recall from pre-exposed items (false alarms). Accuracy increased with frequency of presentation, orienting condition also exerted a significant effect on accuracy and an effect that was consistent across frequency of presentation as indicated by the lack of frequency X orienting condition interaction. Individual comparisons of the orienting condition effect were conducted with independent t-tests. These comparisons showed reliably higher accuracy in the CP condition than in either the CC or PP conditions which themselves did not differ.

Geraci and Rajaram (2004) investigated the distinctiveness effect in the absence of conscious recollection: Evidence from conceptual priming. The goal of their experiments was to test the hypothesis to determine whether conscious awareness was necessary for obtaining the effects of distinctiveness on memory.
Experiment 1 was designed to separate the contribution of conceptual processing at study and test by using an implicit test that recapitulates the evaluative process presumed to mediate this distinctiveness effect. In this study, distinctiveness was created using study lists that contained categorically isolated items. At test, participants were given an implicit test of category verification. They used this test to try to mimic the hypothesized evaluative processes that participants carry out at study when they encounter an incongruent item. They used both an explicit and implicit version of the category verification test. In the explicit version, participants saw a category label and a word and judged whether or not they had studied the word. In the implicit version, participants also saw the category label and word. However, in this version, participants were simply told to indicate whether or not the word was a member of that category. In the present paradigm, the study conditions are designed to lead the participants to note that the isolated items does not fit within the context without requiring an integration of the incongruous item within the context. Thus, if the process of evaluating mismatch at study is recapitulated at test, participants should be faster to engage in this same evaluation for the isolated items at test compared to the non-studied and non-isolated items. 166 SUNY Stony Brook students participated for research credit. 52 participants received an explicit category verification test and the remaining 84 participants received the implicit category verification test. Of these 84 participants, 60 were categorized as “unaware” of the connection between study and test based on the data from the post-test questionnaire. Data from these “unaware” participants were included in the implicit analysis. The experiment was a 2X2X2 mixed factorial design. Test type (explicit vs. implicit) was the between subjects variable and study status (studied and non-studied) and item type (common and distinct) served as the within
subjects variables. 24 categorized lists of 8 words each were derived from the Battig and Montague (1969) and Shapiro and Palermo (1970) category norms. All items were chosen from among the top 10 most common examplars for each category. For each list, one of the eight items served as the critical items. The critical item was either studied as common or as distinct, depending on the structure of its study list. Participants studied 12 lists. Across participants, the order in which the 12 study lists were presented was semi-randomized, with the exception that no more than two common or distinct lists appeared in a row. For each list, the critical item appeared in either the 3rd, 4th, 5th, study position to avoid primacy and recency list positions. Six of the 12 study lists were structured such that the critical item was distinct because it was semantically isolated. In these lists, all the words in the list were from the same category except the one critical word that was from a different category. For example, the word, "table" might be presented in a list of types of fish (e.g., trout, herring, shark, catfish, perch, salmon, and tuna). The other 6 study lists were structured such that the critical item "table" was not distinct. In these lists, all the words, including the critical item "table", were from different semantic categories (e.g., opium, cookies, sponge, salmon, Iroquois, spokes, and four). Across participants, the same or a common item by virtue of the study list construction or as a non-studied item studied as common or distinct and whether it was studied or non-studied produced 4 study lists sets. In the test list, there were 24 critical test items, half of which were studied and half of which were non-studied, and 24 filler items. The purpose of the filler items was to disguise the nature of this task as memory test, and to include "Yes" responses as described later. Of the 12 studied critical items, half had been studied as distinct and other half had been studied as common. At test, the distinct items were paired with category labels that corresponded
to the category membership of the background items in the study list. If the word, “table” had been studied as distinct in a list of fish, then at test participants would see table paired with the category label, “Type of fish”. The common items were paired with other unique category labels that corresponded to one of the background items in the study list. For example, if “table” were studied as a common item in a mixed list of unrelated items including only one type of fish, then the question would again be whether “table” is a type of fish. However, in this case, the word would have been encoded as common instead of distinct. Thus, the category names and examplars were held constant and only the encoding of the examplars as a distinct or common item was varied. For all studied items, common and distinct, the correct answer to the category verification question was “no”. This was done for two reasons. First, they wanted to mimic the evaluation process that they assumed to have occurred when people saw the different item in either a list of other similar items (the distinctive processing condition) or in a list of intermixed items (the common processing condition). Second, they wanted to be able to directly compare category verification times for common and distinct items. Previous research using the category verification test has consistently shown that people are faster to confirm it (Vaidya and Gabrieli, 2000; Vaidya et al., 1997). Therefore, we kept the “no” responses constant across distinct and common items. The 12 category verification questions for non-studied items also required a “no” response for purpose of comparison across studied and non-studied conditions. To reduce the bias to respond “no” to the verification questions, they included 24 filler questions to which the correct answer was “yes”. The order of category verification questions in the list was randomized with respect to common/distinct encoding, studied/non-studied status, and yes/no response.
Data from the explicit and the implicit version of category verification test were analyzed separately because of procedural and baseline differences. The level was set at .05 for all analyses. For both tests, they analyzed participants mean accuracy and median reaction time. For the explicit version of the test, they compared recognition accuracy. As predicted, result from the explicit version of the test showed that participants had better recognition for isolated than non-isolated category exemplars, consistent with previous literature using the explicit memory test of category verification (Srivinas et al. 2000), there was no differences in reaction time to correctly recognize isolated versus non-isolated exemplars. For the implicit version of the test, they again analyzed accuracy and reaction time. They obtained overall priming in the accuracy data and found that participants were better at correctly disconfirming category membership for studied versus non-studied category exemplars. However, they were no more accurate to disconfirm category membership for isolated exemplars than to the non-isolated exemplars. The authors compared reaction times to disconfirm category membership for studied versus non-studied exemplars and found that the pattern of reaction times was in the expected direction, although the differences was not statistically significant. Importantly, this lack of overall priming appeared to be due to the comparable reaction times between the non-studied and common or non-isolated items. However, reaction times for the isolated items were significantly faster than reaction times for the non-studied items. Importantly, priming was significantly greater for the isolated items than the non-isolated items. Thus, they found greater priming, as measured by speeded reaction times to disconfirm category membership, for category members that had been isolated at study than those that had not been isolated. Results suggested that the distinctiveness effect can be obtained using an implicit test if the
test engages the evaluative processing that took place at study (i.e., evaluating the way in which an item is different from surrounding items). This result is generally consistent with the transfer appropriates processing framework, which states that performance on explicit and implicit tests is based on the match of processing. However, their results demonstrated that it is important to cue the specific conceptual processing that occurred at study. As such, the results supported the more recent elaborations of multiple conceptual processes that mediate memory (Cabeza, 1994; Hamilton and Rajaram, 2001; Vaidya et al. 1997; Weldon and Coyote, 1996). Ultimately, these findings demonstrated that although awareness effect is sufficient to mediate the distinctiveness effect and often plays an important role, it is not always necessary.

The results were interpreted to suggest that the distinctiveness effect can be obtained on an implicit memory test if it recapitulates the exact evaluative processing engaged at the time of study. However, it is not clear whether the distinctiveness effect was obtained on this test because the test accessed category information alone, or because it required the same evaluative processing across study and test. In other words, it could be that simply cueing the general category information was sufficient to produce an advantage of the isolated items (and not necessarily the accompanying evaluative process per se). To distinguish between these two possibilities, experiment 2a and 2b used a different conceptual implicit test, this test again accessed general category information, but it did not require an active evaluation of category membership that would recapitulate the evaluative processing that occurred at encoding. They used explicit and implicit category production tests. In the explicit version, participants saw a category label and were told to write down an example of that category that they had seen earlier.
Priming would be observed if people were more likely to write down a category exemplar when it had been studied than when it was not studied. 106 students from SUNY Stony Brook and 88 students from Washington University in St. Louis (for a total 194 students) participated for research credit. 72 participants (40 from Stony Brook and 32 from Washington University in St. Louis) received the explicit category cued recall test. The remaining 122 participants (66 from Stony Brook and 56 from Washington University in St. Louis) received the implicit category exemplar production test. 122 participants were tested in the implicit version to obtain a completely counterbalanced group of 72 participants (43 from Stony Brook and 29 from Washington University in St. Louis) that were categorized as “unaware” of the connection between study and test based on the post-test questionnaire data. Their data were included in the initial analyses. The experiment was 2X2X2 mixed factorial design. Test type (explicit vs implicit) was the between subjects variable and study status (studied and non-studied) and item type (common and distinct) were manipulated as within subject variables. The study materials were identical to those used in experiment 1. The test materials differed in that participants were now presented with only the category label in experiment 2a (e.g., A type of fish), whereas in experiment 1 they were presented with category labels and targets (e.g., A type of fish: table). The labels in experiment 2a corresponded to the critical items—thus, unlike experiment 1, all of the to-be-recalled targets were consistent with the labels presented. For example, people would be given the label, “piece of furniture”? and write down examplars of this category, regardless of whether they had studied “table” in a list of all type of fish, or in a non-isolated, random list. Once again data from the explicit category cued recall test and the implicit category exemplar production test were analyzed separately because of procedural and baseline
differences. For both tests, they analyzed participant’s accuracy. Result from the explicit category cued recall test replicated earlier work (Smith and Hunt, 2000) and showed that participants had better category cued recall for isolated than non-isolated category examplars. In the implicit category examplar production test, participants were first sorted into test “aware” and “unaware” groups based on the post-test questionnaire responses. Overall priming for participants classified as unaware showed that these participants produced more studied exemplars than non-studied examplars, but this effect did not reach significance using a two-tailed test. This result was attributable to a lack of significant priming for non-isolated items. Priming for distinct items was significant. This pattern was similar to the pattern of results from the category verification test in experiment 1. Priming was observed only for isolated but not for non-isolated items. Results from the test unaware participants supported that people were no more likely to produce examplars that had been studied as distinct than those that were studied as common. Given the theoretical significance of this result, they conducted additional analyses to confirm the lack of a distinctiveness effect in the implicit category production test. First, they conducted a power analysis; the results showed that the effect size for the implicit category production test in experiment 2a was small, such that they would need over 1000 participants to have an 80% chance of detecting this effect. Second, they also examined the data from the test aware participants; interestingly, the data from the test aware participants followed the same pattern and showed no difference in category production for isolated and non-isolated items. This finding suggested that the extent of awareness was not strong enough to motivate participants to engage in intentional strategies. Two possible reasons were pointed out to support this interpretation. One was that the procedure for this test required speeded responses in that it required rapid production of
responses (within 30s), thereby reducing the opportunity for deliberate retrieval. Two, the responses on the post-test questionnaire indicated that, although many people guessed the purpose of the experiment, they did not report deliberately recalling items to write down in the short time they were given (Richardson-Klavehn and Gardiner, 1995 and Richardson-Klavehn and Gardiner, 1996; Richardson-Klavehn, Gardiner, and Java, 1996). It is also likely that participants became aware of the purpose of the test only upon answering the questionnaire. Because these two groups did not appear to perform differently, they examined all the implicit participants (both aware and unaware) together to increase power. Even when they increased power by combining these two groups for a total of 122 participants, they found no evidence for a distinctiveness effect: there was no difference in priming for isolated and non-isolated items. Result from experiment 2a suggested that the distinctiveness effect cannot be obtained on the implicit test of category production, and that this outcome is not attributable to a lack of reasonable power. Although they failed to find a distinctiveness effect on this test of category production, this result is weakened by the fact that they did not obtain significant priming for the non-isolated items. Experiment 2b was conducted to determine whether the isolation effect would be obtained on a test of category production when significant priming was obtained for both isolated and non-isolated items. They attempted to increase priming in this experiment by using critical items of a lower frequency. In experiment 2b, they used less frequency category exemplars for the critical items (e.g., bookcase for the critical furniture item). They also used a more elaborate cover story to attempt to further reduce the possibility of explicit contamination. With these changes, they examined whether the distinctiveness effect would be obtained on an implicit test of category production when priming was significant for both the isolated and non-isolated items.
students from Washington University in St. Louis were tested in the implicit version to obtain a completely counterbalanced group of 72 participants that were categorized as “unaware” of the connection between study and test based on the post-test questionnaire data. The experiment was 2X2 mixed factorial design: study status (studied and non-studied) and item types (common and distinct) were both manipulated within subjects. The material were identical to those used in experiment 1 and 2a, with the exception that the critical items were now of lower frequency (average POSITION = 13.58, as compared to 3.96 in experiment 2a).

Again, participants were first sorted into test “aware” and “unaware” groups based on the post-test questionnaire responses. Examining overall priming for only the participants classified as unaware, a t-test showed that these participants produced significantly more studied examplars than non-studied examplars. The overall level of performance was lower in this experiment; this reduction was to be expected given that lower frequency category examplars were used in this experiment. Priming was also significant for both the isolated and non-isolated items. They found no difference in priming for words that were isolated or non-isolated at study. Thus, again, they found no evidence for an advantage for isolated words on the implicit test of category production. Results from experiments 2a and 2b suggested that the distinctiveness effect cannot be obtained in conceptual priming when the task does not recapitulate the distinctive processing from study.

Brandt, Gardiner, and Macrae (2006) conducted a study to investigate the distinctiveness effect in forenames: the role of subjective experiences and recognition memory. The main aim of the present research was to further investigate the qualitative nature of the
distinctiveness effect using socially relevant stimuli that, unlike the faces in the Brandt et al. (2003) study, are familiar to participants.

Participant in the experiment 1 were 34 undergraduate from the University of Bristol (20 women and 14 men). The experimen. had a single-factor (forename: typical verses distinctive) repeated measures design. The forenames used in the present experiment were taken from a study by Macrae, Mitchell and Pendry (2002) in which participants rated the familiarity of 160 male and female forename. Two sets of forenames (familiar and unfamiliar) were constructed then sequentially presented to 20 participants on an Apple Macintosh microcomputer (Power Mac). The presentation of forenames was randomized and for each forename, participants were required to make a typicality rating by means of a key press. A 9-point scale was used whereby 1 represented a very distinctive forename and 9 represented a very typical forename.

Participants’ overall corrected recognition scores (i.e. hits-false alarms for distinctive verses typical forenames) were submitted to a paired t-test. This revealed that distinctive forenames were recognized significantly more memorable than typical forenames has replicated and extended the distinctiveness effect to a class of socially relevant stimuli other than faces (Brandt et al., 2003). In addition, the results demonstrated that distinctive forenames were more likely to elicit ‘remember’ responses in comparison to typical forenames thereby supporting predictions made by the distinctiveness fluency model (Rajaram, 1996). The present results additionally demonstrate that the distinctiveness effect in forenames is present using a direct recognition test, and hence it is not necessary to manipulate both the immediate context as well as participants previous experience with the stimulus.
items in order to obtain a distinctiveness effect (Stanhope and Cohen, 1993).

The aim of experiment 2 was to investigate the effect of retention interval on the distinctiveness effects found in experiment 1. Participants in this experiment were 34 undergraduates from the University of Sussex (21 women and 13 men). The experiment had a 2 (forenames: typical vs. distinctive) X 2 (retention interval: day vs. week) mixed factorial design. Participants arrived at the laboratory individually, and a female experimenter explained that the study was an investigation of people’s memory for forenames. Participants were told that they would be presented with a series of forenames on an Apple Macintosh computer and that they were to try to remember these items as their memory for them would be tested at a later stage. Participants were randomly divided into either the day or the week conditions. Each forename was presented for 2,000 ms with a 2,000 ms inter-stimulus interval. In total, 80 forename were presented (40 distinctive and 40 typical). In addition, half of the presented forenames were female and the other half were male. The presentation of the forenames was counterbalanced such that forenames presented as targets for half of the participants were used as foils for the other participants. After the study phase participants were reminded of their test date (either 1 day or 1 week later). During the test phase, participants were presented with the 80 targets and additional 80 foils, and were required to make old/new decisions for each item and, following an ‘old’ response, were further required to make a remember/know/guess decision.

Participants’ overall corrected recognition scores (i.e. hits-false alarms for distinctive verses typical forenames) were submitted to a 2 (forename: distinctive vs typical) X 2 (retention interval: day vs week)
mixed model ANOVA. This revealed a main effect of forename, whereby
distinctive forenames were recognized more accurately than typical
forenames. The analysis also revealed a main effect of retention interval,
such that recognition accuracy was greater following the 1-day compared
with the 1-week retention interval. The forename X retention interval
interaction was not significant.

