JOURNAL OF VERBAL LEARNING AND VERBAL BEHAVIOR 23, 669–682 (1984)

Decisions and Memory: Differential Retrievability of Consistent and Contradictory Evidence

DENISE DELLAROSA AND LYLE E. BOURNE, JR.

University of Colorado, Boulder

Three experiments are reported that investigate the effect of decision-making on memory. In Experiment 1, subjects were found to recall, following a delay, more facts that supported decisions they had made concerning three texts than facts that contradicted their decisions. Recognition of both types of facts was equivalent, however. The same results were obtained even when recall of both types of facts was equated prior to decision-making (Experiment 2) and when decisional processing was eliminated by simply informing the subject of the correct decision for each text (Experiment 3). On the basis of these results, we concluded that (1) decisions, whether internally generated or externally provided, produce a reorganization of memory traces, (2) this reorganization produces differential accessibility of supporting and contradictory facts, and (3) this differential accessibility produces biased memory performance that can be removed by the use of strong retrieval cues. Two models of memory performance following decision-making are proposed to account for these results.

In the course of making a decision, many facts are usually evaluated and compared. Having reached a decision, how well do people remember the facts which led them to it? Intuitively, it would seem that because those facts were deeply processed, they should enjoy rich memory traces, and hence be easily accessible. Memory for decision-relevant facts should therefore be quite good. Alternatively, because decision-making is a goal-directed process, the decision itself may influence the type of information retained in memory. Once a decision has been made, the information that led to it need be retained only insofar as one believes it will be necessary to justify the decision in the future. For this reason, memory for information encountered during a decision-making episode may be biased toward retention of facts supporting the decision.

Greene (1981) explored memory subsequent to decision-making, and in fact found such biased memory performance. She had subjects listen to ambiguous detective stories, and decide which of two characters was guilty of the crime described in the stories. She found that subjects tended to recognize sentences that were consistent with their decision and often failed to recognize statements that did not implicate their chosen character.

These results, however, do not provide an entirely clear picture of the effects of decisions on memory, for several reasons. First, it is not clear that encoding of information about both characters was equivalent prior to decision-making; Greene provided no measure of subjects' memories for the materials *prior* to announcing their decisions. In a series of studies, we found that enhancing the memorability of some decision-relevant information relative to other

This research was supported by the National Institute of Mental Health under Training Grant 5-T32-MH14617-07, the Office of Naval Research, under Contract N00014-78-C-0433, Contract Authority Identification NR 157-422, and by funds from Martin Marietta Corporation. We thank Seaneen Donnelly for her help in data collection. We also thank Walter Kintsch and Alice Healy for helpful comments and suggestions on earlier versions of this manuscript. Requests for reprints should be sent to Denise Dellarosa or Lyle E. Bourne, Jr., Department of Psychology, Campus Box 345, University of Colorado, Boulder, Colo. 80309.

0022-5371/84 $3.00

decision-relevant information reliably produced decisions that were consistent with the enhanced information (Antos, Bourne, & Kintsch, 1981; Dellarosa & Bourne, 1982). It is entirely likely that the decisions reached by Greene's subjects were themselves the product of biased memory for information implicating one of the two characters in that her subjects knew prior to hearing the texts that they were required to choose a character, and the texts were read to them twice. It would therefore not be surprising that subjects better recognized statements that were consistent with their decisions, since those decisions themselves were a reflection of differential memory for facts implicating the two characters. A more convincing test would require that the texts be processed under incidental learning conditions, without subjects knowledge of a subsequent decision-making task.

Second, if decisions do selectively enhance the memorability of information encountered during a decision-making episode, how does this selective enhancement occur? Greene's results imply that information that is inconsistent with one's decision is forgotten. We believe this is unlikely for two reasons. First, studies reporting memory biases have typically found such biases only in false alarms (as opposed to hits), and only after a delay interval (e.g., Dooling & Christiaansen, 1977). Since Greene found biased performance for both hits and false alarms on an immediate recognition test, it is more likely that memory biases existed prior to decision-making, as stated previously. Second, recovery of "lost" memories, particularly those presumably lost through "overwriting" (e.g., Loftus, Miller, & Burns, 1978) has been shown to be dependent on the instantiation of proper retrieval cues (Bekerian & Bowers, 1983). (We will return to this point in more detail later.) We propose that a decision serves as a powerful retrieval cue, but only for supporting information, thereby producing memory performance that is biased in favor of such information. However, given proper external cues, such as those present in a recognition test, contradictory information should become accessible, thereby eliminating memory biases. To test this hypothesis, we conducted three experiments. In the first experiment, we attempted to replicate the memory biases produced in Greene's study with other materials. Subjects in Experiment 1 read three texts and made decisions about them. They returned 48 hours later and were asked to recall and recognize information from the texts. We predicted, contrary to Greene, that memory biases would be apparent in recall but not recognition performance; that is, subjects would be able to recognize contradictory information that they would not be able to generate during free recall. In the second experiment, we repeated this procedure with the modification that subjects received a recall test immediately following text presentation, thereby providing a measure of memory prior to decision-making. We predicted that memory biases would be apparent only on the second (delayed) recall test, indicating a change in relative accessibility of consistent and contradictory information over the delay period. In the third experiment, the procedure of Experiment 2 was repeated except that subjects were simply informed of the correct decision for each text instead of being required to reach a decision themselves. We predicted that the same pattern of results would be obtained under these conditions, thereby providing evidence that the decision itself is responsible for memory biases, and not the cognitive processes involved in reaching it.