Park, Arndt, and Reder (2006) investigated a contextual
interference account of distinctiveness effect in recognition. In this
article, they seek to add to the extant literature in two ways, first, they
seek to further test the extent to which font fan effects are driven by
encoding-or retrieval-based processes with two experiments that hold the
perceptual distinctiveness of study items constants, while varying the
distinctiveness of perceptual features reinstated at retrieval in a
recognition memory task. Second, they seek to examine the extent to
which the memorial advantage of presenting test words in distinctive
fonts is associated with recollection-or familiarity-based recognition.

To examine whether the number of contextual associations to a
font has a greater effect on retrieval competition at test than a differential
influence at encoding. They attempted to equate item processing and
overall distinctiveness of study items. This was accomplished by
presenting each list item with two features, one distinctive (low fan) and
one less distinctive (high fan). At test only one of the two features was
reinstated, either the low-fan or high-fan feature. By equating the
distinctiveness of the stimulus during encoding, it was possible to
measure the effectiveness of two types of cues at retrieval and hereby
examined the SAC retrieval-based account.
In experiment 1, two perceptual features (color and font) were associated with each word. The goal was to keep the overall distinctiveness of each word constant during encoding by presenting the studied word in one high-fan feature (i.e. used to present many other words) and one low-fan feature (i.e. used to present few other words). At test, only one of the two features was reinstated with the word, while the other feature was kept neutral at test, the word would be presented in a common font such as Time Roman; if color was to be neutral at test, the color of the font would be black instead of the colors utilized to present items at study. 26 Carnegie Mellon University students participated in experiment 1a and 24 participated in experiment 1b for research experiment credit. They conducted two versions of the experiment, in which they varied the nature of the perceptual features associated with the study words. During study, each word was presented in an unusual font as one contextual feature, experiment 1a varied the color of the background on the computer screen, and experiment 1b varied the color of the font. Given that colors have for greater pre-experimental exposure than do these a typical fonts, the degree to which varying experimental fan of the color cue would affect performance was unclear. Because of this uncertainly, they instantiated color as a contextual feature in two different ways. For expositional purposes, they treated these two different implementations of the color feature as the same manipulation, because the two variants of this manipulation had the same effects on the results. Each color and font was randomly assigned to be either a high fan or low-fan feature for a given participant. High-fan features (colors or fonts) were shown with 24 different words, whereas low-fan features were shown with 3 different words. Which feature (color or font) was high fan or low fan for any particular word was randomly determined for each participant, with the constraint that each word had one high and one low-
fan feature and that each feature type (color and font) was used equally often as a high-and as a low-fan feature. Words were also randomly assigned to study and test conditions for each participant. Given these constraints, nine different colors and nine different fonts were required to present 48 words at study.

During the recognition memory test, 48 studied words-48 new words were presented. For studied items, only one feature from the study phase was reinstated and the other feature was neutral. For example, if a color feature was reinstated with an item, the item was presented in a neutral font (Time Roman); if the studied font was reinstated as the test cue, then the words was presented on a white screen using a black font so the color feature at test was neutral. An equal number of words were tested with their low-fan and high-fan features. Additionally, an equal number of words were tested with the font of study presentation reinstated as with the color of study presentation reinstated. 48 new words were also tested with one of the two contextual features, using the same distribution of the features tested for targets. These were two study test cycles, that there were 96 old words ands and 96 new words used in the entire experiment.

Experiment 2 employed 9 different voices and 9 different fonts, just as experiment 1a used 9 different colors and fonts. In experiment 1, color or font was made neutral at test, whereas in this experiment, only one of two encoding features was reinstated at test. 21 Carnegie Mellon University Students participated in this experiment, either to satisfy a research experience requirement or in exchange for $7 payment. 9 different voices (5 male and 4 female) and 9 unusual fonts were used to present the words aurally and visually during study. Assignment of voices and fonts to fan condition was randomly determined for each participant.
as was assignment of fonts and voices to words. In all, 48 words were presented at study, each word presented with two features, one low fan (shared with 3 words) and one high fan (shared with 24 words). Only one of the two encoding feature was presented at test, either the voice or font. Likewise, the 48 new words used as foils were either only heard or only read at test, always with one of the voices or fonts used with the studied items. Half of these foils were presented aurally and half visually, just like the target probes, using the same distribution of high-and low-fan features that was used for studied items on the recognition test.

They assessed whether the recognition advantage for items tested in a low-fan feature was associated with recollection-based processes or familiarity-based processes and found evidence that the advantage was reflected in changes in recollection-based processes. The present result may not provide unique support for SAC, in the sense that alternate explanations such as the recollection-fluency perspective of Rajaram (1996, 1998) can also account for these results; there were two primary reasons why they found the SAC account of these data to be informative. First, the SAC model provides a formal specification of the mechanisms producing the perceptual distinctive effects in these experiments: contextual competition. It is important to note that this mechanism is a fundamental property of the theory, and has been applied not only to explain fan effects, but also to other recognition memory phenomena, such as the word frequency effect (Reder et al. 2000), the list-length effect (Cary and Reder, 2003), and list-strength effect (Diana and Reder, 2005). Second, the SAC model advances the strong prediction that perceptual fan effects should be evident in recollection-based processing, because of the view that details of study presentation are associated with episodic representations. Thus, the SAC model can explain not only why
perceptual fan effects occur, but also why these effects are reflected in recollection-based processing, and when the perceptual feature present at study was reinstated at test (Reder et al. 2002).

SECTION III

As stated earlier, this section exclusively reviews those studies which are directly or indirectly relevant to the problem under investigation. More specifically under this section of the chapter an attempt is made to highlight the relationship between emotion and memory in general and between emotion and conversation memory in particular. Research on the relation between emotional states and cognitive processes has burgeoned in the past ten year. Although this area of research has a much earlier history of activity. It lay dormant for many years. Beginning in the mid.....1970’s, this research area began to accelerate to the point where it has now become a major area of activity. A number of article, books and edited volumes now atleast to the vitality of this research area (e.g. Clark and Fisco 1982, Clark and Isen 1982; Fielder and Forges, 1988; Isen 1984) and new journal “cognition and emotion” appeared in 1987 which is devoted entirely to relations among emotional states and full range of cognitive processes typically studied by psychologist. This section describes the important aspects of research on emotional states and impact of different shades of listener’s mood on conversation memory.

Kleinsmith and Kaplan (1964) investigated interaction of arousal and recall interval in nonsense syllable paired-associate learning. The present study was designed to test the independence of the phenomenon from the association values or other qualities unique to the individual
stimuli. Nonsense syllables of zero association value (Hilgard, 1951) were selected as stimuli in order to obtain random arousal effects. In this way both high and low arousal curves would contain essentially the same items and thus any differential behavior in learning could not be explained in terms of qualities unique to the stimuli. The subjects were 36 University of Michigan undergraduates obtained from introductory courses in psychology. They were run in three subgroups of 12 subjects each. The subjects were given a single learning trial with a list of six nonsense syllable-number pairs. The following six 0% association value nonsense syllables were used: CEF, QAP, TOV, JEX, LAJ, DAX. The response items were single digits from 2 to 7. A slide projector with an exposure time of 4 sec. was used to present the stimuli. During the training trial, subject first saw the nonsense syllable repeated with a single digit response. To separate the arousal effects from one set of pairing to the next, two slide containing four colors each were inserted between the paired associates and subject was instructed to name the colors. (Red, green, orange, blue, black, and yellow were used randomly on these slides). The subject was instructed to “concentrate care ully on both colors and nonsense syllable-number pairs” as he called them out loud, to avoid rehearsal subject was not specifically told that he would be tested for recall. In order to determine the specific arousal effects of each nonsense syllable, skin resistance was recorded during learning. The electrodes were of the zinc variety described by Lykken (1959). To ensure constancy of conditions, electrodes and recording apparatus were also used during the recall sessions. During the recall session subject was instructed to indicate the correct number for each nonsense syllable as it appeared, and to guess if he was uncertain. The correct numbers were not repeated. Colors were used as an interpolated task as before. The three groups were tested on at each of three recall intervals: immediate (about 2
min.), 20 min., and 1 wk. To correct for serial order effects six different training lists were used, each list given to two subjects in each group. The lists were designed so that each of the six nonsense syllables appeared once in each position in the list (Fisher and Yates, 1938). The order of the recall lists was varied in the same manner. Any drop in subject's skin resistance which occurred within 4 sec. of presentation of a nonsense syllable was considered an arousal deflection; these were converted to percent deflections. Each subject's six GSR deflections were then ranked. (In case of ties, a deflection occurring at a low level of absolute skin resistance was considered higher arousal than a similar deflection occurring at a higher absolute level). The three highest deflections for each subject were designated “higher arousal learning”, and the three lowest were designated “low arousal learning”.

The findings revealed that at immediate recall, numbers associated with low arousal nonsense syllables were recalled four times as often as numbers associated with high arousal nonsense syllables. The capacity to recall numbers associated with low arousal nonsense syllables decreased as a function of time in a characteristic forgetting curve pattern. On the other hand, the capacity to recall numbers associated with high arousal nonsense syllables showed a considerable reminiscence effect. After 20 min., the increase was found 100%, by a week it has increased 200%. Significant interaction was primarily attributable to the effect of the immediate condition and 1-wk condition. Thus at immediate recall, low arousal learning was significantly greater than higher arousal learning; after 1 wk. the situation reversed and high arousal recall was found significantly better than low arousal recall. Since there was some tendency for adaptation of GSRs during presentation of a list, the possibility aroused that serial order interacts with the arousal
classification. Thus it would be expected that the first two items would tend to fall in the high arousal category, while the last two items would be in the low arousal category. This was not borne out by the data. Analysis of recall as a function of interpolated interval and arousal classification, plotted separately for pairs of consecutive items, was essentially identical to the results.

In brief, high arousal learning showed poor immediate recall due to the relative unavailability of the rapidly reverberating neutral trace, and strong permanent memory due to the great amount of perseverative consolidation. Low arousal learning, on the other hand, showed good immediate recall due to lack of rapid reverberation, and yet poor long-term memory for the same reason.

Natale (1977) investigated effects of induced elation-depression on speech in the initial interview. The effects of depression and elation on speech were sought in an experimental analogue (mood induction) that is both generalizable and valid in term of the authenticity of the affective states. Subjects were 45 female college students, ages 17 to 23, enrolled in an introductory psychology course. Subjects dialogue was obtained in an audiotape automated standardized interview developed by Webb. All recording equipment was located in a chamber adjacent to the interview room. Three separate tape decks (sony-104) were used—one for the delivery of the interviewer’s emergency statements, and the third for the continuous playing of appropriate background noise. The later deck was used to audibly mask the stopping and starting of the other two tape decks. Velten’s technique was used to induce either elated or depressed mood. The subject read preparatory statements silently and then aloud for the mood induction task. The actual mood induction consisted of the reading of 60 self-referent statement; each statement was typed on a
separate card and presented to the subject for 20 sec. A sample elation statement is “If your attitude is good, then things are good, and my attitude is good”. A sample depressive statement is “I have many bad things in my life”. There is also a set of 60 neutral cards used by a control group. The difference scores of the pre-induction and post-induction writing speed score and the mean semantic differential ratings were used as independent criteria to verify the experimental MIP. Post-interview questioning ascertained that none of the subjects discerned the automated nature of the standardized interview and none of the subjects were cognizant of the true nature of the experiment. Newman-Keuls analysis demonstrated that the “personal feelings” associated with the three levels of the state variable were all significantly different from one another. The subjects’ performance on the semantic differential scale indicated that Velten’s MIP did evince feelings congruent with the moods of depression or elation. It was stated that depressed subjects manifested motoric retardation, with elated subjects exhibiting motoric acceleration as compared to a neutral group. The verbal behavior, as measured by all five variables, of subjects in the depressed and elated treatment groups was found not to differ from individual in the neutral mood condition. The post-induction results revealed that subjects made to feel elated, had significantly shorter response latencies, faster speech and articulation rates, and less total speech time occupied by pauses longer than 2 sec (silence quotient), as compared to individual in a neutral mood state. The induction of depression in subjects was shown to produce a larger interviewee silent quotient as contrasted with the neutral subjects. The results also revealed that the second half of the interview (post-induction), regardless of the experimentally induced mood state, elicited fewer words per response, longer reaction time, slower rates, and an
increase of lengthy hesitation (silence quotient) in the speech of the interviewees.

Bower, Monteiro, and Gilligan (1978) conducted a study on emotional mood as a context for learning and recall. In this experiment, emotional moods were produced by hypnotic suggestions. They chose happiness (euphoria) and sadness (grief) as the moods because they are apparently incompatible opposites. Highly suggestible subjects were induced to feel very happy or very sad during learning of a word list, and later induced again to feel happy or sad during free recall of that list. The issue was whether performance is superior if the same mood prevails during learning and recall. Experiment 1 used a single learning list and a short retention interval. One group learned a word-list under hypnosis while in a happy mood; a second group learned while in a sad mood. After a brief interpolated task, they were placed in the opposite mood and asked to recall the word list. Then their original mood was reinstated and they were asked to recall again. The contrast of interest was the first recall in the opposite mood vs. the second recall in the same mood as in original learning. Further, half the words in the learning list were pleasant or happy words, and half were unpleasant or sad words. The question was whether happy subjects recall more pleasant words, and sad subjects recall more sad words. The word list consisted of 16 abstract because Weingartner et al. (1976) found greater SDR (state-dependent retention) effects with abstract rather than concrete nouns. Eight of the words connoted happy things (e.g., humor, pleasure), and eight connoted sad things (e.g., atrocity, misery). The subjects were 10 Stanford undergraduates selected for high susceptibility to hypnosis. They had scored from 10 to 12 (maximum) on an earlier test using the Sanford Hypnotic Susceptibility Scale, Form C (Weitzenhoffer and Hilgard,
They were experienced hypnotic subjects, able to enter trance easily, and to verbally assess the "depth of trance" they reported vivid emotional experiences when these were suggested. The subjects were run individually in a 1-hour session.

Happy subjects frequently laughed as they found humor in everything; sad subjects were long-faced, morose, slow to respond, and often on the verge of tears. Initial learning was equivalent for the happy and sad subjects. The groups showed no forgetting nor did they differ reliably from one another. Thus, a switch in mood caused no loss in memory whatsoever. Mean for second recall did not differ reliably from the first-recall means nor from the end-of-learning means. The results thus showed no mood-dependent retention effects at all. This was not due to a ceiling on recall, since recall averaged around 65% but the fact that items available in immediate recall were still available on the later retention test regardless of the mood state. Happy subjects were not more likely to recall happy words, and sad subjects were not more likely to recall sad words. The sole difference that occurred in this respect concerned the order of recall; the first word recalled by subjects matched the mood during recall. Of 40 recall protocols (10 subjects times four recall each), in 37 the first word recalled matched the mood of the recaller (a highly significant bias).

There is always the possibility that more sensitive experimental conditions could be arranged to show the SDR effect. This was attempted in experiment 2. This experiment had a 2X2 factorial design with four groups of four subjects: happy or sad mood during learning and happy or sad mood during retention testing the next day. The 20-word learning list consisted of eight happy words, eight sad words, and four neutral words. Thus, word-type was a within-subject variable. The 16 subjects were
Stanford undergraduates who had scored very high (10 to 12) on a screening with the Stanford Hypnotic Susceptibility Scale. They were experienced in hypnosis, quickly entered trance, could estimate “depth of trance” and could vividly experience the emotions suggested by the experimenter. The subjects were assigned in random rotation to the four groups defined by the “learning-mood by testing-mood” factors. They found that the original learning was about the same for the Happy and Sad learners and learning mood had a slight effect on later recall. However, the difference between happy and sad learners was not significant in the factorial analysis of variance. Subjects in the same mood as learning recalled an average of 6.8 words, and those in different mood recalled an average of 7.0 words.