General Method

Materials. Passages were constructed that contained positive or negative information pertaining to 1 of 12 fact categories for 1 of 3 content areas: stock market (S), medical diagnosis (M), and criminal trial (C). Positive passages were defined as ones

which supported a positive decision and negative passages were those which supported a negative decision. Positive decisions for the stock, medical, and criminal texts were, respectively, "buy the stock in question," "hospitalize the patient," and "find the accused guilty." Positive passages included information such as "earnings have increased," "the patient has suffered dizzy spells," or "the accused's fingerprints were found inside the house." An example of positive and negative versions of one passage for each text is presented in the Appendix. The particular mix of positive and negative passages in a text for any given subject was chosen in a pseudorandom fashion, as described below.

Stock market texts were created in the following manner: One positive and one negative passage were constructed for each of 12 fact categories (Sales, Earning, Dividends, Capitalization, State of the U.S. Economy, Growth, Stock Activity, Operating Costs, Market Analysis, Dun and Bradstreet Rating, Credit Worthiness, and Management). The 12 positive and 12 negative passages resided in two separate computer files. Prior to testing a subject, six positive passages (corresponding to six fact categories) were randomly chosen from the positive file by a computer randomization procedure. The negative passages corresponding to the six unchosen fact categories were chosen from the negative file. These 12 passages constituted the text presented to that subject. Thus, the selection of valences for the categories within each text was pseudorandom in nature. Order of passage (category) presentation was randomly determined by a computer randomization procedure. Order of text presentation was balanced across subjects for Experiments 2 and 3 (i.e., SMC, MCS, or CSM). In Experiment 1, the stock text was presented first, followed by the medical and finally the criminal text.

The passages consisted of sentences describing the condition of the category in question. For example, a positive passage for the Sales category contained information to the effect that sales had increased for the company by some specified amount, etc. A negative passage stated that sales had declined, etc. The passages were factual in nature, much like the prose found in a typical annual report of a company. An introductory passage was also constructed describing the text to be read.

Materials for the medical and criminal text were constructed similarly. The 12 medical categories were Vomiting, Fever, Abdominal Cramps, Fatigue, Shortness of Breath, Hematocrit Test Results, Status of Lymph Glands, Numbness, Shoulder Pain, Fainting Spells, Elevated Blood Pressure, and Elevated Pulse Rate. The 12 criminal trial categories were Eyewitness Identification, Possession of Stolen Property, Motive, Alibi, Prior Criminal Convictions, Character Witnesses (2), Knowledge of the Burglarized Home, Courtroom Behavior of the Accused, Fingerprints, Association with Criminals, and Articles Left at the Scene of the Crime. The passages in the medical text itself consisted of sentences describing the condition of a patient with regard to each of the above symptoms. Again, passages were factual in nature, much like the report of a diagnostic lab to a referring physician. An introductory passage for the medical text was also constructed describing the patient whose condition was to be discussed. The criminal text consisted of testimony given during the course of a criminal trial by various witnesses. The case was that of a man accused of burglarizing a house and making off with a camera, a portable television, and some cash. An introductory passage to the criminal trial text was constructed which described the case being tried.

Apparatus. The texts were displayed on a CRT; subjects entered their responses on the keypad of the keyboard. Text presentation and response collection were monitored by a FORTRAN program, which was run on a DEC VAX 11/70 computer.

Procedure. Subjects were tested individ-

ually, and were randomly assigned to either the decision or no-decision group. The FORTRAN program was initiated on the terminals, and randomization of text files and presentation order was completed prior to arrival of the subject. Once the initialization procedure was completed, the subject was seated in front of the terminals and given instructions. The instructions included no mention of decisions to be made or recall tests. Instead, subjects were told that they were assisting in the testing of materials to be used in future experiments, and their task was briefly described. The task involved rating each passage in terms of how positive or negative the content of that passage was. A passage in the stock text was rated as positive if it contained information that would lead one to believe the company would be financially successful (e.g., sales went up); a negative passage was one that indicated financial trouble for the company (e.g., sales went down). In the medical text, a positive passage was one that made the patient appear sick, and a negative passage was one that made her appear healthy. Positive passages in the criminal text were ones that implicated the accused in the crime, while negative passages were ones that supported the innocence of the accused. Following instructions, the subject initiated the presentation of the texts by pressing the "return" key. The introductory passage of one of the texts then appeared. The subject was instructed to press the "return" key again when ready to begin. The passage pertaining to the first randomly chosen category was then displayed. The rating scale, which appeared just under the passage, ranged from 1 to 6, with 6 clearly marked "Positive" and 1 "Negative." (For all the medical and criminal texts, the words "Sick" and "Guilty" appeared under the word "Positive," respectively, and the words "Healthy" and "Not Guilty" appeared under the word "Negative," respectively.) Underneath, prompts were displayed requesting a rating response. The subject indicated a rating choice by typing the appropriate number from the rating scale. The next passage corresponding to the next category was then displayed, and so on. When the last response had been made to the last text, the program displayed a termination message to the subject, and requested that he or she report to the experimenter. Subjects were then given a distractor task lasting approximately 15 minutes. The distractor task was a simple arithmetic test. The experimental tests, as described in the specific procedure sections for each experiment, then ensued.

Experiment 1

In Experiment 1, we attempted to replicate the memory performance biases reported in Greene's study with our stimulus materials, thereby assuring that such biases would be produced with our materials before employing the other manipulations of interest. Subjects read the three texts; one-half were then required to make a decision about each text (i.e., Buy/Don't buy, Hospitalize/Don't hospitalize, /Guilty/Not guilty), while the remaining half were not. All subjects returned 48 hours later, and were given surprise recall and recognition tests. We predicted that subjects' recall performance would show biases favoring recall of information consistent with their decisions, but that recognition performance would be unaffected by such biases due to the strong cues provided by this type of test.

Method

Subjects. Sixteen subjects were chosen from introductory psychology classes at the University of Colorado, Boulder. Subjects were given class credit for participation in the experiment.

Procedure. As described in the General Method section, subjects first rated each passage for positive and negative content and were then given the 15-minute arithmetic distractor task. Following completion of the distractor task, decision materials containing the following questions were

distributed to the decision group: (1) Suppose you were a stock broker. Would you or would you not recommend buying stock in the company you read about? Why or why not? (2) Suppose you were a medical doctor. Would you or would you not recommend hospitalizing the patient? Why or why not? (3) Suppose you were on the jury that heard the testimony given in the burglary trial. In your opinion, is the defendant guilty or not guilty? Why or why not? The subjects wrote down their decisions, and gave brief rationales concerning them. The subjects then engaged in another distractor task for 10 minutes and were dismissed.

Each subject returned 48 hours later, and was required to write down as many of the passages seen during the previous session in as detailed fashion as possible. The subject was then seated at the terminal once more, and given a forced-choice recognition test. In this test, both the positive and negative passages for a particular category appeared simultaneously on the screen. The subject indicated which passage was the one originally presented by typing a "1" or "2" as appropriate. After the subject indicated a response, the next two passages pertaining to the next category appeared, and so on.

Results

Unless otherwise indicated, rejection probability was .05 for all statistical tests.

Recall. Recall protocols were scored in terms of (1) the category recalled, and (2) the valence. More specifically, reproduction of a previously presented statement was scored as a correct recall if *both* the subject matter (e.g., sales) and the valence (e.g., increased/decreased) were correctly reproduced. If an error occurred in either, the entire reproduction was scored as an intrusion error. Maximum correct recall was, therefore, the number of categories presented in each text (12). A passage was scored as a positive or negative passage based on the subject's rating of that passage on the rating task.

Separate unweighted means analyses were conducted on the three texts using as variables: decision type (Positive, Negative, or None), and passage type (Positive and Negative), with repeated measures on the last variable. Because the results for each text were substantially the same, the data from each text were combined, and a single unweighted means analysis was performed on the whole pool using text (Stock, Medical, or Criminal) as an additional between subjects variable.

A significant main effect of text was found, $F(2,39) = 4.68$, $MS_e = 2.66$. Tukey's HSD test of pairwise comparisons indicated that the criminal text was recalled better than the medical text, which was in turn recalled better than the stock text. While these comparisons no doubt reflect a certain degree of differing memorability for the texts themselves, this effect is also confounded with order of presentation, which was not counterbalanced in this particular experiment.

More important, a significant Passage Type × Decision Type interaction was found, $F(2.39) = 8.54$, $MS_e = .86$, $p < .01$, and is presented in Figure 1. As predicted, simple effects tests indicated that positive deciders recalled more positive than nega-

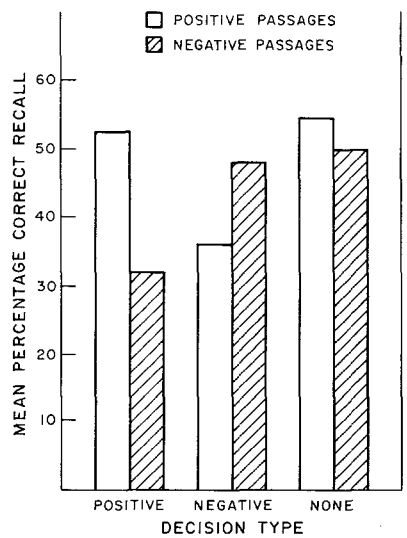

FIG. 1. Mean percentage correct recall as a function of passage type and decision made—Experiment 1.

tive passages, $F(1,30) = 13.47$, $MS_e = .86$, $p < .001$, negative deciders recalled more negative than positive passages, $F(1,39) = 15.52$, $MS_e = .86$, $p < .001$, and nondeciders recalled an equivalent number of both, $F < 1$.

Recognition. The percentages of items recognized in the various conditions are presented in Table 1. No significant results were obtained from analysis of these data, indicating similar recognition memory for all groups.

Discussion

The results of Experiment 1 supported our hypothesis: decision-makers tended to recall more information that was consistent than inconsistent with their decision, while subjects who made no decision recalled an equal number of both types of information. Moreover, no differences were found among the groups on recognition performance, thereby arguing against differential memory structures for decision-makers and in favor of differential cuing afforded by their respective decisions. These results are not definitive, however, for two reasons: First, recognition performance was very near ceiling; hence any true differences in memory structures may have been masked due to this limit on performance. Second, subjects could have made their decisions based on the valence of the majority of information remembered on Day 1. It would not be surprising in that case that the groups differed in their ability to recall the two types of information.

TABLE 1
MEAN PERCENTAGE OF PASSAGES CORRECTLY RECOGNIZED—EXPERIMENT 1

	Passage type	
Decision type	Positive	Negative
Positive[a]	92.3	88.7
Negative[b]	97.2	80.6
None[c]	98.1	93.4

[a] $n = 14$.
[b] $n = 10$.
[c] $n = 24$.