Experiment 3 was designed to investigate whether hypnotic mood could be used to enhance or diminish the retroactive interference that normally arises when a person learns two lists of words. The two 16-word lists were abstract words (rated between 2.0 and 3.0 on concreteness in Paivio et al. (1968). In each list there were eight happy words (e.g., affection, victory, gaiety) and eight sad words (e.g., atrocity, tragedy, misery). The subjects were 14 female and 10 males assigned as evenly as possible to the three main experimental conditions. Ten subjects were Stanford undergraduates selected for having scored very high (10 to 12 points) on the Stanford Hypnotic Susceptibility Scale, Form C. The other 14 subjects were non-students selected from a pool of former students in hypnosis workshops conducted by the third author. They were excellent hypnotic subjects; nine were tested on the Stanford Scale, Form C and all scored in the 10 to 12 range. There were six experimental conditions. Using first and second letters to denote mood states during learning of List 1 and List 2, and the third letter to denote mood during recall of both
lists, the six conditions were: SHH and HSS (maximum interference for List 1); HHH and SSS (controls); and SHS and SHS and HSH (maximum interference for list 1). These conditions are referred as the max, min, and control groups. Four subjects were assigned in random alternation to the six groups with the restriction that the male-female composition should be nearly the same in each group. The two sub-groups with a given condition hardly differ in original learning or recall. Thus, for example, subgroups SHS and HSH in the max conditions are quite comparable at all points. This means that mood state per se was not an influential factor during either learning or recall; only the relation of learning o recall moods was influential. The overall analysis of variance for list-1 retention percentages showed a significant effect due to groups. With he max interference group showing poorest list-1 retention (46%) the control group a moderate amount (54%), and minimum interference group showing the best list-1 retention (66%). A similar overall analysis of variance on percentage recall scores (arc sine transforms) for list-2 declared highly significant group differences (with percentages at 90%, 57% and 48% for groups max, control, and min, respectively). Recall was best when the mood-state during recall was the same as the mood-state during learning of the list being recalled and when the learning mood-state of the interfering list was different (e.g., SHS or HSH).

Lewinsohn and Amenson (1978) investigated some relation between pleasant and unpleasant mood-related events and depression. The present research was conducted within a framework that postulated that a low rate of response-contingent positive reinforcement is a critical antecendent of clinical depression (Lewinsohn, Weinstein, & Shaw 1969). That is, it was, assumed that the behavior of depressed individuals does not elicit reinforcement to a sufficient degree to maintain behavior and
that a low rate of response-contingent reinforcement leads to a feeling of
dysphoria, which is central to the phenomenology of depression. Thus,
the concept of the reinforcement is crucial to this approach to depression.
Data from three separate samples were obtained. Sample 1 (Lewinsohn &
MacPhillamy, 1974) consisted of 251 participants, about equally divided
between depressed (D), non-depressed psychiatric (PC) and normal
control (N). Sample 2 (Lewinsohn & MacPhillamy, 1974) consisted of
459 normal participants. Sample 3 consisted of 199 participants about
equally divided between D, PC, and N participants. The participants in
sample 1 and 2, and the non-depressed controls in sample 3, were
recruited through a variety of public announcements to townspeople of
the Eugene-Springfield area of Oregon inviting paid participation in
psychological research. The participants were thus very heterogeneous in
regard to age and educational and occupational levels. The diagnostic
groups did not differ from one another in regard to socioeconom ic status
(Hollinghead, 1957). The depressed subjects for sample 3 were recruited
by announcing, through both public media and community agencies, the
availability of a depression treatment program. For samples 1 and 3 a
two-stage screening process was used to assign participants to the
diagnostic groups. This classification procedure employed the M innesota
Multiphase Personality Inventory (MMPI) in complete (sample 3 or
abbreviated (sample 1) form (Byrne, 1964) and the structured clinical
interview rating form constructed and validated by Grinker, Miller,
Sabshin, Nunn, and Nunnally (1961). The normal controls in sample 2
were selected using only the MMPI with the Byrne (1964) scale criteria.
Depressed participants were mildly to severely depressed with depression
as the major problem. Psychiatric controls had mild to severe
maladjustment other than depression. All participants signed informed
consent forms and except for the depressed subject in sample 3. The
relation between the mean aversiveness rating of each UES item (based on sample 3, N = 199) and the percent of the population for whom the item significantly correlated with mood was less strong. Examination of the scattergram suggested that the relation was attenuated by the fact that some of the most aversive items occurred with low frequency and, therefore, were significantly correlated with mood for only a few participants. There was a highly significant main effect due to diagnostic class for all measures. Depressed participants have lower PES (Pleasant Events Schedule) frequency, enjoyability, and cross-product scores and higher UES (Unpleasant Events Schedule) frequency, aversiveness, and cross-product scores than the non-depressed participants. There was also a highly significant main effect due to item type. Mood-related items were rated higher than non-mood-related items on all measures. It was demonstrated that mood-related items discriminate more powerfully between depressed and non-depressed participants than do non-mood-related items. Strong positive correlations were found among pleasant events and among unpleasant events. It was concluded, therefore, that no significant correlations were observed between the participants' impact rating of pleasant and unpleasant events. Results also showed that low but statistically significant positive correlations were observed between PES MR (mood related) frequency and UES MR frequency and between PES MR enjoyability and UES MR aversiveness. All other correlation did not differ significantly from zero.

Teasdale and Fogarty (1979) investigated differential effects of induced mood on retrieval of pleasant and unpleasant events from episodic memory. The aim of the present study was to obtain unambiguous evidence of the relationship between depressed mood and speed of retrieval of pleasant and unpleasant past experiences. Each
subject was seen on two occasions, receiving a depressed mood induction on one occasion and a happy mood induction on the other. On each occasion subjects were presented with a series of words. For each word they were asked to recall a real-life experience, specified to either pleasant or unpleasant, that they associated with the word latency of retrieval was measured for each experience. 30 student subjects were recruited from Oxford Polytechnic and Dorset House School of Occupational Therapy, Oxford. Of the 16 subjects who met the acceptance criteria for the study, 12 were women and 4 were men, with a mean age of 19.8 yr. A latin square with repeated measures design was used. Each subject received both happy and depressed mood inductions, counterbalanced for order across subjects. Subjects rated their mood using cards bearing 0-100 line scales, 10 cm long, on which 0 was labeled “I do not feel at all X” and 100 was labeled “I feel extremely X”. For different cards, X was anxious, despondent, and happy. “At this Moment” was typed at the top of each card to indicate that it was instantaneous mood that was to be rated. The ratio of latency for retrieval of unpleasant experiences to latency for retrieval of pleasant experiences was significantly less following depressed mood induction than following happy mood induction. This result was largely due to an effect of mood on retrieval of pleasant experiences. The judges’ ratings of the pleasantness of the experiences recalled were similar for the corresponding experiences in the two conditions. Following depressed mood induction, the subjects were also significantly more anxious, less happy, showed longer speech pause-time in counting from 1 to 10, and took longer, overall, to retrieve memories associated with the stimulus words. Following the depressed mood induction, latency of retrieval of pleasant experiences was significantly longer and the U/P ratio significantly smaller than after the happy mood induction. Latency of
retrieval of unpleasant experiences was very similar following two mood inductions. A significant effect of word list was obtained for the U/P ratio, the mean for list A being significantly smaller than that for list B. The researcher found that the subjects were reasonably balanced for the order in which they received mood inductions (14 had the order happy-depressed, 16 the order depressed-happy), but were not balanced for order of mood induction and word list. For the total sample of subjects, significant correlations were obtained between the difference in mean latency of retrieval of pleasant memories in two conditions and differences in mean happiness, as well as the differences in an overall measure of “bad mood” calculated by adding the mean scores for despondency and anxiety, and subtracting the mean score for happiness. Non-significant trends in the same direction were obtained for differences in mean anxiety and despondency. The “worse” the mood was in the depressed condition relative to the happy conditions, the greater was the slowing of recall of pleasant experiences relative to the happy condition.

Maslach (1979) investigated “Negative Emotional Biasing of Unexplained Arousal”. The current research was designed to further investigate the process of a cognitive search for the causal antecedents of unexplained arousal. It involved a modified replication in which procedural elements were altered to correct some of the original methodological problems. These improvements included the use of hypnosis to produce an arousal state, rather than a drug injection. Second, the behaviors of the two emotion conditions were made directly comparable the nature, level, and timing of the activity were equivalent in both conditions, and only the confederate’s obvious emotional state (happy or angry) was varied. The reasons given for the confederate’s emotion were ones that the subject could readily share, and the
confederate’s behavior appeared believable and appropriate in the laboratory setting. Finally, several measures were added to the study, both to assess the validity of the experimental manipulations and to reflect more precisely any differences in emotional behavior. 48 undergraduate subjects (25 males and 23 females) were drawn from the introductory psychology course of Stanford University. All of the subjects had received scores of high hypnotic susceptibility on the Harvard Group Scale of Hypnotic Susceptibility (Shor and Orne, 1962). 36 of these subjects were randomly assigned to training in hypnosis prior to the experiment, while 12 subjects were randomly assigned to control groups that did not receive such training. The hypnotic training, which averaged about 10 hours per subject, was conducted in small groups and utilized a relaxed and permissive approach (Zimbardo, Maslach and Marshall, 1972). The training program sampled a wide variety of hypnotic experiences and included as criterion tests (a) the successful completion of posthypnotic suggestions with accompanying amnesia (b) increased tolerance of ischemic pain (a measure developed by Lenox, 1970). Throughout the experimental session, trained and untrained subjects were treated identically except for a brief hypnotic induction for the trained subjects. There were six such intervals following each stimulus, and their individual scores were combined into a total mean score for each stimulus. A repeated measures analysis of variance for the highest heart rate scores showed that there was a significant main effect of visual stimuli, such that the highest heart rate following the arousal cue was significantly higher than the highest heart rate following the neutral stimuli (which did not differ from each other). However, there was a significant interaction between stimuli and hypnosis condition. The F tests for the two hypnosis conditions revealed that this main effect was due entirely to the hypnotized subjects. The unhypnotized subjects did
not show a significant increase in highest heart rate in response to the arousal cue. An analysis of mean heart rate revealed exactly the same pattern of results. A similar pattern of arousal was reflected in the total number of GSRs (Galvanic skin response) following each stimulus. There was a significantly greater number of GSRs following the arousal cue than following the neutral stimuli. However, there was also a significant interaction between stimuli and hypnosis condition and the F tests for the two hypnosis conditions showed that the main effect was due entirely to the hypnotized subjects and not to the unhypnotized group. This difference in arousal response is even more striking given the fact that both hypnotized and unhypnotized subjects showed a more rapid recognition of the cue word (as evidenced by a shorter latency and higher amplitude of the first GSR following the cue). The changes score for the hypnotized subjects indicate the marked shift in emotional state that resulted from their arousal. Their mood became significantly more negative, less positive, and less passive. In contrast, the unhypnotized subjects showed no changes on the negative, active, or passive mood indexes, although there was a tendency for them to report less positive feelings. Overall, the hypnotized subjects showed a more extreme shift on a composite of these four indexes mood than did the unhypnotized subjects. There was no between-group difference in overall recall accuracy on either list, subjects in the arousal groups made significantly fewer errors of commission than did the unaroused and unhypnotized groups on the second learning list. From 92% to 100% of the subjects in the aroused conditions reported experiencing each of the four suggested symptoms, whereas only 26% to 52% of the unaroused and unhypnotized subjects reported having even a single one of the symptoms. Subjects in the happy condition evaluated the confederate’s mood as significantly more positive than the neutral midpoint of zero. In contrast, the treatment
mean for subjects in the angry condition was significantly more negative than zero. The difference between the two confederate conditions is highly significant. The overt behavior of the subjects varied as a function of the confederate’s mood, but seemed unaffected by experience arousal. Subjects in the presence of a happy confederate exhibited a significantly higher amount of sociable behaviors than did subjects who were with the angry confederate. This was true for both verbal behaviors and nonverbal behaviors. Although the confederates’ mood had a demonstrable effect on the subjects’ overt behavior, it did not exert the expected influence on their reported emotional state. A comparison of means revealed that this was due to the more negative scores of the aroused condition. Analysis of each of the happy and the angry ratings separately revealed the same significant difference in emotional reactions between the aroused subjects and the other subjects without arousal. There was a significant main effect of arousal condition and comparison of means revealed that the aroused groups reported a significantly more negative emotional state than both the unaroused and unhypnotized groups (which did not differ from each other). This same pattern of results was found for each individual scale. On the overall index, the aroused conditions reported negative emotions that were significantly different from the neutral midpoint of zero, while the unhypnotized groups were neutral in their emotions (their mean score did not differ significantly from zero), the unaroused subjects rated themselves as feeling a positive emotional state. Apparently, the unaroused subjects were responding appropriately to the posthypnotic suggestion that they would feel relaxed and good when they came out of hypnosis. The positive feelings of the unaroused group offer a dramatic comparison level for the negative reaction reported by the arousal subjects (despite the same posthypnotic suggestion to feel good and identical prior to hypnotic training). The number of aroused subjects
(67%) who thought they knew the cause of their current feelings was as high as that of the unaroused and unhypnotized groups. They showed the same increase in heart rate, similar changes on the Mood-Adjective Checklist, the same pattern of recall errors, but not a similar change in GSR. The simplified arousal subjects reported increased breathing and heart rate. Surprisingly, however, all of them also reported sweaty palms, and two thirds reported a sinking feeling in the stomach. The perception of the confederate by the two control groups was the same as that of the happy aroused group, and they showed the same pattern of sociable behaviors (both verbal and non-verbal) in their interaction with the confederate. Overall, the two arousal control groups were similar to the happy aroused group in terms of physiological arousal, veridical perception of the confederate, and overt behavior. However, they differed in experienced emotional state. The informed arousal subjects reported a neutral set of feelings (which did not differ from the zero midpoint), as opposed to the negative feelings expressed by the happy aroused group. They all stated that the reason for their experienced arousal was the hypnotic suggestion. In contrast, the simplified arousal subjects reported the same negative emotion as the happy aroused group (on each of the self-report measures) and gave the same type of causal explanations. Their mean rating on the emotion index was significantly more negative than the neutral midpoint of zero.

Riskind, Rholes, and Eggers (1982) conducted a study entitled “The Velten Mood Induction Procedure Effect on Mood and Memory”. This study was conducted to determine whether the memory bias observed by Teasdale and Fogarty (1979) is produced by the somatic statements in the VMIP, as the study of Frost et al. (1979) suggested. It is possible that the self-devaluative aspects of the VPIM have an effect on
memory, even if they do not affect mood. 52 undergraduates (13 per experimental condition) took part in the study. Slightly less than half of the subjects were male and the rest were female. (There were no sex effects for any analysis). Prior to the VMIP mood manipulation, subjects filled out a 7-item mood scale and took part in a task in which they recalled four pleasant and four unpleasant life experiences (in a random order). The time between memory trial onset and beginning of a subject’s recitation of experience (response latency) was recorded. After the premeasures, the VMIP was administered to manipulate mood. Following the procedures of Frost et al. (1979), this study included three depressed mood conditions. In the combined-depression condition, subjects read VMIP statements that concerned both self-devaluation (e.g., I am worthless) and “depressed” somatic states (e.g., I feel listless). In the self-devaluative and somatic conditions, subjects read only self-devaluative or somatic item, respectively. Subjects in the fourth condition read Velten’s set of positive mood items. This “elation” condition served as reference point for assessing the effects of the three depressed conditions. Following the mood-induction procedure, the original recall and mood measures were readministered.

An analysis of covariance was conducted on the posttest mood scores, with the pretest scores as the covariate. The elation group (contrast weight: 3) was found to be significantly more positive in mood than the other three groups. Further comparison showed that the elation group was significantly higher in mood than the self-devaluation group, the somatic-suggestion group and combined-depression group. Posttest recall latencies for positive experiences were significantly faster than for negative experiences among subjects in the elation condition. In contrast, among the self-devaluative, somatic-suggestion, and combined-
depression conditions, positive experiences were not recalled significantly faster than negative ones. Further comparisons contrasted the elation condition with each depression condition. The difference between the speeds of positive versus negative recall in the elation condition was significantly different only from the speeds in the self-devaluation condition. The correlation for positive recall was not significant. The correlation for negative recall was not only not significant but opposite to the expected direction. The results do not support the hypothesis, suggested by Frost et al. (1979), that somatic VMIP statements have a greater effect on memory and mood than self-devaluative ones. All three depression conditions produced lower mood than the elation condition. Only the self-devaluative statements produced a significant difference with the elation condition in positive versus negative recall. The self-devaluative group, however, was not significantly different from the other two depression conditions. Regarding memory, the present results extended the findings of the memory-distortion literature by showing that self-devaluative statements alone can produce a difference in positive versus negative recall that is significantly different than the one for elation statements. The other two depression conditions were not significantly different from either the elation or self-devaluative condition. Results suggested that changing self-devaluative cognitions may be beneficial therapeutically by (a) improving mood and (b) ameliorating a negative memory bias that may otherwise maintain a depressed person's mood overtime.