EXPERIMENT 2

In Experiment 2, a stronger test of our hypothesis was undertaken by comparing subjects' memories for the texts prior to and following decision-making. In this experiment, all subjects received a surprise recall test immediately after reading all three texts, and surprise recall and recognition tests after a 1-week delay. We predicted that memory biases would be apparent only on the delayed recall test, that is, only *after* a decision had been made. Memory for both positive and negative material was predicted to be equivalent prior to decision-making.

Method

Subjects. Twenty seven students from introductory psychology classes at the University of Colorado, Boulder, served as subjects in the experiment. Subjects were given class credit for participation in the experiment.

Procedure. To briefly recapitulate the general method, subjects first rated each passage for positive and negative content and were then given a 15-minute arithmetic distractor task.

The materials and procedure in this experiment were the same as in Experiment 1, with five modifications: First, after completing the first distractor task, subjects were given a surprise recall test in which they were asked to write down as many of the previously presented passages as possible. Following this, subjects in the decision group were required to make the same three decisions as in Experiment 1; subjects were not allowed to look back at their recall protocols while making their decisions. The final distractor task followed decision-making. Subjects in the no-decision group were given the distractor task immediately following their recall of the passages.

Second, the delay interval was lengthened from 48 hours in Experiment 1 to 1 week in Experiment 2. (While this produces a possible confound with the new testing

sequence, Experiment 3 will rule out this interpretation.) Third, to more nearly equate the number of subjects in the three resulting groups, two-thirds of the subjects were randomly assigned to the initial decision group and one-third to the no-decision group. Fourth, because recognition performance was so near ceiling in Experiment 1, the recognition test employed in Experiment 2 was not a forced-choice test (which subjects apparently found too easy); instead each of the 12 positive and 12 negative passages from each text was presented to subjects on a CRT one at a time and in random order. Subjects indicated whether each item was old or new and gave a confidence rating for their judgment. Finally, three orders of text presentation were counterbalanced between groups.

Results

Recall and recognition protocols were scored as in Experiment 1. Rejection probability for all tests was .05. For each retention measure, separate unweighted means analyses were conducted on each text, with decision type (Positive, Negative, or None) and order of text presentation (SMC, MCS, or CSM) as between-subjects variables, and Day (Day 1 and Day 2) and passage valence (Positive and Negative) as within-subjects variables. Because the results of each analysis were substantially the same, the data from each text were combined into a single analysis for each retention measure, with text (Stock, Medical, or Criminal) as an additional between-subjects variable.

Recall. The main effects of Text $F(2,54) = 4.52$, $MS_e = 3.49$ and Day $F(1,54) = 53.04$, $MS_e = .44$, were significant. These were modified by a significant Text × Day interaction, $F(2,54) = 4.08$, $MS_e = .44$. Simple effects tests on this interaction indicated no differences in recall of the three texts on Day 1, $F(2,108) = 1.33$, $MS_e = 1.96$, but on Day 2, recall of the three did differ, $F(2,108) = 3.79$. Tukey's test of pairwise comparisons indicated that Day 2 recall of the criminal text exceeded recall of both the stock and medical texts, but that recall of the latter two did not differ.

The Decisions × Passage Type interaction was significant $F(2,54) = 3.37$, $MS_e = 2.85$, but more important, this interaction was modified by a significant Decision × Day × Passage Type interaction, $F(2,54) = 10.07$, $MS_e = .51$, $p < .001$. This interaction is illustrated in Figure 2. Simple effects tests revealed several important results. First, recall of positive and negative information was equivalent among the three decision groups on Day 1 ($F < 1$). On Day 2, however, the groups differed in type of information recalled, $F(2,108) = 9.81$, $MS_e = 1.67$, $p < .01$. Subjects who made positive decisions recalled more positive information than negative information on Day 2, while the reverse was true for those who made negative decisions, F's$(2,162) = 11.58$ and 5.00, $MS_e = 1.07$, p's $< .01$, respectively. Subjects who made no decision recalled an equivalent number of both types of information, $F < 1$.

Second, the recall differences on Day 2 were a result of differential forgetting of the two types of information across the two days by the decision groups. Positive deciders tended to forget negative items, $F(1,162) = 35.56$, $MS_e = .48$, but not positive, $F < 1$, while negative deciders tended to forget positive as well as negative information, F's$(1,162) = 28.42$ and 3.39, $MS_e = .48$. Subjects who made no decisions tended to forget more negative than positive items as well, but this trend did not reach significance, $F < 1$. Taken together, these results would seem to indicate that subjects exhibit a predisposition to forget negative information and that this tendency is facilitated by positive decisions, and attenuated by negative decisions. This relationship is clearly depicted in Figure 2.

Recognition. The mean percentages of hit rates and false alarm rates of the three decision groups for each type of item are presented in Table 2. Unweighted means analyses on hits for each text and all texts

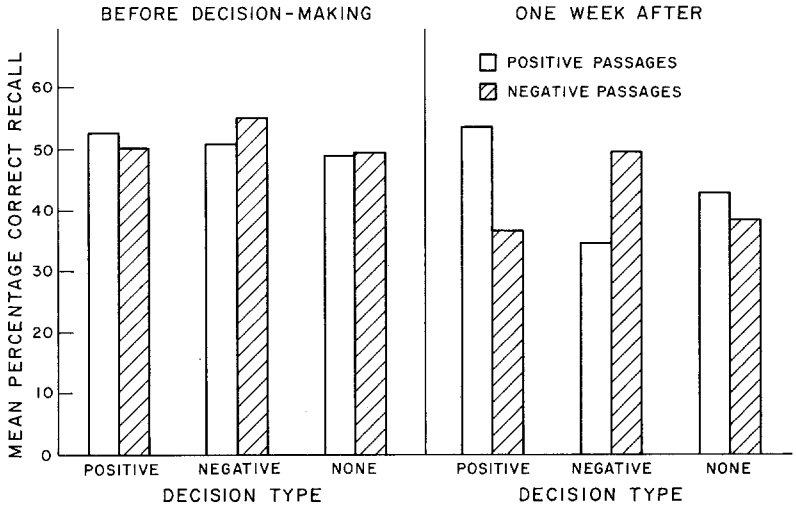

FIG. 2. Mean percentage correct recall as a function of passage type, decision made, and time of test—Experiment 2.

together revealed no significant effects. The level of recognition performance was lower than in Experiment 1, presumably due to the change in presentation format. The failure to find significant differences does not appear to be due to ceiling effects; the highest mean recognition performance was 78%.

Analyses on false alarms returned a single significant effect, that of text, $F(2,54) = 7.77$, $MS_e = .87$, $p < .01$. Tukey's test of pairwise comparisons indicated that subjects made fewer errors to the criminal text than to either the stock or medical texts, while error rates for the latter two did not differ.

Discussion

The results of Experiment 2 are quite clear. Differences in recall of the two types of information after decision-making were attributable solely to the types of decisions made by subjects, and not to differences in memorability of the information prior to reaching a decision. Moreover, memory biases were found only in recall performance and not in recognition, indicating that such biases are a result of differences in accessibility of the two types of information given the decisions made, and not of differences in their respective memory structures. Once sufficient cues are provided for their retrieval, information that is inconsistent with one's decisions can be reproduced.

TABLE 2
MEAN PERCENTAGE OF PASSAGES RECOGNIZED—EXPERIMENT 2

	Passage type	
Decision group	Positive	Negative
Positive	77.8 (8.3)	73.2 (10.9)
Negative	66.7 (10.7)	67.2 (14.2)
None	74.0 (12.3)	64.8 (13.6)

Note. False alarm rates are in parentheses. $n = 27$.

EXPERIMENT 3

The results of Experiment 2 indicated that differential recall of information that is consistent or inconsistent with one's decisions can be attributed to differential accessibility of the two types of information. We interpret this result as a reflection of the power of decisions to cue recall of consistent information only. An alternative explanation would be that it is not the decision per se which produces these results as much as it is the processing subjects employ while reaching their decisions. For ex-

ample, if subjects reach their decisions by successively eliminating certain facts as salient to the determination of guilt, the facts that are eliminated early in the iterative process would receive far less rehearsal than other facts which are retained and evaluated repeatedly. The importance of decisional processing as opposed to the outcome decision itself in producing memory biases was tested in Experiment 3 by eliminating the need for decisional processing. This was accomplished by repeating the procedure described in Experiment 2 with one important change: Subjects were not required to make a decision concerning the facts, but were instead simply told the correct decision for each text. For example, in the criminal text, subjects were told that the defendant confessed to the crime and was found guilty, or that a third party confessed to the crime and the defendant was found not guilty.

Method

Subjects. Twenty-seven students from introductory psychology classes at the University of Colorado, Boulder, served as subjects in the experiment. Subjects were given class credit for participation in the experiment.

Materials and procedure. As in the previous experiments, subjects first rated each passage for positive and negative content and were then given the 15-minute arithmetic distractor task. The texts and recognition tests were the same as in Experiment 2, with one exception. Rather than being given the decision forms, two-thirds of the subjects were shown one of each of the following statements:

Stock

Positive. The passages you read about Eckol (not the company's real name) were chosen from a real report prepared by a stockbroker. The broker thought the company was a sound investment opportunity, and highly *recommended* buying stock in Eckol. A few weeks later, Eckol's stock *increased* significantly in value.

Negative. The passages you read about Eckol (not the company's real name) were chosen from a real report prepared by a stock broker. The broker thought the company was a poor investment prospect, and did *not recommend* buying stock in Eckol. A few weeks later, Eckol's stock significantly *decreased* in value.

Medical

Positive. The passages you read about the runner were chosen from a real report from a medical clinic. The doctor who prepared the report *recommended* hospitalizing the runner because her symptoms indicated the presence of a cardiac inflammatory disorder. The runner was hospitalized under intensive care for 3 weeks, and was unable to resume training until several months later.

Negative. The passages you read about the runner were chosen from a real report from a medical clinic. The doctor who prepared the report did *not recommend* hospitalizing the runner because her symptoms were typical of ordinary exhaustion from over-training. After a few days rest, the runner returned to the clinic. Her symptoms had disappeared, and she was able to resume her normal training activity.

Criminal

Positive. The passages you read about the criminal trial were excerpts chosen from a real trial. During the course of the trial, the defendant *confessed* to having burglarized the house, and as a result was found *guilty*. He was sentenced to 90 days in jail.

Negative. The passages you read about the criminal trial were excerpts chosen from a real trial. The defendant was found (correctly) *not guilty* of burglary. The real thief (who confessed to having burglarized the home) was apprehended several days later.

The statements were typed on a sheet of paper; four statement combinations were used. The remaining third of the subjects served as a control, and were not shown any of the forms. After all retention measures had been collected on Day 2, subjects who were shown a judgment sheet were asked to indicate which judgment had been returned for each text, and to indicate whether they agreed or disagreed with that judgment.