Whinney (1982) investigated the role of arousal in memory for conversation. This study examined the extent to which high interactional content can lead to superior memory for the surface form of whole sentences, even when the interaction is viewed on videotape. If the effect
of interactional content holds up, even for the nonparticipants, it may be possible to begin to study memory for natural conversation by using videotape as a stimulus. In order to test for the effect of participation on memory for natural conversation, three groups of subjects were studied: (1) participants, subjects who knew each other and who directly engaged in natural conversation; (2) familiar nonparticipants, subjects who knew the participants but who did not themselves participate in the conversation and who listened to and watched a videotape of the conversation generated by the participants; and (3) nonfamiliar nonparticipants, subjects who neither knew the participants nor engaged in the conversation and who listened to and watched the discussion on videotape. These three groups were viewed as varying along a general dimension of involvement: the participants were the most involved and the nonfamiliar nonparticipants were the least involved. Unfortunately, it was not possible to deconfound participation and familiarity in this design, because a conversation between nonfamiliar participants could hardly qualify as an instance of natural conversation. Thus the main comparison of interest was between the participants and nonfamiliar nonparticipants. Although familiarity is partially confounded with participation, by comparing the performance of the familiar nonparticipants with the nonfamiliar nonparticipants, it was possible to get some idea of the relative importance of familiarity as opposed to participation in memory for natural conversation. The participant and familiar nonparticipants groups were composed of graduate students from the clinical area of the department of psychology at University of Denver. Four conversations, each involving three students plus the experimenter (P.R), who was also a psychology graduate student, were recorded on videotape. Simultaneously, electrodermal responses were recorded on polygraph tracings. Each conversational group was made up of students
from the entering year. Thus, there was a 4th-year conversational group, 3rd-year group, a 2nd-year group, and a 1st-year group. Because there were 3 subjects in each group, there were a total of 12 subjects in the familiar participants group. The familiar nonparticipants group also contained 12 subjects. Each familiar nonparticipant was shown a videotape of the conversation in which the participants were from his own class. In other words, 4th-year familiar nonparticipants watched 4th-year participants, 3rd-year familiar participant's watched 3rd-year participants, and so forth. The familiar nonparticipants had extensive social contact with the participants within the context of the numerous close interactions in the clinical program. This was true because all the students within a given class spent a great deal of time with each other in supervision meetings, area meetings, and courses. The nonfamiliar nonparticipants were 12 graduate students in the professional psychology program at the University of Denver. All were complete strangers to those they watched on videotape. However, in terms of clinical interests, academic ability, age, social class and so forth, the nonfamiliar nonparticipants were quite comparable to the other two groups of clinical graduate students. Results suggested that none of the correlations between EDR (Electrodermal response) and memory was significant, whereas all the correlation between interactional content and memory for surface form were significant. Furthermore, a scatter plot of memory performance as a function of EDR showed that these low correlations were not the result of masking of an underlying quadratic function. Thus, these results provided little support for the arousal hypothesis. However, these results are consistent with a model in which there are two paths from interactional content, one to memory and one to EDR, but no path from EDR to memory. Such a model clearly fits these data better than the alternative, in which EDR and interactional content make separate contributions to memory. There was a highly
significant effect of involvement. Participants showed high ERDs on an average of 39% of the sentences, whereas familiar nonparticipants showed high EDRs on only 20% of the sentences and nonfamiliar nonparticipants showed high EDRs on only 9% of the sentences. It is interesting to note that, while both familiarity and participation affected arousal, the increase due to familiarity (19%) was almost twice that due to participation (11%). There was also a significant effect of interactional content on EDR. An average of 30% of the high interactional content statements produced high EDRs, compared with only 17% for the low interactional content statements. Finally, there was a significant Involvement by Interactional Content interaction. The large difference between high and low interactional content statements observed in the familiar participants (27%) and decreased for the familiar nonparticipants (9%) and was virtually eliminated in the nonfamiliar nonparticipants group (3%).

So for as overall memory performance is concerned, it was found that memory for the meanings of high interactional content statements exceeded that of low interactional content statements. The surface forms of high-interaction items were recognized significantly better than the surface forms of low-interaction items. The main effect of involvement also just reached significance. Planned comparisons showed that nonfamiliar nonparticipants remembered less about surface form than did either participants or familiar nonparticipants. The main effect of class was significant, the 4th-year class remembered more than the 5th-year class, the 2nd-year class, and the 1st-year class. An important and potentially useful finding of the study was the absence of an interaction between interactional content and involvement. Whereas in the EDR data the difference between high and low interactional content statements was
quite large for the participants and increasingly smaller for the familiar nonparticipants and nonfamiliar nonparticipants, in the memory data the difference between high and low interactional content statements was virtually the same for all three experimental groups. The main effect of interactional content was again highly significant. The main effect of involvement was also significant. Planned comparisons showed that the nonfamiliar nonparticipants remembered less than both the familiar participants and the familiar nonparticipants. Finally, the effect of class was significant, indicating once again that the 4th-year class remembered more than any of the other class.

In order to test the possibility that factors other than interactional content may have been responsible for the difference in memory between high-and low interaction statements, a control study was performed to determine whether the high-interaction statements used in this experiment were simply more inherently memorable than the low-interactional statements. 36 students from the University of Denver were recruited by an advertisement in the school newspaper. Subjects listened to a tape consisting of 150 conversational-tape sentences. The list of sentences included the 56 high and 56 low interactional content sentences from the four conversations, along with 38 filler sentences. Filler were included to prevent ceiling effects on the memory test. All of the sentences were recorded by a single speaker who spoke in a conversational tone. The order of the sentence presentation was randomized, with the constraint that the first 10 and last 10 sentences of the list be filler items to control for primacy and recency effects. Subjects were instructed to listen carefully to the sentences in order to do well on a subsequent memory test; the nature of the test was not specified. As part of the informed consent procedure, it was necessary to warn the subjects that a few of the
sentences contained profanity, highly suggestive sexual references, or both and that, if they did want to listen such language, they did not have to participate in the experiment. All subjects, however, agreed to participate. Immediately after listening to the 150 sentences, the subjects were given a composite of the four memory tests used in the conversation experiment. The items were the same four-alternative forced-choice questions; however, in the test given to control group, the speaker was not identified by name. The test consisted only of the 112 high- and low-interaction statements from the original conversations.

An analysis of variance was performed in which the 36 experimental subjects were compared with the 36 control subjects. Overall, the control group had better memory for surface form than did the experimental group. However, this superiority held only for low interactional content statements. In terms of memory for the surface form of high interactional statements, the two groups did not differ. It was found that the size of this difference in memorability is much smaller (9%) than that observed for any of the experimental group (21%). This is reflected in the highly significant interaction between group and interactional content. The control group remembered more of the content of the high-interaction sentences (82%) than of low-interaction sentences (76%). Again this difference (6%) is much smaller than the average difference observed in the experimental group (31%), as reflected in the significant interaction between group and interactional content.

Clark, Milberg, and Ross (1983) conducted a study on arousal cues arousal-related material in memory: implications for understanding effects of mood on memory. Their first two studies examined the idea that information about arousal can be stored in memory and that subsequent arousal can serve as a retrieval cue for material presumably
linked to that information. Under the guise of a study on learning under distracting conditions, subjects learned a list of phrases while experiencing normal arousal. Later, they were given a recall test for phrases on both lists when they were experiencing either enhanced arousal or normal arousal. Subjects were 37 college students enrolled in their first psychology class. Two lists of phrases consisting of an adjective and a noun were generated (e.g., failed exam, wide road). Then 8 judges rated how abstract and how emotionally toned each phrase was. Finally, 20 phrases were selected for each of 2 lists, hereafter referred to as List A and List B, such that the average abstractness and emotionality ratings of the phrases on each list did not differ. Results supported the effectiveness of the arousal manipulation. Exercise produced significantly greater increases in pulse from pre-test-to post-task than did relaxation. For subjects who were highly arousal at recall this proportion was greater for phrases originally heard when experiencing high arousal (85%) than for phrases originally heard when experiencing normal arousal (70%). For subjects who were experiencing normal arousal at recall this proportion was greater for phrases originally heard when experiencing normal arousal (91%) than for phrases originally heard when they were experiencing enhanced arousal (69%). Results revealed no effect of arousal at learning, no effect of arousal at recall and a significant interaction between arousal at learning and arousal at recall. A planned comparison indicated that when subjects were highly aroused at time of recall, recall of material heard when highly aroused was significantly greater than recall of material heard when normally aroused. A second planned comparison indicated that when subjects were experiencing normal arousal at recall, recall of material heard when normally aroused was significantly greater than recall of material heard when highly aroused. A 2X2 mixed ANOVA was conducted with emotionality of
phrases as one variable, arousal at recall as the second independent variable, and percentage of phrases recalled as the dependent variable. There were no significant effects.

In second study sixteen college students participated. The word lists and procedure were the same as in the first study with the exception of the manipulation of arousal at recall. For subjects who were highly aroused at recall this proportion was greater for phrases originally heard when experiencing high arousal (80%) than phrases originally heard when experiencing normal arousal (60%). In addition, for subjects experiencing normal arousal at recall this proportion was greater for phrases originally heard when experiencing normal arousal (77%) than phrases originally heard when experiencing high arousal (65%). Analysis revealed a marginally significant effect of arousal at learning, such that subjects tended to recall more phrases learned when aroused than phrases learned when not aroused, no effect of arousal at recall and a significant interaction between arousal at learning and arousal at recall. When subjects were highly aroused at recall, recall of material heard when aroused was significantly greater than recall of material heard when not aroused. A second planned comparison indicated that when subjects were not highly aroused at recall, recall of material previously heard when not aroused was significantly greater than recall of material heard when aroused. As in the first study, a second 2X2 mixed analysis of variance was conducted with emotionally of phrases and arousal at recall as the independent variables, and number of phrases of each type recalled relative to the number learned as the dependent variable. This analysis revealed no significant effects.

Third study was designed to demonstrate that additional arousal at the time a person is experiencing a positive mood will enhance priming of
positive material from memory. Under the guise of a study on distraction and memory, subjects listened to a story while either stepping up and down on a block (High Arousal) or stringing cardboard disks together (Normal Arousal). Afterwards, all subjects took a memory test. Half the subjects from each arousal condition were then told they had done well (Positive Feedback Conditions); half were told that test had not yet been scored (No Feedback Conditions). Finally, all subjects filled out a survey about their university. 44 college students participated. Each was randomly assigned to one of the following conditions: (1) High Arousal-Positive Feedback, (2) High Arousal-No Feedback, (3) Normal Arousal-Positive Feedback, or (4) Normal Arousal-No Feedback. The change in subjects' blood pressure and pulse from just before the distraction task to 2½ minutes after the task was the manipulation check of arousal. In the High Arousal condition subjects' systolic and diastolic blood pressure increased +3.5 and +1.8, respectively. In contrast, in the Normal Arousal conditions, subjects' systolic blood pressure dropped as did their pulses. The difference between changes in systolic blood pressure between the two groups was significant as was the difference between changes in pulse. The difference between changes in diastolic pressure was not significant. Scores in the Positive Feedback condition were higher than scores in the No Feedback condition. In the Positive Feedback Condition, arousal seems to have heightened the effects of the positive feedback. Subjects who received positive feedback gave their University higher ratings than did others. The interaction between Arousal and Feedback approached significance and there was no main effect for Arousal. The High Arousal Positive Feedback subjects rated the university as being significantly more positive than the Normal Arousal Positive Feedback subjects. A second planned comparison between the Normal Arousal-No
Feedback and High Arousal-No Feedback Condition revealed no significant difference.

Finally, the blood pressure and pulse increase for each subject were correlated with the positivity of their judgments regarding their university separately for each of the four conditions. None of these correlations was significant. The result of third study supported both predictions. At least in the presence of arousal, a positive feeling state, induced by success, enhanced the favorability of subjects’ rating of their university. The results also supported new prediction that if something positive happens to a person, arousal, in addition to that which was assumed that the person feels as a result of the mood state induction, will enhance the impact of that mood on judgments.

Warr, Barter & Brownbridge (1983) conducted a study on the independence of positive and negative affect. The objective of this article was to explore three different possible explanations of the paradox relating to association between positive and negative effect.

In summary, the research described in this article explored three possible explanations of the statistical independence of Bradburn’s measures of positive and negative affect. The explanations are of different kinds and are not mutually incompatible, so that each may be viewed as an alternative account of one part of reality.

520 volunteer undergraduate students from Sheffield University served as subjects in four separate samples. Sample 1 comprised 24 men and 26 women, sample 2 comprised 100 men and 90 women, and samples 3 and 4 each comprised 70 men and 70 women. Subjects’ median age was 20 years. Respondents were contacted in university halls of residence or
in teaching departments (other than the psychology department) and asked to complete anonymously the questionnaire provided. Respondents returned this questionnaire through boxes placed at several locations in the university. Items were drawn from the literature and from discussion with undergraduate students. Following a pilot study with 89 items, a 65-item inventory was administered to sample 1 subjects in order to identify for subsequent inquiry events in university students' lives that were rated consistently as either desirable or undesirable. The events covered those broad areas examined by previous investigators: health, work, financial, home and family, and personal and social. Respondents from sample 1 were asked to assess the desirability or undesirability of the occurrence of each event in their lives on a 7-point scale from _3 to +3 (extremely, moderately, and slightly undesirable; neither desirable nor undesirable; slightly, moderately, and extremely desirable). Some examples of events were personal illness or injury that required a period in the hospital, an important piece of work favorably assessed, major financial difficulties, planned pregnancy (if female) etc. On the basis of the sample 1 results, 50 events were chosen for the principal investigation. These had been consistently assessed as desirable or undesirable (25 items each), and mean desirability rating were available for use in weighting reports of the occurrence of each item in the main questionnaire. This was administered to sample 2(n = 190), and respondents were asked to place a tick beside “any event which has happened to you in the last six months”. The mean numbers of desirable and undesirable reported were 9.15 and 3.24 (SD = 3.41 and 2.3, respectively). In addition to these unweighted totals, each person's responses were weighted by the mean item desirability rating obtained from sample 1. Two separate desirability rating were examined: those from sample 1 respondents of the same sex as a sample 2 respondent, and those from all members of sample 1 irrespective of sex.
For reported life events from sample 2, three scores were thus computed for both desirable and undesirable events: unweighted totals, total weighted by same-sex desirability ratings, and totals weighted by combined-sex desirability ratings. Lei and Skinner (1980), Ross and Mirowsky (1979), and others reported that several forms of differential weighting of life events yield scores whose pattern is almost identical to that of unweighted totals and that associations with other variables are virtually unchanged by weighting. This was also the case in the present study, with correlations between weighted and unweighted scores close to unity and external correlations almost identical for each type of weighting.

Scale from the Eysenck Personality Questionnaire (EPQ; Eysenck & Eysenck, 1975) was used to measure extraversion and neuroticism in sample 3 and 4 (n = 140 in each case). For extraversion, mean for sample 3 and 4 were 13.16 and 14.47, respectively (SDs = 4.29 & 4.12); for neuroticism, mean were 9.96 and 8.94, respectively (SDs = 4.65 & 5.18). These values are slightly above the age norms for extraversion and slightly below that for neuroticism. The two scales from EPQ are typically significantly negatively intercorrelated (e.g., Eysenck & Eysenck, 1975; Rocklin & Revelle, 1981), and in this case the value were -.33 and -.28 for the two samples, respectively. Bradburn’s (1969) 10 items served as the basis for construction of the new measures of positive and negative affect. However, two of the original items were omitted because responses to them are a function of specific interpersonal events (cf. Cherlin and Reeder, 1975). These items are “proud because someone complimented you on something you had done” and “upset because something criticized you”. Five additional items were devised for both positive and negative affect, which were drawn from the literature on
minor psychiatric symptomatology and from studies of mood. Data from the two samples were combined to yield an overall sample size of 330 in order to carry out several forms of item analysis. Despite an initial expectation that it would be appropriate to omit a small number of items, the two 9-item scales proved to be most acceptable. The correlation between them was .01, and alpha coefficients were .66 and .64 for positive and negative affect, respectively. Successive omission of the four items in each scale that had lowest item-whole correlations hardly affected the interscale correlation (values all less than .06), and decided to retain the 9-item scales because these scale covered the widest range of experiences. Mean values were 7.42 (SD = 2.00) and 5.98 (SD = 2.13) for positive and negative affect, replicating a typical finding with the Bradburn items that positive affect scores tend to be higher than negative affect scores.

The results showed that number of desirable and undesirable episodes were found in general uncorrelated and that both affect and life-event measures were members of the general category of episodes. The significant correlation between evaluatively matched affect and life-event scores provided direct support for this possibility. An independence between counts of desirable and undesirable life events seems reasonable on logical grounds in the case of psychologically healthy samples, but a negative association might be expected for sample of, for example, depressed patients; depth of depression is likely to covary both with reports of many undesirable events and few desirable ones. Positive associations between numbers of the two kinds of events would be less easily explained, although a correlation of .19 (N = 337 mothers) had been reported by Gersten, Langner, Eisenberg and Orzech (1974) and value of .16 and .47 (N = 720 adults in a community survey) have been
reported by Ross and Mirowsky (1979). In both of these studies, however, the number of undesirable items in the inventory considerably outweighed the number of desirable one, and judgments about desirability were made by the investigators rather than by a sample of respondents. One or both of these features may have been responsible for the positive correlations. When responses were given in terms of proportion of time, a strong negative relationship was observed between positive and negative experiences.

Another issue that was examined was the relationship between positive and negative affect and personality measures of extraversion and neuroticism. The findings demonstrated that positive affect is significantly associated with extraversion whereas negative affect was found to be associated with neuroticism.

Rholes, Riskind and Lane (1987) investigated emotional state and memory biases: effects of cognitive priming and mood. They formulated three hypotheses as given under

1. In the self-evaluative conditions, positive experiences would be recalled faster than negative ones when subjects were exposed to positive statements, and negative memories would be recalled faster than positive ones when subjects were exposed to negative statements.

2. The impact of the self-evaluative cognitions would be largely unaffected by mood; thus, they expected to find that self-evaluative cognitions would influence memory retrieval even among those subjects whose mood states were little affected by the self-evaluative statements.

121
3. Finally, the somatic conditions would have little impact on memory retrieval relative to the self-evaluative conditions.

96 students from the introductory psychology course at Texas A & M University took part in the study. 46% of the subjects were men. The percentage of men in each condition was 50, 39, 44, 44, and 50 for the positive somatic, positive self-evaluative, negative somatic, negative self-evaluative, and neutral conditions respectively. The questionnaire consisted of four positive mood adjectives were: elated, cheerful, happy, and good (mood), and four negative mood adjectives were: sad, discouraged, depressed, and bad (mood). Subjects indicated how elated, cheerful, sad and so forth they were on 11-point scales. An overall mood score was obtained by subtracting the sum of a subject’s responses to the positive adjectives. After the first eight recall trials, each subject read one of the five types of VMIP statements. They read either positive somatic (PSS) statements (e.g., I feel energetic), positive self-evaluative (PSE) statements (I am a very capable person), negative somatic (NSS) statements (I feel tired and listless), negative self-evaluative (NSE) statements (I am worthless); or neutral, nonaffective (NU) statements (Oklahoma City is the capital of Oklahoma). There were 43 statements in each list. Subjects were asked to read each list twice. As many as possible of the statements from Velten’s (1968) original lists of statements were used. The data were analyzed with t-test that revealed that the mood scores of subjects in both of the negative statement conditions (NSE and NSS) were significantly more negative than the scores of subjects in the neutral group and that the mood scores of subjects in both of the positive statement conditions (PSE and PSS) were significantly more positive than those of subjects in the neutral group. Other comparison showed no significant differences between either the positive self-evaluative and
negative somatic conditions or the negative self-evaluative and negative somatic conditions. These results indicated that overall the self-evaluative and somatic conditions had equally strong effects in mood. However, it is important to note that the analysis of variance (ANOVA) also revealed a significant Sex X Treatment Condition interaction. T-test showed that the mood scores of women in the negative self-evaluative condition were significantly more negative than those of men in this condition. Thus, women were more responsive to the negative self-evaluative statements than were men. There were no significant differences between the scores of men and women in any of the other conditions.

A preliminary analysis showed that the sex-of-subjects variable did not influence the recall latencies either as a main effect or in interaction with other independent variables. Thus, the interaction of conditions and sex that appeared in the mood data had no parallel in the recall data. Furthermore, they tested for male versus female recall differences in the one condition, NSE, that showed a sex effect on mood and found that recall latencies were not different in this condition. The results further showed a main effect for type of memory, with positive memories recalled faster than negative ones, in the two somatic conditions (PSS and NSS), and an interaction, with positive memories recalled faster than negative ones among subjects who read positive statement and negative memories recalled faster than positive one among subjects who read negative statements, in the self-evaluative conditions (PSE and NSE). A contrast that tested the significance of this predicted pattern of means proved to be highly significant. Further analyses showed that positive memories were recalled significantly faster than negative ones in the PSE condition and negative memories were recalled significantly faster than positive ones. The overall ANOVA revealed a main effect for
positive versus negative memories which indicates that, on the whole, positive memories were recalled faster than negative memories (cf. Natale & Hantas, 1982). It also revealed a significant interaction between positive and negative statements and recall of positive versus negative memories which indicates that positive memories recalled faster in the positive statement condition (PSE and PSS), whereas negative statements tended to be recalled faster in the negative statement condition (NSE and NSS). Subjects in the upper half of this distribution showed large changes in mood in the expected direction in response to the Velten procedures, and subjects in the lower half of this distribution showed small changes. This overall analysis revealed the same effects as the overall ANOVA reported previously, a main effect for positive versus negative memories which indicates that positive memories were recalled faster than negative ones, and an interaction between positive versus negative statements and the recall of positive versus negative memories. Compared the negative memory latencies in the control group with those in the four experiment groups revealed that only the NSE group led to significantly faster recall in one-tailed. All of the other groups produced a non-significant comparison with the control. Compared the latencies for positive memories revealed that the PSE group had shorter latencies than the control and that the PSS group also had shorter latencies in one-tailed. The other comparisons of positive memories yielded non-significant effects. The main effect for positive versus negative memory trials which simply indicates that the memories recalled on the positive memory trials were judged by the raters to be more positive than the one recalled on negative trials. This analysis also revealed a Condition X Memory Valence interaction. Post hoc comparisons indicated that the unpleasant memories of the negative self-evaluative groups were more negative in valence than memories of all other VMIP (Velten Mood Induction
Procedure) groups. Surprisingly, comparison also indicated that pleasant memories were less positive in the negative somatic group than in all other groups.

The results, however, do not invalidate the state-dependent model. There are, in fact, results that are in keeping with it, namely the covariation between mood change and memory effects that occurred in the somatic conditions. Thus, the results of this study as a whole suggest that (a) self-evaluative and somatic mood induction do not seem to influence memory retrieval in the same way, and (b) some of these results are consistent with the state-dependent model but others quite clearly are not consistent with it and seem best explained by the priming model. The results, in other words, seem to suggest that, for the present, it may be best to regard the effects of mood on memory as determined by multiple processes.

A number of recent studies have argued that negative versus positive affect states have effects on a wide range of behaviors in addition to memory retrieval. Positive mood states, relative to negative ones, have been found to alleviate learned helplessness (Kilpatrick-Taback & Roth, 1978), increase prosocial behavior (Rosenhan, Karylowski, Salovey, & Harris, 1982), and increase perceptions of control over events (Alloy, Abramson, & Viscusi, 1981). Some of these studies of affect have used cognitive or autosuggestive techniques to induce mood (e.g., the VMIP), whereas others have manipulated situational factors (e.g., success or failure on a task, or adventitious positive or negative events) to induce mood states. Regardless of their procedures, the results have been interpreted as demonstrating the influence of affect on behavior. The present study under review suggests, however, that it may be the cognitive (self-evaluative) as well as the affective experiences that
accompany mood-induction procedures that influence behavior. For example, Alloy et al. (1981) found that the illusion of control was increased by having subjects read positive mood statements from the VMIP. Was it the affective component of mood that influenced perceptions, or was it the positive self-evaluative cognitions (e.g., I am very able) that may have been produced by the Velten procedure? The results of this study suggested that it may partly have been the cognitive set induced by the positive evaluative statements in the Velten procedure.

To summarize, this study seems to provide support for both the priming and the state-dependent models as they relate to memory accessibility.

Forgas, Burnham, and Trimboli (1988) investigated mood, memory, and social judgments in children. This study sought to extend this work by looking at the role temporary moods play in children's encoding and retrieval of social information, as well as in their impression-formation judgments. Do affective states influence children's memory performance in school? Does affect facilitate the learning of affect-consistent information? Are judgments of a friend or a peer dependent on how a child feels at the time and, if so, are these mood influences similar to the pattern found with adults (Forgas and Bower, 1987a)? These were some of the questions the investigators looked at in a large and heterogeneous child population. Groups of 8 to 10 year-old children were presented with counterbalance pictures and descriptions of two stimulus characters, identified as Lisa or Jenny. Each presentation was preceded by a short video tape segment designed to elicit a manipulation happy or sad mood state. One day later, children free recall and recognition memory for detailed of the two stimulus characters were assessed while they again experienced either a happy or a sad state, and
their impression-formation judgments of the target characters were also obtained. Subjects were 161 schoolchildren (85 boys and 76 girls), between 8 and 10 years of age. The experiment was based on 2X2X2 design: children's mood state while learning about the first stimulus character (happy/sad), their mood state while learning about the second stimulus character (happy/sad), and their mood state one day later while remembering information about either of the two characters (happy/sad). In addition, the effect of the affective loading of the stimulus material (positive vs. negative) was also assessed as a within-subject variable. The dependent variables included two memory measures (cued recall and recognition) and impression-formation judgments. The use of short video sequences with strong and predictable emotional impact appeared to be the most suitable mood-manipulation technique in the present circumstances. They needed six short video-film sequences (three sad and three happy) for the study. A small preliminary investigation was run in order to select and establish the reliability and validity of these films as suitable mood manipulators for children. As a first step, they encoded fourteen 6 to 10 min story sequences (6 happy and 8 sad) from popular children's programs such as “Lassie”, “Little House on the Prairie”, Happy days”, “Swiss Family Robinson”, “Disneyland”, “Different strokes”, and “Gilligan’s Island”. These episodes were selected on the basis of the consistently happy and cheerful, or sad and depressed, emotional tone of events they depicted. Care was taken to exclude episodes with potential direct priming effects on the stimuli. In all, 14 episodes were presented to a sample of twenty-seven 8 to 10 year-old children, who were told “you will be watching some stories. At the end of each story.....write down how happy or sad that story made you feel”. Judgment were made by children by checking one of seven faces with stylized expressions ranging from the very sad to the very happy. Of the
14 episodes, those 6 (3 happy and 3 sad) that produced the most consistent and clear-cut mood effects on the -3 to +3 sad-happy scale were selected. The mood effects of these selected sequences were then further validated by showing each sequence to a second group of 18 children and asking them to indicate how the films made them feel. Average mood ratings were significantly different following sad sequences as against happy sequences, indicating the validity of these films as effective mood manipulators. The effects of the films were also reliable across the first and second group of children who rated them.

Happy learning moods resulted in slightly better recall memory than did sad learning moods for both characters. Similarly, children who felt happy at the time of recall also remembered significantly more details about the target characters than did children who felt sad. They found no significant main effects in recognition memory due either to the learning mood or the recognition mood of the subjects. A significant interaction of learning and retrieval moods showed that total recall for both stimulus characters was better when mood at the time of learning was identical to rather than different from mood at the time of recall. Happy subjects remembered more about a character they encountered in a happy rather than in a sad learning mood and sad subjects remembered more about target seen in sad rather than in happy learning moods. This finding clearly supports the mood-state dependent retrieval hypothesis. Consistent with priming models that assume that the strong retrieval cues provided by recognition tasks will mask the much weaker mood effects (cf. G. Bower, 1981), they found no mood effects on recognition memory. They also found that the interaction of the two learning mood (when learning about target 1 and when learning about target 2) had a powerful effect on overall recall. In other words, details of both stimulus
characters were recalled much better when the targets were encountered in contrasting rather than identical learning moods, irrespective of the quality of the retrieval mood. A happy-sad learning sequence resulted in better overall recall than did a happy-happy sequence and a sad-happy learning sequence led to better recall than did a sad-sad learning sequence. They found that recognition memory was also better when the 2 characters were encountered in contrasting rather than in matching learning mood states. Results showed that, overall, children remembered almost twice as many negative as positive details about both stimulus characters. They also found a significant quality (positive vs. negative details about the target child). Contrary to their expectations as well as most previous findings (Bartlett et al. 1982; Bower et al. 1978; Forgas and Bower, 1987a; Nasby and Yando, 1982), they found that more negative details were remembered by children who felt happy rather than sad at the time of encoding a target character. There were no significant mood biases in the encoding of positive details. They also found a significant mood main effect on judgments. In a happy mood, children were more likely to give extreme endorsement to both positive and negative descriptions of the target characters, whereas in a sad mood, both kinds of features were rated more conservatively, as being less typical of the stimulus person. In short the findings of the study under review are partially consistent with existing mood-priming models (Bower, 1981; Clark and Isen, 1982) and research with adult subjects. Moreover, they also found limited support for the mood-state dependent retrieval hypothesis.

Moreover, Parrott and Sabini (1990) conducted a series of experiments on mood and memory under natural conditions and presented evidence for mood incongruent recall. Experiment 1 and 2 were
field studies in which the effects of everyday moods were observed outside the laboratory on a memory task that was designed to be naturalistic. In both field studies, evidence of mood incongruent recall was obtained. Experiments 3, 4, and 5 were conducted in the laboratory to establish the reliability of the field studies and to evaluate alternative explanations of their results. Experiment 1 took advantage of a reliable elicitor of mood in the ecology of the undergraduate student: the return of the graded midterm exam (Forsyth & McMillan, 1981; Russell & McAuley, 1986). The mood-inducing exams were handed back at the beginning of class, and the lecture that followed included a memory task as part of a class demonstration. This memory task involved recalling several autobiographical memories, a task selected for its high relevance to mood's influence on memory in everyday life (Teasdale, Tylor, & Fogarty, 1980). By comparing the affective content of memories recalled by students who did well with that of memories recalled by students who did not do well, they hoped to study the effect of naturally produced moods on a realistic memory task in a natural context. The inevitable tradeoff, however, was that for obvious ethical reason they were not able to assign their subjects their grades (and thus mood) at random, nor were they able to ensure that all students present would consent to participate. Thus, this study used a quasi-experimental design, and, like all such studies, required greater attention to ruling out threats to internal validity than would a true experiment (Cook & Campbell, 1979). Students in a social psychology class at the University of Pennsylvania participated in this study on a volunteer basis. There were 233 students in attendance on the day of the study. Of these, 124 students (58 women, 34 men, 32 missing sex data) contributed usable data to the study. The study was conducted during the class in which the graded midterm examinations were handed back to the students. At the beginning of class, the students
were introduced to a confederate posing as a visiting researcher conducting a study on students’ expectations and perceptions of exams. The researcher handed out blank index cards and asked the students to write two things: the letter grade that they expected to get on their midterms and a secret identification code of their own invention. The researcher explained that he was going to know what the students thought about the exam after they got it back, and that the purpose of the identification code was to allow him to match those later perceptions with their expected grade while guaranteeing anonymity for each student. The completed cards were then collected. During the next 15 min the course instructor and teaching assistants handed back the graded midterms to the students and explained the graded scale and the grading of the essay questions as is usually done in this class. The visiting researcher then handed out a one-page exam questionnaire on which the students were asked to record their actual grade and identification code and to rate the exam on a number of scales designed to enhance the credibility of the cover story. The second part of the study was designed in such a way as appear unrelated to the first. The course instructor announced that he would lecture during the remainder of class and that he would begin with a class demonstration of an experiment on personality and memory. The experiment, the instructor explained, would consist of two parts. First, the students were be asked to write a brief description of the first three memories they recalled from a period of time that were be specified. The students were assured that their responses would remain completely anonymous and that they therefore need not worry about censoring their recollections. Second, the students were be asked to fill out a brief personality questionnaire. At this time booklets were handed out to the students. A blank cover sheet prevented the students from learning the specific time period of the memory task before being asked to begin.
After reporting the last memory, subjects continued on the final page. Consistent with the cover story, this page was titled PART TWO: Personality Questionnaire, but its true purpose was to provide a manipulation check on the effect of receiving the exam grade on the subjects’ mood. This manipulation check was conducted after the memory task so as not to make subjects aware that their moods were relevant to the experiment (an arrangement that necessarily entailed not being able to check moods prior to recall). Subjects were instructed to indicate how they were feeling at the present time on a series of Likert scales. Students completed 7-point unipolar scales for Happy, Sad, Bad, Emotional, and Confused (anchored by not at all ‘1’, moderately ‘4’, and extremely ‘7’,) and bipolar scale for Ashamed-Proud, Angry-Grateful, Lonely-Belonging and worried-optimistic, anchored by extremely (1 and 7) and neither (4).