Results and Discussion

Rejection probability was .05. Two sub-

jects disagreed with the judgments delivered on the stock texts, one of them also disagreed with the judgment delivered on the criminal text, and a third subject incorrectly recalled the judgment delivered on the criminal text. Data for these subjects on these texts were excluded from the analyses. The same analyses performed in Experiment 2 were performed on these data.

Recall. The main effect of text was significant, $F(2,50) = 23.44$, $MS_e = 2.39$, $p < .01$. Tukey's test of pairwise comparisons indicated that recall of the criminal text was superior to that of the stock text; the level of recall of the medical text fell between that of the stock and criminal, and did not significantly differ from either.

The main effects of Day and Passage Type were significant, F's$(1,50) = 65.91$ and 4.61, $MS_e = .557$ and 2.17, p's $< .001$ and .05, respectively. These were modified by a significant Decision × Day × Passage Type interaction, $F(2,50) = 7.55$, $MS_e = .48$, $p < .01$, which is depicted in Figure 3. Simple effects tests revealed significant Passage Type × Day interactions within each decision group, F's$(1,100) = 10.18$, 11.44, and 68.94, $MS_e = .48$, p's $< .01$, .01, and .001 for the groups shown positive, negative, and no judgments, respectively. Finer analyses on these interactions indicated the following: Subjects who were shown positive judgments recalled an equivalent number of both positive and negative information on Day 1 ($F < 1$), but recalled a significantly greater number of positive than negative information on Day 2, $F(1,150) = 6.40$, $MS_e = 1.04$, $p < .01$. Moreover, a significant loss of information across the 2 days occurred only for negative information, $F(1,150) = 21.59$, $MS_e = .51$, $p < .01$; positive information was retained nearly intact across the 2 days ($F < 1$). The pattern of results for those subjects shown negative judgments was slightly different. These subjects recalled more positive than negative information on Day 1, $F(1,150) = 4.89$, $MS_e = 1.04$, but recalled an equivalent number of both types of information on Day 2 $F(1,150) = 1.07$, $p > .05$. Here the loss of information across the 2 days occurred for positive information, $F(1,150) = 39.04$, $MS_e = .51$, $p < .001$; negative information was retained well across the 2 days, $F(1,150) = 2.51$, $p > .05$. Thus, while these subjects displayed an advantage for recall of positive information on Day 1, far more of that information was lost during the retention interval than negative information, resulting in equivalent recall of the

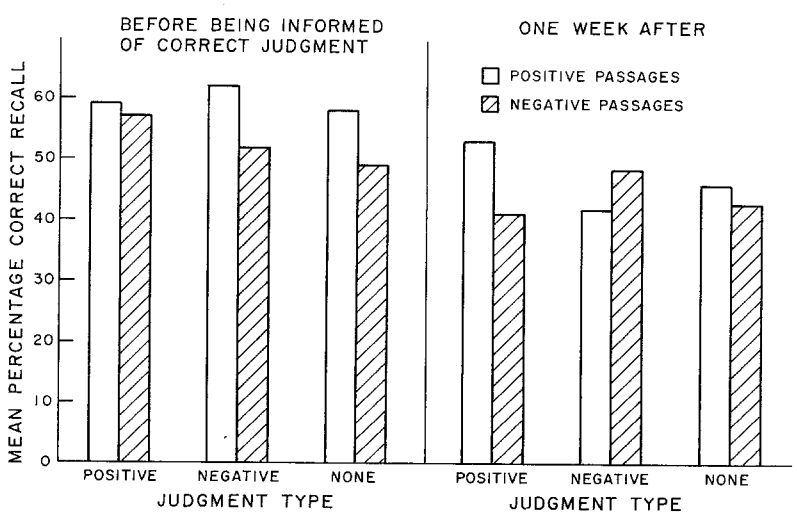

FIG. 3. Mean percentage correct recall as a function of passage type, time of test, and type of judgment presented—Experiment 3.

TABLE 3
MEAN PERCENTAGE OF PASSAGES CORRECTLY RECOGNIZED—EXPERIMENT 3

Judgment group	Passage type	
	Positive	Negative
Positive[a]	76.5 (13.3)	70.7 (16.6)
Negative[a]	74.3 (12.7)	75.7 (17.3)
None[b]	74.7 (10.5)	73.5 (17.9)

Note. False alarm rates are in parentheses.
[a] $n = 25$.
[b] $n = 27$.

two types of facts on Day 2. While these results may reflect the influence of the judgments seen by the subjects, they are also consistent with simple regression to the mean. However, the results of the analyses on the recall of the control group argue against the latter interpretation. These subjects also recalled significantly more positive than negative facts on Day 1, and an equivalent number on Day 2, $F(1,150) = 4.46$ and 2.05, $MS_e = 1.04$, $p < .05$ and $> .05$, for Day 1 and Day 2, respectively. Moreover, significant losses of *both* types of facts occurred across the retention interval, F's$(1,150) = 13.88$ and 7.88, $MS_e = .51$, p's $< .01$ for positive and negative facts, respectively. The overall pattern of results is a preferential retention of positive facts by subjects who were shown positive judgments, negative facts by those shown negative judgments, and equal forgetting of both types of facts by the control group; this overall pattern argues strongly for the influence of the type of post-recall judgment information on subsequent retention performance for the two types of information.