Finally, the students were asked to write in their own words how they believed that getting their exam back affected their recall of past events. They were then thoroughly debriefed by the course instructor. The students were told that they could elect not to hand in their data if they did not wish to contribute to the study. As noted above, 109 students elected either to do so or never to participate in the class demonstration in the first place. This number of non-participants is quite high but is difficult to control in a field study. Students agreeing to contribute data wrote their secret identification code on their data booklet. To determine the affective tone of the student’s memories of high school, each memory was rated by two independent judges who were unaware of all information about its author. Each memory was rated for affect at the time the event occurred on two 7-point scales. One scale rated positive affect, the other negative affect, none (1), intense (7). In this and in all
subsequent experiments, judges were trained to use these scales to rate, as best they could, the affects specified by the student without making unsupported inferences. Students who received grades that were higher than average reported being in more of a happy mood, less of a bad or sad mood, more proud (less ashamed), and more grateful (less angry). Multivariate comparison of all nine ratings revealed a significant difference between groups. These differences were also apparent when comparing students who did worse than expected, although the differences are not as large and the multivariate comparison did not reach significant. Thus, the outcome of the exam, both relative to the absolute to be associated with students’ subsequent moods and emotions in the manner predicted. This experiment provided evidence of mood incongruent recall under conditions similar to those that naturally occur in daily life. Students who received good grades, and thus were in good moods, tended to recall events that were significantly less positive and more negative than were those recalled by students who received lower grade and thus were in worse moods. This outcome seems to suggest that under natural conditions the predominant effect of good moods relative to bad moods was not based on automatic associative priming or other processes producing mood congruent recall. Perhaps it results, in part, from students’ attempts at mood regulation (Clark & Isen, 1982). Mood incongruence was found not only when exam performance was measured in terms of absolute grade but also when it was measured in terms of the deviation of the actual grade from the expected grade.

Experiment 2 used clear and cloudy weather to produce happy and sad moods. The effect of sunlight on mood and behavior is well documented. For example, Cunningham (1979) found sunshine to be the most significant of 13 weather variables in predicting self-reported mood
and the likelihood of helping behavior, and Schwarz and Clore (1983) found sunny and rainy days to influence subjects' reports of life satisfaction. The study was conducted inside the entrance of the undergraduate library on the campus of the University of Illinois at Urbana-Champaign on weekdays between 11 am and 4 pm. Every third person entering the building from outdoors was approached and asked to participate, excluding people who appeared to the university employees. Between 11 and 16 people were approached on each of three sunny days and three cloudy days. A total of 8 people (four in each condition) of the 73 approached declined to participate. Thus, data were collected on sunny days from 30 people (13 men, 17 women) and cloudy days from 35 people (14 men, 21 women). On sunny days subjects reported being in better moods than did subjects on cloudy days. There was no difference between sexes, nor was there an interaction, so the separate univariate measures were inspected individually. Negative affect associated with subjects' memories was substantially greater on sunny days than on cloudy days, a difference that is incongruent with mood. They found no effect of sex or of an interaction between sunlight and sex. The positive affect associated with subjects' memories revealed a negligible difference between sunny days and cloudy days. Although the direction of this small difference was in the direction of congruence with mood, this difference did not even approach significance. Again, there was no significant difference between sexes, nor was there an interaction between sunlight and sex both.

The goal of experiment 3 was to induce happy and sad moods in the laboratory using a cover story and musical mood induction that inform subjects of the relevance of mood to the research and that enlist subjects' cooperation in changing and maintaining their moods. The
memory task given to subjects was the same as that used in experiment 1, to recall three events that happened during subjects’ high school years. 34 female undergraduates at the University of Illinois participated in this experiment in partial fulfillment of a course requirement. The unavailability of male subjects for this experiment was not regarded as a serious limitation because there were no sex differences in experiment 1 and 2. One subject was not allowed to complete the experiment because she did not follow instructions. Of the remaining 33 subjects, 22 met the acceptance criteria for the experiment. Subject mood prior to the MMIP was assessed in two ways, as in the procedure of Clark and Teasdale (1985). First, subjects were asked to rate 11 affective and cognitive states on a series of self-report scales. Three scales measured sadness (Sad, Depressed, Hopeless), happiness (Cheerful, Happy, Lighthearted), and anxiety (Anxious, Apprehensive, fearful), and two measured uncertainty (Confused, Uncertain). The mean of the two or three scales measuring each affect was used in data analysis. Each scale asked, “At this moment how do you feel”? Afterward appeared a 9-point Likert scale anchored by not at all (1) and extremely (9). The second mood assessment was a writing speed task used to measure psychomotor retardation (Natale & Hantas, 1982; Velten, 1968). Subjects were given a blank sheet of paper and asked to write out numbers in descending order starting with 100; they were stopped after 60s. The number of digits written served as the measure. Subjects were left alone during each of these tasks. The MMIP was powerfully successful. Inspection of the cell means for self-reported mood and for the writing speed task revealed that after the happy MMIP, subjects’ ratings of happiness increased relative to their ratings at the beginning of the experiment. While their ratings of sadness decreased and their counting speed increased; the sad MMIP had the reverse effects. Rating of anxiety and confusion decreased after both MMIPs. This
pattern of results suggested that the effect of the MMIP was focused primarily on the intended moods. Analysis of simple effects of the time variable in these three interactions revealed that the happy MMIP increased subject’s self-related happiness but did not significantly change ratings of sadness. The sad MMIP, on the other hand, significantly increased ratings of sadness, decreased ratings of happiness and decreased counting speed. The means for the effective content of subjects’ memories of high school, as rated by the judges, suggested that mood had a strong effect on the affective content of memory, and that this effect was congruent with the subjects’ moods. This MANOVA revealed a significant multivariate effect of mood, Univariate ANOVA confirmed that the memories of happy subjects both contained more positive affect and contained less negative affect than did those of sad subjects. Although there were no significant effects of memory or of the interaction, note that the size of the effects for the first memory recalled was about double what it was for subsequent memories. This observation is supported by one-way ANOVAS performed separately for each of the three memories, which were significant for the first memory but not for the second or third memories.

The fourth experiment is different from the previous ones in the sense that in this experiment the mood was induced by musical notes. Thus happy and sad mood were induced by music and incongruent recall was investigated. 49 undergraduates (28 women) at the University of Illinois served as subjects. They received credit toward an introductory psychology course requirement for participating. There were no significant effects of sex or order, or any interactions. Subjects who heard the happy music reported experiencing both more happiness and less sadness than did subjects who heard the sad music. The subtle musical
mood manipulation did not appear to differentially affect anxiety or irritation. Although the comparison between these means and those from experiment 3 is overall remarkably similar, the differences in mood ratings between the happy and sad conditions are not as large as they were in the previous experiment. This difference does not necessarily mean that the subtle musical mood induction produced less intense moods, however, since the subjects in this study rated their moods at the end of the experiment, while those in experiment 3 rated them immediately after the MMIP. The significance of these differences was evaluated by performing a four-way, mixed-model MANOVA having music, task, order, and sex as between-subject variables and memory number (first, second, third) as within-subject variable. The MANOVA revealed a significant interaction between music and memory number. Subsequent univariate ANOVAS showed this interaction to be significant both for the ratings of positive affect and for the ratings of negative affect. There were no other main effects or interactions. Thus, as in the midterm experiment, the effect of mood on memory depended on the serial position of the memory. For the first memory recalled, there was a significant incongruent effect of mood on memory. Thus, mood incongruent recalled occurred on the first memory recalled, and there were no significant effects on the second and third memories. These results are very similar to those obtained in experiment 1, suggesting that the more subtle musical mood induction may be a good laboratory analogue of the natural mood inductions used in experiments 1 and 2. The occurrence of mood incongruent recall in the present experiment refutes the hypothesis that this phenomenon can occur only outside the laboratory. Instead, it supports the hypothesis that mood incongruent recall occurred when subjects are unaware of the relevance of their moods to the experiment. The researchers replicated the fourth
experiment using a different memory task in experiment 5. Hence the fifth experiment was conducted with this objective. In this experiment, therefore, subjects were asked to recall three events that happened to them within the past week. In all other respects the present procedure was identical to that of experiment 4. 31 undergraduates (23 women) at the University of Illinois served as subjects. They received credit toward an introductory psychology course requirement for participating. The apparatus, materials, and procedure were identical to those of experiment 4, with the exception of the memory task. Subjects were asked to recall three events that happened to them within the past week, a task designed to minimize the possibility that subjects could think of this time as contrasting with the present. As in experiment 3 and 4, the number of subjects was insufficient for evaluation of all five self-reported affect ratings with a Mood X Task Order X Sex MANOVA, so only the two variables directly testing the success of the manipulation were entered into the multivariate analysis. A significant main effect of music confirmed the difference between the happy and sad conditions. Subsequent univariate analysis revealed that the ratings of sadness differed significantly, and those of happiness fell just short of significance. There were no significant differences in reported anxiety, irritation, or confusion, again suggesting fairly focal effects of the mood manipulation. They suggested that mood incongruent recall occurred, just as it did with the high school memory task. A MANOVA was performed on the ratings of positive and negative affect in the memories, with music, task order, and sex as between-subject variables and memory number as the within-subject variable. The MANOVA yielded a significant main effect for music, confirming the occurrence of mood incongruent recall. Subsequent univariate ANOVAS revealed significant effects of music for both positive and negative affect. These ANOVAS revealed the same
pattern exhibited in experiment 1 and 4, and the reverse of that in experiment 3. For the first memory recalled, both the ratings of positive and negative affect differed significantly in the direction of mood incongruence. For the second and third memories recalled, the ANOVAS failed to reach significance. Thus, although the present experiment failed to produce a significant interaction between mood and memory number (as occurred in experiment 1 and 4), the same general pattern of results nonetheless occurred.

These results strongly contradict the hypothesis that mood incongruent recall was the result of a contrast between subjects past high school years and their present collegiate life. On the other hand, these results replicate the phenomenon of mood incongruent recall occurring when subjects are unaware of the relevance of mood to the experiment, and generalize it to a recall task different from that used previously.

Eich, Macaulay and Ryan (1994) conducted a study on mood dependent memory for events of the personal past. There were two reasons for conducting this research. One was to replicate and extend prior research relating to mood congruent memory: the “enhanced encoding and/or retrieval of material the affective valence of which is congruent with ongoing mood” (Blaney, 1986). To clarify, and to consider the results of a study by Teasdale and Fogarty (1986), in which elated and depressed moods were induced by instructing subjects to internalize the emotion implied by a series of self-referential statements, the technique devised by Velten (1968). Using a constrained test of autobiographical memory, Teasdale and Fogarty found that although the mean latency to generate negative events was constant across mood conditions, positive events were retrieved more rapidly on average in the elated than in the depressed state. This research provided an opportunity
both to replicate these results and to pose a number of novel questions concerning mood congruence. For instance, is greater importance ascribed to positive than to negative autobiographical memories generated during a pleasant mood, and do people experiencing an unpleasant affect recollect negative events more vividly than positive ones?

The second, more pressing reason was to promote the occurrence of mood dependent memory. Although the subjects did not know at the time, the autobiographical memories that they generated today (during the encoding session) would serve as the targets of a surprise test of free recall given 2 or more days later (during the retrieval session). Thus, on returning to the laboratory, subjects were reminded that they had previously generated events from their personal past when primed with common-word probes, following the induction of a mood that either did or did not match the one they had experienced earlier, subjects were asked to recall as many of these events as possible, in any preferred order, without benefit of any observable reminders or cues.

The experiment 1 consisted of two session, here called encoding and retrieval, and conformed to a 2X2X2 mixed design. Whereas both mood at encoding (pleasure (P) vs. displeasure (U)) and mood at retrieval (again, P vs. U) were varied between subjects, the valence or type of autobiographical event generated during the encoding session (either positive or negative) was varied within subjects. The crossing of the first two variables defined four encoding/retrieval conditions, each of which was represented by 16 randomly assigned subjects. At the encoding session, subjects were told that the study was part of a research program aimed at understanding how moods influence the performance of various cognitive tasks and how the performance of these tasks in turn influence
mood. It was explained that the experiment would be divided into 60 to 90 min sessions (designated simply as “first” and “second”), spaced 2 or 3 days apart. Subjects were also advised that they would be tested individually throughout the course of the experiment and that each session would entail a different set of tasks.

Next, participants were provided with a page bearing the matrix, an adaptation of the “affect grid” invented by Russell et al. (1989). Subjects were told that the matrix would be used to measure two aspects of their prevailing mood: level of pleasure (indicated along the horizontal axis of the matrix) and level of arousal (identified by the vertical axis). It was further explained that, reading from left to right, the columns connote a mood that is extremely unpleasant, very pleasant, moderately unpleasant, slightly unpleasant, neutral (the shaded center square), slightly pleasant, moderately pleasant, very pleasant, and extremely pleasant. Similarly, the rows signify a state that ranges from extremely high arousal at the top through neutral at the shaded center to extremely low arousal at the bottom (with slightly, moderately, and very high or low levels in between). On receiving these instructions (along with the examples cited by Eich & Metcalfe, 1989), subjects marked the one square that best exemplified the levels of pleasure and arousal that they were experiencing at that moment. Participants were then seated in a lounge chair, bordered by stereo speakers. Through these speakers were played, at a comfortable listening volume, one of four cassette tapes, each carrying 60 min of instrumental music. Two of these cassettes contained different selections of “happy” music (excerpts of Beethoven’s Minuet in G, Boccherini’s Minuet in E, Mozart’s Divertimento No. 136 Pachelbel’s Canon in D, and Vivaldi’s Four Seasons: Spring, Summer, and Fall), and the other two contained different selections of “sad” pieces (segment of Alfinoni’s
Adagio in G Minor, Barber’s Adagio pour Cordes, Grieg’s Peer Gynt: The Death of Ase, and Sibelius’ violin Concerto: Second Movement). The two tapes representing each type of music were assigned at random; with the proviso that subjects tested under matched mood conditions (P/P or U/U) hear one selection of the appropriate music during the encoding session and the alternate selection during the retrieval session. 5 minutes after music onset, and every 5 min thereafter, subjects marked their current levels of pleasure and arousal on a clean copy of the mood matrix (which, once marked, was promptly collected by the experimenter and never used again). The music continued to play while subjects made these mood ratings and did not stop until the end of the session. On the rare occasions in which subjects took more than 60 min to complete the session, they rewound their tape and replayed it from the beginning. On reaching the requisite level, subjects read a list of 16 common unrelated words (such as banana and bride) and rated each item in terms of its concreteness, imagery, and goodness of meaning on 7-point scales. Subjects were told that the intent of this word rating task was determined whether people perceive the properties of a word differently depending on their mood. After rating the final item, subjects marked a new mood matrix and then began the personality assessment task. The material for this task consisted of 50 adjectives, displayed one at a time in a random order, representing 25 pairs of trait-descriptive opposites (e.g., warm, cold, sociable, shy). Subjects were told that an adjective would appear on a monitor and that they should indicate, as quickly as possible, whether the term is self-descriptive. The task, which was also administered during the retrieval session in the same manner using the same materials, was included in order to test some ideas stemming from Campbell’s (1990) work on self-concept clarity. Following the personality assessment task, subjects rated their current levels of pleasure and arousal in the familiar
fashion. Those who rated themselves as feeling at least moderately pleasant (in the P-mood condition) or moderately unpleasant (in the U-mood condition) proceeded directly to the task of autobiographical event generated. Those who did not, were asked to concentrate once again on ideas or images that make them feel pleasant (unpleasant). This “mood boost” period lasted 5-10 min, long enough to ensure that all subjects were experiencing at least a moderate degree of pleasure or displeasure prior to the start of the event generated task. For purposes of this task, subjects were asked to generate an emotionally positive event from anytime in their personal past, in response to each of 8 word probes and to generate an emotionally negative episode in response to each of 8 additional probes. These two types of probes were read by the experimenter in a randomly alternating order; to avoid item-selection artifacts, any probe that was associated with a positive event for one subject was linked to a negative event for another. All 16 of the probe, plus 2 other that were used for practice, were common, concrete, semantically unrelated, and affectively neutral nouns culled from Brown and Ure’s (1969) word norms. On presentation of a given probe, subjects were instructed to say OK as soon as they generated an event that was both appropriately valenced (either positive or negative, as stipulated by the experimenter) and suitably specific (one they could date with precision and describe in detail). If the subjects failed to generate such an event within 2 min, then that probe was skipped and the next one was read. If they succeeded, then the experimenter, after logging generation latency (in seconds), asked the subjects to recount aloud the particulars of the event: where and when (month and year) it occurred, what happened, who was involved, and so forth. All of these details were transcribed by the experimenter. Next, subjects rated (a) how intense was the emotion (perforce either positive or negative) they experienced at the time the
event transpired; (b) how personally important did the event seem to them; and (c) how vividly can they recollect it now. To simplify matters, subjects were provided with a set of index cards containing the pertinent 5-point scale (1 = not at all intense/important/vivid, 5 = extremely intense/important/vivid). Subjects spoke their ratings, and the experimenter wrote them down. After generating, recounting, and rating the final (16th) event, subjects marked a new mood matrix. P-mood subjects were then discharged with a reminder to return, in 2 or 3 days, for a second session of testing, details of which would be divulged at that time. U-mood subject stayed longer—sharing cookies and conversation with the experimenter while rock or reggae music played in the background so that they left the lab feeling no worse than neutral.