Recognition. The percentage of correct recognition for both types of facts by the three groups are presented in Table 3. An unweighted means analysis on these data returned a single significant result: The main effect of Text, $F(2,50) = 5.99$, $MS_e = 1.67$, $p < .01$. Tukey's test of pairwise comparisons indicated no significant differences although the difference in recognition performance for the stock text and the criminal text was marginally significant. (The required mean difference was .868; the mean difference between the stock and criminal texts was .861.)

Analyses of false alarms returned two significant effects. First, the main effect of Text was significant, Text, $F(2,50) = 5.99$, $MS_e = 6.21$, Tukey's HSD test of pairwise comparisons indicated that, once again, fewer errors were made on the criminal text than on either the stock or medical texts, while the latter two did not differ. Second, the main effect of Passage Type was also significant, $F(1,50) = 5.99$, $MS_e = .66$, indicating that subjects made fewer errors on positive passages than on negative passages.

As predicted, the pattern of results in Experiment 3 mirrored those of the previous experiments. It should be noted that the replication of the results of Experiment 2 where subjects were randomly given judgments, minimizes the importance of any possible confound between testing sequence and delay interval changes in producing the results of Experiment 2. More important, however, the results of this experiment support the claim that it is not decisional processing following initial recall that produces memory biases on subsequent recall performance, but instead the storage of the decision itself in memory.

GENERAL DISCUSSION

The results of these experiments show quite clearly the influence of decisions on memory for decision-relevant facts. In Experiments 1 and 2, subjects recalled more facts that supported their decisions than facts that contradicted their decisions. Despite this clear bias in recall however, recognition for both types of facts was equivalent. The same pattern of results was obtained in Experiment 3, where subjects were simply told the correct decision for each text rather than reaching a decision themselves. In all three experiments, therefore, subsequent information or cognitive activity produced differential accessibility for the two types of information.

These results are similar to those of oth-

er studies in which intervening information is presented to subjects following presentation of stimulus materials. The nature of this intervening material falls roughly into two categories. The first category consists of studies that expose subjects to facts that clearly contradict some aspect of the previously presented stimuli (e.g., Hertel, 1982; Loftus et al., 1978). The second category of studies consists of those in which subjects are required to read a text describing a character and are subsequently given information that identifies the character as a famous person (Dooling & Christiaansen, 1977) or a member of a culture-stereotyped group, such as a lesbian (Snyder & Uranowitz, 1978). The first category of studies differs from the second primarily by virtue of the fact that contradictory information is presented to subjects in the first but not in the second. In the Loftus et al. study, for example, subjects are told that they saw a yield sign when in fact they saw a stop sign; Dooling and Christiaansen, however, simply told their subjects that the text they read was about (e.g.) Adolph Hitler, a fact that was completely consonant with the information contained in the text.

In both cases, subjects exhibit memory performance which is biased toward some aspect of the intervening information. These results, like ours, suggest that subsequent cognitive activity can alter memory for facts presented previously. In particular, they suggest that memory structures are changed or updated in some way to accommodate the introduction of new information. It is the manner in which this updating occurs that is unclear. Bjork (1978), for example, distinguishes between destructive and structural updating of episodic memory. In destructive updating, memories are overwritten (substituted) by new memories. In structural updating, new and old memories are retained intact. Loftus et al. took their results to be an instance of destructive updating. However, in a replication of the Loftus et al. study, Berkerian and Bowers (1983) significantly reduced recognition errors simply by presenting the slides in their proper causal order on the recognition test rather than in random order. The authors concluded that forgetting, in the Loftus et al. study, was not due to "overwriting" of memories, but was instead due to the absence of critical retrieval cues at test time, which, when provided, reveal "old" memories intact. The authors offer no explanation, however, as to why such strong retrieval cues are necessary to uncover original memories.

An alternative explanation, similar to Bjork's structural updating, was proposed by Dooling and Christiaansen. They suggested that new information produces an "active cognitive reorganization" of memories, yielding differential accessibility of various subparts. In cases that involve activation of existing knowledge structures (famous figures or stereotypes), restructuring would yield differential accessibility of information consistent with previous knowledge. As time passes and memory for the details of the experimental episode becomes less and less accessible, the only readily accessible information about the episode may be the fact that a passage was read about a famous figure or stereotype. Once this fact is retrieved, it may serve as a retrieval cue for information consistent with one's general knowledge. This explanation can be extended to cases involving contradictory or misleading information. In such cases, differential accessibility produces a retrieval bias for the most recent input, or information that is consistent with the most recent input.

Our results appear to be an instance of reorganization of episodic memory. The question still remains, however, as to what form this reorganization takes. Figure 4 illustrates one possible model of memory reorganization that would account for our results. Facts that are processed during a decision-making episode are stored in a cluster and are linked together through prior semantic associations or new inter-

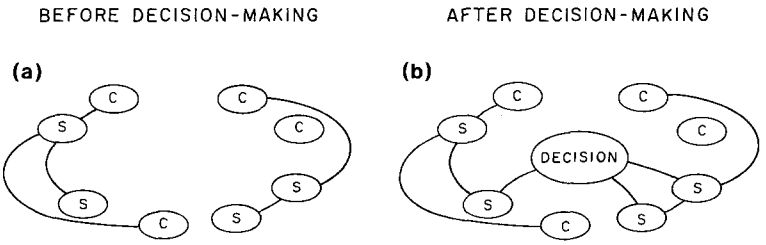

FIG. 4. Descriptive model of the memory representation resulting from an episode of decision-making. *Note:* Nodes marked S represent supporting information; nodes marked C represent contradictory information.

item associations produced during encoding. Following decision-making, the decision itself is stored as a focal point within the information cluster. Moreover, the decision is directly linked to particularly salient or influential supporting facts via retrieval paths. Inconsistent and unimportant facts enjoy only the connections formed during encoding and are not linked directly to the decision itself. Successful retrieval of any fact depends on the success of encountering a retrieval path which leads to it. Since the supporting facts are all interconnected through the decision, the likelihood of retrieving any one fact given retrieval of either the decision itself or one other supporting fact is quite high; inconsistent facts are less likely to be retrieved since they share only a few paths with each other and with supporting facts.