Procedure involved in manipulating and measuring mood during the retrieval session duplicated those used in the encoding session. On reporting the requisite rating of mood (viz., at least very pleasant or very unpleasant, irrespective of degree of arousal), subjects began formal testing, which consisted of three tasks: stem completion, personality assessment, and autobiographical event recall. As was the case during the encoding session, subjects recorded their current levels of pleasure and arousal on a new mood matrix both before and after each task. Subjects first task was to speak the first word they thought of that began with a particular three-letter stem, such as ban or bri. Of the 32 stems that were read to the subjects, 16 corresponded to old words (banana, bride, and others they had rated for concreteness, imagery, and goodness of meaning 2 or 3 days earlier), and 16 corresponded to new words (ones they had neither seen nor heard in the course of any encoding session task, including word rating). At the outset of the third and last task, autobiographical event recall, all subjects rated themselves as feeling, at a
minimum, either moderately pleasant or moderately unpleasant (in the P-mood and U-mood conditions, respectively). The experimenter reminded the subjects generated both positive and negative autobiographical events in response to common-word probes. She then asked them to spend 5 min trying to recall, in any preferred order, the gist of as many of these events as possible. Subjects were told that they need not give detailed replies; indeed, it was sufficient—even preferable—for them to recall only the probe that was associated with a given event. The experimenter kept a written record of the subject’s spoken responses. Next, subjects were asked to specify the date (month and year) on which every recalled event had occurred. To make matters easier, the experimenter reminded the subjects of the relevant events by reciting aloud the corresponding probe in the order in which they had been recalled. By comparing these dates with these supplied during the encoding session, they sought to determine whether the subjects were recalling the same events they had generated earlier and whether these events pertained to actual, specific experiences of a personal nature. After marking yet another mood matrix, subjects were debriefed and then discharged once their level of pleasure was neutral or higher. 64 undergraduates (44 women and 20 men) participated in the study in return for course credit. Criteria for selection included first or second-year standing, fluency in English, and no prior participation in research involving either autobiographical memory or experimentally induced moods.

It was found that mean differences in rated pleasure between P-mood and U-mood subjects were significant at the conclusion of both event generated and event recall (viz., AEG and AER ratings). P-mood subjects rated themselves as being more aroused than did their U-mood counterparts, and this was true for all ratings taken after the CMT had
began (viz., BEG, AEG, BER, and AER ratings). A 2X2 (Encoding Mood X Event Type) mood-design analysis of variance (ANOVA) was applied to each of the six measures. The results of these analyses revealed three significant main effects for event type and one marginal main effect for encoding mood. Regarding the former, negative events were not only older than positive events but also more intense and more important. Regarding the latter, U-mood subjects took somewhat longer to generate a given event than did P-mood subjects, a result consistent with the claim that depressed mood impedes cognitive processing. There were no reliable Encoding Mood X Event Type interactions, meaning that there was no sign of any sort of mood congruence. P-mood subjects did not ascribe greater importance to positive than negative events nor did U-mood subjects recollect negative events more vividly than positive ones. Moreover, and in contrast to Teasdale and Fogarty’s (1979) findings, there was no evidence that P-mood subjects generated positive events more rapidly than negative ones. The recall data were analyzed per a 2X2X2 mixed design (Encoding Mood X Retrieval Mood X Event Type). Results were negative in all respects but two. First, there was a marginally significant main effect of encoding mood, such that the probability of recall was higher for events (of either type) that had been generated under U-mood as opposed to P-mood conditions. Given that generation latency was somewhat longer for encode-U than for encode-P subjects, this result may merely reflect a difference in time spent thinking about the events. Second, and more important, there was a significant interaction between encoding and retrieval moods, a greater percentage of events were recalled when encoding and retrieval moods matched than when they mismatched. The difference between the large/small and large/large groups was not significant, the overall pattern is similar to that found by Eich and Metcalfe (1989), thus providing a small measure of
added support for the idea that stronger MDM effects are associated with two-dimensional (pleasure plus arousal) than with one-dimensional (pleasure only) shifts in mood state. Of the 62 correlations that were computed (one for every subject who recalled at least three events), 61 were significant, the overall mean correlation was 0.97. Additionally, of the 407 total events recalled by these subjects, 358 (88%) were calibrated within 1 year of their original date of occurrence.

Experiment 1 showed no sign of mood congruent memory—even in the context of a supposedly sensitive measure (event generation latency)—is at odds with conventional wisdom, which holds that mood congruence, unlike mood dependence, is a robust and readily reproducible phenomenon. These considerations call for experiment 2 geared toward replicating and extending the results of the first. One difference from experiment 1 concerned the requirements of the event generated task. Although subjects in this experiment, like those in the first, were asked to generate a specific autobiographical incident in response to every probe (16 common, neutral nouns), the type or valence of event generated was determined by the subjects themselves rather than by the experimenter. Thus, after describing and dating a given event, but before appraising its importance and vividness, participants in Experiment 2 rated both the original valence and intensity of the event on a 9-point scale ranging from 4 (extremely positive) through 0 (neutral) to 4 (extremely negative). This measure replaced the 5-point scale used to assess intensity alone in the initial study. In Experiment 2, however, these tasks were undertaken first in their respective sessions. Accordingly, subjects now had to rate themselves (on the mood matrix) as feeling at least very pleasant or very unpleasant, irrespective of level of arousal, before they could advance to either task, thereby obviating the need for a mood boost. Another change
in methodology was that all (rather than most) subjects were tested for free recall 2 days after event generation. Finally, the task of event redating was omitted in Experiment 2. This task had served its purpose in Experiment 1 by showing that subjects complied with the request to generate specific autobiographical events during the encoding session and to recall these same events during the retrieval session. Because subjects in the experiment 2 received these same instructions, they saw no point in saddling them with the relating task. 64 undergraduates (39 women and 25 men) participated and criteria for subject selection were the same as those for Experiment 1. Also like the initial study, 16 students were randomly assigned to each of the four combination of encoding and retrieval mood, and all testing was conducted individually. They found that there were no discernible differences among encoding/retrieval conditions in ratings registered at the start of either session. P-mood subjects rated themselves as feeling very pleasant or better prior to both event generated and event recall, whereas U-mood subjects felt very unpleasant or worse. Differences between P-mood and U-mood subjects in posttask ratings, although smaller than those seen pretask, were significant. The mean correlation between the six principal pair of ratings made by subjects in this experiment was .53. A 2X3 mixed-design analysis of these scores revealed a reliable effect of event type. There was also a reliable Event Type X Encoding Mood interaction. Relative to their U-mood counterparts, P-mood subjects rated more of their events as positive. Both the main effect of event type and the Event Type X Encoding Mood interaction were reliable. Whereas P-mood subjects perceived positive events to be more intense than negative events, U-mood subjects saw them the same. None of the four remaining measures of event generation performance-latency, importance, vividness, and age-provided evidence of either mood congruence (Encoding Mood X Event
Type interaction probability levels greater than .10) or a main effect of encoding mood. However, every measure except age did vary reliably as a function of event type. Concerning latency, negative events were centered between positive events at the short end of the time-to-generate distribution and neutral events at the long end & pairwise comparison. Regarding importance, negative events were rated highest and neutral events lowest. Finally, with respect to vividness, neutral events were rated lower than were either positive or negative events. Although neither main effect was reliable, the interaction was robust. Replicating the mood dependent effect found in the first experiment, proportionately more events were recalled when encoding and retrieval moods matched than when they mismatched. 2X2X3 analysis of these percentages revealed a reliable interaction between encoding and retrieval moods, indicating an advantage in recall of matched over mismatched mood conditions. Neither the main effect of event type nor any of its affiliated interactions were significant, suggested that positive, negative, and neutral events were equally susceptible to mood dependence. Results of Experiment 2 did not provide even slim support for the idea that a two-dimensional change in mood is more conducive to demonstrating MDM than is a one-dimensional change. Two results of Experiment 2 rise above the rest in importance. First, when asked to generate specific events from their personal past in response to affectively neutral probes, participants in Experiment 2 recollected more positive episodes, and fewer negative ones, if they were in a pleasant as opposed to an unpleasant mood. Second, when asked to freely recall the gist of these events 2 days later, subjects remembered more if they were tested in the same mood they had experienced during event generation than in the contrasting emotional context. Together, these results mark Experiment 2 as the only single study we know of that has succeeded in demonstrating both mood
congruent and mood dependent memory. This is reason enough to wonder whether the results are replicable—hence Experiment 3. There was another reason for running a third study, one that is traceable to the results of the first. Recall that in Experiment 1, the interval separating the encoding and retrieval sessions was set at 2 days for some subjects and at 3 days for others. They reanalyzed total event recall (i.e., positive and negative events combined) with a view to this difference, and an intriguing pattern emerged. Selected according to the criteria applied in Experiment 1 and 2, 128 undergraduates (87 women and 41 men) served as subjects. The sample contained substitutes for students who either (a) generated specific events in response to less than 75% of the autobiographical memory probe (2 cases); (b) failed to attend the retrieval session (5 cases); (c) became too upset during U-mood induction to proceed further (2 cases); or (d) seemed unable (or unwilling) to attain the requisite pretask level of pleasure or displeasure (24 cases). For a random half of the subjects in this study, the retrieval session met 2 days after the encoding session; for the remaining subjects, the intersession interval was set at 7 days. The factorial combination of these two intervals with two encoding moods (P vs. U) and two retrieval moods (again P vs. U) defined eight conditions, each of which was represented by 16 individually tested subjects. In an effect to enhance the sensitivity of memory measurement, participants in Experiment 3 were asked to generate a specific autobiographical event in response to each of 24 emotionally untinged probes. Otherwise, their task was identical to the one performed by subjects in Experiment 2 (who had generated a maximum of 16 events, divided as they saw fit among positive, negative, and neutral types). As in Experiment 2, autobiographical event generated was the first task undertaken during the encoding session of Experiment 3, and autobiographical event recall was the initial task performed during
the retrieval session. Participants in Experiment 3 were required to reach a critical level of mood (viz., very pleasant or very unpleasant, irrespective of degree of arousal) before they could advance to either task. The methods by which moods were modified and measure paralleled those used in the two preceding studies.

The findings of Experiment 3 demonstrated that rating of pleasure and arousal, averaged across intersession intervals, which did not differ. Subjects in Experiment 3 generated a maximum of 24 rather than 16 events. The patterns proved to be virtually identical in terms of event generation latency, event age, and ratings of event intensity, importance, and vividness. In terms of the number of positive, neutral, and negative events generated, this Experiment showed a smaller mood congruent effect than did Experiment 2. The Encoding Mood X Event Type interaction was reliable, with P-mood subjects generating more positive and fewer negative events than their U-mood peers. A 2X2X2 analysis of the percentages of total events recalled (i.e., positive, neutral, and negative types combined) yielded two notable results. First, the overall level of recall was higher at the 2-day than at the 7-day intersession interval. Second, although the triple interaction among encoding mood, retrieval mood, and intersession interval was significant, there was a reliable advantage of matched over mismatched mood conditions at the 2-day intersession interval. A separate analysis of the recall data supplied by 2-day subjects who had generated at least one positive, one negative, and one neutral event showed a significant interaction between encoding and retrieval moods, but no appreciable effect (whether simple or interactive in nature) of event type. However, instead of seeing a stronger MDM effect at the 7-day interval, they found no effect at all. Analysis by planned comparisons showed that subjects who had experienced a large as
opposed to a small change in both pleasure and arousal recalled fewer total events. Neither of the other two pairwise contrasts was significant.

Eich (1995) investigated the mood as a mediator of place dependent memory. In his study each experiment was conducted in three locations termed the homeroom, the outside environment, and the inside environment. Participants were escorted by the research assistant from one site to another, and as they walked they often talked about matters (e.g., classes, hobbies, and movies) unrelated to the experiment. The main task undertaken during the first session of each study was autobiographical event generation. Participants in experiment 1 and 2 performed this task in either the I (inside) or the O (outside) environment; those in experiment 3 always did so in outside environment. The main task undertaken during the second session of each experiment was autobiographical event recall. This task was performed in the I environment by half of the subjects in each study, and in the O environment by the other half. Like event generation, the task of event recall began following a 5-min acclimation period of which the participants had been apprised in advance. In experiment 1, 32 participants were randomly assigned in equal numbers to one of two conditions: O/I. The letter to the left of the slash mark signifies the environment where participants undertook autobiographical event generation; the letter to the right symbolizes the environment where, 2 days later, participants completed autobiographical event recall. After expressing their views on the study’s aims and methods, participants were asked to describe, in their own terms, the moods they had experienced during the event generation and event recall tasks. The intent here was to obtain a rough idea of whether the participants did indeed feel different (presumably better) being outside than they did being inside. As it means
of assessing performance on this task, six dependent measures were derived for each participant: (a) number of event generated (maximum = 16) (b) mean event generation latency (in second), (c) median event age (month since the event occurred), (d) mean event emotionality (range = -4 to 4), (e) mean event importance (range = 1 to 5), and (f) mean event vividness (range = 1 to 5). Participants undertook the identical event generation task while experiencing either a very pleasant (P) or a very unpleasant (U) mood, affects that had been instilled through a combination of music and imagery. No appreciable difference was found between P and U mood subjects in five of the six measure of generation performance. The overall correlation was positive and significant with both conditions contributing about equally. Thus, even though all participants undertook event recall in an environment that plainly looked different from the one they had inhabited during event generation, they varied considerably in how similar they felt on the two occasions. Those who felt most similar recalled the most events. After being asked to verbalize their moods, 16 of the 32 participants reported feeling happier in the O than in the I environment; how much more cannot be gauged with confidence given the concise and occasionally cryptic nature of many of the participants’ replies. Being outside made 4 participants feel more anxious or agitated. One participant each said that he or she felt either more serene, sadder, or sleepier while in the garden as opposed to the office. The remaining 7 participants claimed that they felt basically the same in both environments. Although most of these participants acknowledged the garden’s aesthetic appeal, it appeared not to have moved them emotionally, for better or for worse.

This study sought to achieve four aims. One was to replicate the positive relation between event recall and subjective similarity seen in the
first study. A second was to discover whether more events are recalled when generation and recall environments match (condition I/I and O/O) than when they mismatch (conditions I/O and O/I). The third objective, which was predicated on the success of the second, was to determine whether the advantage in recall of matched over mismatched condition is more accurately construed as an instance of PDM or one of MDM. Finally, in an effort to measure the moods experienced in the O and I environments, all participants were administered the Positive and Negative Schedule (PANAS) at regular intervals throughout the course of the study. Developed and Validated by Watson, Clark, and Tellegen (1988), the PANAS consists of 20 emotion-relevant adjectives such as attentive and irritable. In second study, participants were instructed to respond to each item on the basis of how they were feeling “right now”. The assistant read aloud the items and recorded the participants’ spoken responses, which were made on a 5-point scale ranging from very slightly or not at all (1) to extremely (5). 48 participants were assigned randomly and in equal numbers to one of four condition: two in which event generation and recall environments matched (I/I and O/O) and two in which they mismatched (I/O and O/I). Participants rated their current mood, via the PANAS, on six occasions: (a) in the home-room, moments after the completion of informed consent and about 5 min before they were escorted to the O environment (the intervening time was taken up with instructions and practice on the event generation task). This rating is coded as GSB, for generation session baseline; (b) immediately before and (c) immediately after event generation (rating BEG and AEG, respectively); (d) in the homeroom, 2 day later, just prior to departure for the recall environment (either I or O). The symbol RSB, short for recall session baseline, denotes this rating; and finally, (e) immediately before and (f) immediately after event recall (ratings BER and AER,
respectively). Event generation in the I environment were indistinguishable from those generated in the O environment; this was so for all six measures of task performance. The most puzzling outcome was that O-generated events were not rated as higher in emotionality than were I-generated events. First, as one would anticipate, there were no reliable differences among conditions in rating of PA made at the outset of either the generation or the recall session (ratings GSB and RSB), for the main effect of condition and the Condition X Rating interaction. Average PA, however, was significantly lower at the second baseline than at the first, probably because the experiment was no longer a novel endeavor. Second, both the O and I environment had significant but somewhat transient effect on ratings of PA. PA increased an average of 2.3 points among participants tested in the garden and decreased an average of 2.3 point among those tested in the office. On average, participants recalled more when they mismatch, a clear-cut case of place dependence. Ratings of subjective similarity mirrored the pattern set by event recall, such that participants felt more similar when generation and recall environments matched than when they mismatched. Similarity of feelings accounted for more of the variance recall than did similarity in settings and most of these high subjective similarity participants had been tested under matched rather than mismatched environmental conditions. Participants with atypical similarity scores showed no sign of PDM.

Since previous experiments provided indirect support for the mood mediation theory of PDM, experiment 3 was carried out to test this theory. 48 participants were randomly assigned in equal numbers to the four generate/recall conditions. As in experiment 2, participants rated their current mood, via the PANAS, on six occasions: twice in the homeroom (to establish baseline for event generation and recall), both
before and after generation, and both before and after recall. The method by which moods were modified at recall was similar to the continuous music technique used in recent research on mood congruent and mood dependent memory (e.g. Eich et al. 1994; Eich & Metcalfe, 1989). Specifically, participants were asked to entertain either elating or depressing thoughts about factual or fictional incidents while mood appropriate music played softly in the background. The music, once started, did not stop until the after-recall PANAS has been completed. A high-grade portable cassette deck with detachable speakers was used in the garden; although small in size, its acoustic qualities rivaled those of the much larger stereo system that sat atop a table in the office. After assessment of the subjective similarity between their feelings at recall and generation but before they were debriefed, participants were asked to share their thoughts on the study’s aims and methods. Participants in Experiment 3 were also asked, how real or genuine was the mood you experienced during the test of recall? Participants were encouraged to answer without most Candor and to frame their response in terms of an 11-point scale ranging from totally artificial (0) to extremely realistic (10). The assistant kept a written record of the participant’s spoken responses. Events generated in the present study tended to be older and to have more positive emotionality than those generated in the same (O) environment in either of the previous studies. Results of the analysis were negative in all respects except two. First, there was a marked main effect of recall mood, such that PA increased an average of 2.6 points after the positive mood induction and decreased an average of 5.4 points after the negative mood induction. Second, there was a marginally significant Recall Mood X Rating Comparison interaction. Relative to their baseline ratings, positive mood participants reached a higher level of PA before rather than after the test of recall, whereas negative mood participants
experienced lower PA before recall than after its completion. There was a significant main effect of recall mood, with positive mood participants showed a decrease in NA of 1.1 points and negative mood participants showed an increase of 3.4 points. The Recall Mood X Rating Comparison interaction was also reliable, indicating that the difference between before-recall and baseline rating of NA was greater than the difference between after-recall and baselines scores.

Baumann and Kuhl (2005) investigated positive affect and flexibility: overcoming the precedence of global over local processing of visual information. This study was conducted to test whether preferences for global processing observed in a positive mood condition imply an impaired ability to use local processing. In order to test the hypothesis that positive affect promotes flexibility to shift the focus to non-dominant (local) stimulus features, they assessed reaction times in a target detection paradigm. Their global-local shape task consisted of single figures that unambiguously contained a given target shape or not. They used an affective priming procedure in order to test mood effects within participants. 78 participants (40 women and 38 men) were recruited through flyers at the University of Osnabruck. Their mean age was 24 years (range 15–57). The material used by Derryberry and Reed (1998), the shape task consisted of geometrical objects (circle, triangle, square, or diamond) composed of smaller geometrical objects. Participants’ task was to decide whether a given target shape was present or not. Participants were asked to respond as rapidly and accurately as possible to target items. Personal prime words were introduced as a means to facilitate attention for subsequent tasks. Participants were asked to attend to these words and recall associated feelings. In order to further increase attention to prime words, at random intervals participants were asked to
classify the affective valence of the prime. The shape-detection task consisted of a total of 156 trials (including 12 practice trials): 50% did not include the target shape, 25% included it on the global dimension, and 25% on the local dimension. Four target shapes (circle, triangle, square, or diamond) were alternated block-wise. Within blocks, the presence/absence and global/local dimension of target shapes as well as the affective valence of prime words were completely balanced. Positive negative, and neutral prime words were presented in random order. The findings revealed that participants responded significantly faster to local targets after positive compared to neutral and negative prime words. The typical precedence of global over local processing observed after neutral and negative prime words was reversed after positive prime words. Furthermore, findings suggest that mood-related preferences in global versus local processing cannot be generalized to processing ability.

In addition to the flexibility-enhancing effect of positive affect, the present data also show a flexibility-reducing effect of negative affect: Negative prime words significantly increased latencies to local targets while not slowing latencies to global targets. This finding is consistent with Personality Systems Interactions (PSI) theory (Kuhl, 2000, 2001). According to the negative affect modulation assumption of PSI theory, negative affect reduces access to extension memory and, as a result, reduces cognitive flexibility. Thus, shifting from the default global to the required local features becomes more difficult under conditions of negative affect. The finding suggests that negative affect does not slow down cognitive processing per se, but reduces cognitive flexibility presumably associated with extension memory.

Fredrickson and Branigan (2005) in their study demonstrated positive emotions broaden the scope of attention and thought-action
repertoires. Experiment 1 examined broadening and narrowing of the scope of attention using a global-local visual processing task. Experiment 2 examined broadening and narrowing of thought-action repertoires using an open ended Twenty Statements Test. Participants were 104 university students (60% women) enrolled in an introductory psychology course. Subjective experiences were assessed using Emotion Report Forms (adapted from Ekman, Friesen & Ancoli, 1980). Participants rated the greatest amount felt of the following nine emotions: amusement, anger, anxiety, contentment, disgust, fear, happiness, sadness, and serenity. Ratings were made on 9-point Likert scales (0 = none, 8 = a great deal). Five videotaped film clips served as the experimental manipulation in this research. Two clips elicited two distinct positive emotions: "Penguins" (2 min 6 sec) shows groups of penguins waddling, swimming, and jumping and primarily elicits amusement. "Nature" (1 min 33 sec) shows fields, streams, and mountains in warm, sunny weather and primarily elicits contentment and serenity. Two additional clips elicited two distinct negative emotions: Witness" (1 min 40 sec) shows a group of young men taunting and insulting a group of Amish passers-by in the street and primarily elicits anger and disgust, "Cliffhanger" (2 min 50 sec) shows a prolonged mountain climbing accident and primarily elicits anxiety and fear. A final clip served as a neutral control condition: "Sticks" (1 min 33 sec) shows an abstract dynamic display of colored sticks piling up and elicits virtually no emotion. The findings clearly demonstrated that when compared to a neutral state, positive emotions broadened the scope of attention in Experiment 1 and thought-action repertoires in Experiment 2. In Experiment 2, negative emotions, relative to a neutral state, narrowed thought-action repertoires. These findings provided empirical support to the broaden and build theory of positive emotions as proposed by Fredrickson (1998, 2001).
Goeleven, DeRaedt, Baert and Koster (2006a) investigated deficient inhibition of emotional information in depression. To study whether impaired inhibition of negative affect is a cognitive vulnerability factor explaining (recurrent) depression, inhibition of positive and negative affective stimuli was examined in hospitalized depressed patients, formerly depressed individuals and never-depressed controls. To investigate inhibitory dysfunctions in the processing of emotional material, they used an affective modification of the negative priming task with pictures of sad and happy facial expressions.

They compared to never-depressed controls, depressed patients showed a specific failure to inhibit negative information, whereas inhibition function for positive material was unaffected. Surprisingly, formerly depressed individuals demonstrated impaired inhibition of negative and positive information.

Goeleven, DeRaedt, and Koster (2007) investigated the influence of induced mood on the inhibition of emotional information. In the present study, they examined the influence of negative and positive mood states on the inhibition of emotion-relevant information in a healthy student population. They also examined the relation between negative mood and cognitive inhibition towards affective information. In two experiments, the same pictorial Negative Affective Priming (NAP) paradigm as in the depression study of Goeleven and co-workers (2006a) was used after a mood induction procedure (MIP; Experiment 1: neutral, negative and positive Velten MIP; Experiment 2: neutral and positive Film and Autobiographical MIP) in a healthy student population.

In experiment first 73 first grade psychology students volunteered to participate in the experiment. Provided that depression has been
associated with reduced inhibition of negative material, a cut-off score of 19 or less on the Dutch version of the Beck Depression Inventory (BDI-II-NL; Beck et al. 1996; van der Does, 2002) was used as inclusion criterion. A total of 67 participants between the age of 18 and 41 were included in the study (55 females, 12 males).

The three groups (neutral, negative, and positive) did not differ significantly with respect to male/female ratio. Participants in the negative mood induction reported significant higher scores on Depression, Anger, and Fatigue, after the MIP as compared to the pre-measurement. Furthermore a significant decrease was observed for Vigour. After the positive MIP, participants scored significantly lower on the Depression subscale. Also a significant decline was observed on the Anger, Fatigue, and Tension, subscales. Concerning the Vigour scale, no pre-post difference was found. Finally, participants in the neutral condition reported a significant decrease for the Vigour scale after the MIP.

A similar effect was obtained on the Tension scale. Between group comparisons of mood state after the MIP revealed a significantly higher Depression score in the negative MIP compared with the neutral condition, in the negative compared with the positive MIP, and in the neutral condition compared with the positive MIP. Furthermore, a significant lower Vigour score was found after the negative compared with the positive MIP. The difference in Vigour score between the negative MIP and the neutral condition was only marginally significant. In addition, a significant higher post score was found after the negative MIP compared with the neutral condition for Tension, and Fatigue, and after the negative compared with the positive MIP for Anger. First, a significant interaction effect of Priming condition X MIP emerged. No
interaction effects were found with Picture valence. Furthermore, the ANOVA revealed a marginally significant main effect for Priming condition. Mean reaction time's show that RT is slower in the experimental condition as compared with the control condition, indicating a mild NAP-effect over all conditions. In addition, a main effect of Picture valence was found. Mean reaction time scores indicated that RT was generally slower for negative trials than for positive trials.

An exploration of the significant Priming condition X MIP interaction effect using contrast analyses on the differences in responding to the control and experimental trials, independent of Picture valence, revealed significantly slower responding on the experimental trials in comparison with the control trials, in the negative MIP. The same significant effect was obtained in the neutral condition. However, in the positive MIP no NAP-effect was found. Contrast analyses revealed no difference in NAP-effect (regardless of Picture Valence) between the negative MIP and the neutral condition, whereas a significant difference in NAP-effect was found between the positive MIP and the neutral condition, and between the positive and the negative MIP condition.

In both cases the NAP-effect was smaller in the positive MIP condition. In the second experiment, the effect of positive vs. neutral MIP on NAP performance was examined in a female-only population. For the mood induction a combination of Film and an Autobiographical recall MIP was used. The Positive and Negative Affect Scale (PANAS; Watson et al. 1988) was conducted as pre/post MIP mood measure. In order to assemble information concerning the sustaining effects of the different MIP's, mood was also measured directly after NAP performance. In addition the 11 statements of the Marlowe-Crowne Social Desirability Scale (MCSDS, Crowne and Marlowe 1960) were included in order to
control for social desirability-tendencies. 60 female psychology students volunteered to participate in the experiment. As in Experiment 1 a cut-off score of 19 on the BDI-II-NL was used as inclusion criterion. Based on this criterion three participants were excluded. Two participants were excluded because of an unreliable NAP measurement. The final sample comprised 55 female participants between the age of 18 and 22. The procedure was similar as in Experiment 1: after fulfilling the informed consent form and the BDI-II-NL, participants were randomly assigned to a positive MIP or a neutral condition. The autobiographical recall phase was always preceded by the film clip. Before and immediately after the MIP, current mood was measured with the PANAS. After the induction procedure, participants conducted the NAP. After performing the NAP, current mood was measured again with the PANAS, followed by the Marlowe-Crowne Social Desirability Scale (1960).

An independent t-test revealed a significant difference in BDI-II-NL score between groups. No main or interaction-effects were found for BDI-II-NL in the analyses on NAP. Participants in the positive mood induction reported significant higher scores on the PA after the induction as compared with the pre-MIP measurement. Furthermore a significant decline was observed for the NA score. In the neutral condition, participants reported no significant difference on the PA scale, from pre to post induction. For the score on the NA scale a marginal significant decline was observed. A comparison of post-mood induction scores between groups revealed a significantly higher PA score in the positive MIP group, compared with the neutral condition. For the NA scale, a significant lower score was obtained in the positive group compared with the neutral condition. The same data pattern was preserved during the post-NAP measurement: the positive MIP group still showed a significant
higher PA-score, and a significant lower NA-score, after performing the NAP as compared with the neutral condition. They explored the effect of social desirability as covariate in an ANOVA on the mood analyses. Importantly, no significant main or interaction effect with this variable was found. Results revealed a main effect of Valence, and a main effect of Priming condition. Both main effects were in the same direction as in the omnibus model of Experiment 1. All other effects were not significant, including the MIP X Priming condition effect.

As expected, in the neutral condition results revealed a main effect for Priming condition. Mean reaction times show that RT is slower in the experimental condition as compared with the control condition, indicating the standard NAP-effect or normal inhibitory processing. In addition, a main effect for Picture valence was found. Mean reaction time scores indicate that RT was generally slower for negative trials than for positive trials.

For the positive MIP group, only a main effect of Picture valence was found, RT is slower for negative trials as compared with positive trials. No main effect was revealed for Priming Condition, indicating an absence of inhibitory functioning.

In this study, the data of both experiments show that attentional inhibition as measured by the NAP task is sensitive to positive mood state. Yet, in contrast with Experiment 1, the expected interaction effect between MIP group and priming condition was not significant in Experiment 2. However, using a within subjects ANOVA analysis (priming condition X picture valence) for each MIP condition separately our specific a-priori hypotheses were confirmed: in the neutral MIP condition normal inhibitory processing (NAP effect) was observed, while
the positive MIP group showed an absence of inhibitory function. Although the data as a whole are indicative that positive affect leads to a reduced or even absence of inhibitory functioning, given these data, conclusions are only tentative at this point.

Reduced inhibition of information at the level of attention might be at the basis of enhanced cognitive elaboration and flexibility observed in positive mood state. It can be stated that irrelevant information which is processed as a consequence of weakened inhibition during positive mood might be linked with task-relevant material in working memory, resulting in an "enrichment" of the memory representation (Hasher et al. 2007). Considering reduced inhibitory control as the underlying micro process of this "broadened", more flexible processing mode matches the growing body of literature which suggest that positive affect leads to better problem solving and creativity possibilities and the ability to see alternative cognitive perspectives (Estrada et al. 1994; Isen et al. 1987). However, it is important to note that the results of this study also fit with other studies in which it is demonstrated that positive mood is associated with increased distractibility, causing poor performance on inhibition dependent processes such as deductive reasoning (Oaksford et al. 1996). Indeed, the abovementioned irrelevant information might cause intrusions and memory lapses (Hasher et al. 2007). An increased activation of task-irrelevant information in working memory as a consequence of reduced attention inhibition could thus be an important factor in enhanced creative divergent thinking and decreased deductive reasoning.

The forgoing exhaustive review of literature reveals (1) there is still inconsistency in the findings obtained by numerous researchers regarding the effect of mood on memory and (2) no study has been undertaken so far to investigate conversation memory of humorous-nonhumorous and
congruent-incongruent version of conversation under different shade of listener’s mood. Hence, further research is needed to resolve the existing inconsistency in findings and to explore how conversation memory (relatively less researched aspect of memory) is affected by mood and versions of conversation. In short in the light of existing findings regarding mood dependent memory and mood congruent memory, it is highly logical to assume that if the listener is in a state of elated mood then he may recall humorous content of the conversation more than nonhumorous content of the conversation. Further more it may also be assumed that if the listener is in a state of depressed/sad mood, he may recall humorous content of the conversation less than the subject with elated mood. It is further hypothesized that both subjects with elated mood and subjects with depressed/sad mood are likely to recall incongruent content of the conversation more as compared to congruent content of the conversation. The present study is designed to test these assumptions also. The findings of the present study will not only open a new area of research in memory but will also provide significant information about the development of interpersonal relations and will highlight the process of impression formation. Furthermore, the findings of the present study may also be applied in education settings like preparing syllabai and teaching methods.