The model, while clearly post hoc in nature, is a form of a linked network, and therefore is testable using standard procedures. For example, the configuration of links implies that a decision should serve to prime verification of supporting facts relative to nonsupporting facts. Moreover, the model implies that if the structure is searched beginning at its focal point (e.g., recall is cued with the decision), the configuration of nodes should be apparent in order of recall, with supporting information appearing prior to contradictory information.

An alternative explanation involves viewing the decision or judgment not as a facilitory retrieval cue but as a filter, or response inhibitor.[1] On this account, both consistent and contradictory information is retrieved, but the decision or judgment acts as a filter which both suppresses output of contradictory information and aids in organizing output around a central fact, decision, or judgment. The nature of the output is determined in this case not by the structure of retrieval paths emanating from a central node, but rather by optimal threshold values differentially set by a central filter for the two types of information.

The central filter interpretation is made all the more attractive if we consider the possibility that decisional processing was not completely suppressed during the rating task. For example, subjects may have built up evidence toward a decision as they read each passage by forming a two-column tally sheet with evidence for a positive decision on one side and negative evidence on the other. At time of recall, the decision could serve to simply suppress the output of an entire column, while leaving intact the organization of data in support of the decision. In this case, no reorganization of episodic memory would be required.

While the filter interpretation offers a viable alternative to the reorganization interpretation, postulating a memory structure as highly organized as presupposed by the tally sheet version leads to two objections. First, informal interviews with subjects prior to dismissal indicated that some of them were aware of producing a decision

[1] This account was suggested by an anonymous reviewer.

while reading the criminal text; however, no subject reported making decisions concerning the stock or medical texts. Despite this difference, the same recall and recognition patterns were found in all three texts. Such evidence (albeit weak) of the absence of decisional processing is not very damaging to this account, however, in that any organizational process which involved a tally sheet differentiating the two types of information when combined with a recall filter would produce the observed memory biases. However, the requisite organizational process leads to the second objection. As mentioned earlier, biased memory performance has been reported by several researchers employing a variety of procedures. The materials used in those studies do not readily lend themselves to a bifurcating organizational procedure. It is not likely for example, that Dooling and Christiaansen's subjects divided textual information into that which was consistent with Hitler's activities and that which was not during reading. Nonetheless, differential thresholds imposed by a central fact (i.e., the text was about Hitler) could produce the observed memory biases, as could the addition of retrieval paths from the central fact to supporting information.

APPENDIX: Sample Positive and Negative Passages Used in Experiments

Positive version	Negative version
Stock text	
On its third quarterly report, Eckol reported a 15% increase in earnings over last year's figures, forecasting record net profits for the year.	On its third quarterly report, Eckol reported a 15% decrease in earnings over last year's figures, forecasting a record net loss for the year.
Medical text	
Results of the hematocrit test indicated a significantly depressed red blood cell count, bordering on the low-normal range.	Results of the hematocrit test indicated a normal red blood cell count, falling within the low-normal range.
Criminal text	
Donald Carter: "Yes, sir, I am a neighbor of the Vanderbilts. On the night of the crime, I was out for about 45 minutes walking my dog. It was about 9:00 PM when I saw Michael Davis, the accused in our neighborhood."	Donald Carter: "Yes, sir, I am a neighbor of the Vanderbilts. On the night of the crime, I was out for about 45 minutes walking my dog, around 9:00 PM. I saw many people, but I did not see Michael Davis, the accused, in our neighborhood."

References

Antos, S., Bourne, L. E., Jr., & Kintsch, W. (1981). Text-based decisions: Changes in the availability of facts over time. (Tech. Rep. No. 106-ONR). University of Colorado, Boulder.

Bekerian, D. A., & Bowers, J. M. (1983). Eyewitness testimony: Were we misled? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **9**, 139–145.

Bjork, R. A. (1978). The updating of human memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 12). New York: Academic Press.

Dellarosa, D., & Bourne, L. E., Jr. (1982). Text-based decisions: Changes in the availability of facts due to instructions and the passage of time. Tech. Rep. No. 115-ONR). University of Colorado, Boulder.

Dooling, D. J., & Christiaansen, R. E. (1977). Episodic and semantic aspects of memory for prose. *Journal of Experimental Psychology: Human Learning and Memory*, **3**, 428–436.

Greene, E. (1981). Whodunit? Memory for evidence in text. *American Journal of Psychology*, **94**, 479–496.

Hertel, P. T. (1982). Remembering reactions and facts: The influence of subsequent information. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **6**, 513–529.

Loftus, E. F., Miller, D. G., & Burns, H. J. (1978). Semantic integration of verbal information into a visual memory. *Journal of Experimental Psychology: Human Learning and Memory*, **4**, 19–31.

Snyder, M., & Uranowitz, S. W. (1978). Reconstructing the past: Some cognitive consequences of person perception. *Journal of Personality and Social Psychology*, **36**, 941–950.

(Received July 22, 1983)
(Revision received November 11, 1983)