Technical Report 731

Unaware Memory in Hypothesis Generation Tasks

Judith E. Brooks and Bethany H. Drum

Instructional Technology Systems Technical Area
Training Research Laboratory

U. S. Army
Research Institute for the Behavioral and Social Sciences

December 1986

U. S. ARMY RESEARCH INSTITUTE

FOR THE BEHAVIORAL AND SOCIAL SCIENCES

A Field Operating Agency under the Jurisdiction of the

Deputy Chief of Staff for Personnel

Technical review by

Sharon Mutter
Ok-Choon Park

NOTICES

DISTRIBUTION: Primary distribution of this report has been made by ARI. Please address correspondence concerning distribution of reports to: U.S. Army Research Institute for the Behavioral and Social Sciences, ATTN: PERI-POT, 5001 Eisenhower Ave., Alexandria, Virginia 22333-5600

FINAL DISPOSITION: This report may be destroyed when it is no longer needed. Please do not return it to the U.S. Army Research Institute for the Behavioral and Social Sciences.

NOTE: The findings in this report are not to be construed as an official Department of the Army position, unless so designated by other authorized documents.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER ARI Technical Report 731	2. GOVT ACCESSION NO. ADA179859	3. RECIPIENT'S CATALOG NUMBER

4. TITLE (and Subtitle) UNAWARE MEMORY IN HYPOTHESIS GENERATION TASKS	5. TYPE OF REPORT & PERIOD COVERED June 1984 - February 1986
	6. PERFORMING ORG. REPORT NUMBER --
7. AUTHOR(s) Judith E. Brooks and Bethany H. Drum	8. CONTRACT OR GRANT NUMBER(s) --
9. PERFORMING ORGANIZATION NAME AND ADDRESS U.S. Army Research Institute for the Behavioral and Social Sciences 5001 Eisenhower Ave., Alexandria, VA 22333-5600	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 20263743A794 312H1
11. CONTROLLING OFFICE NAME AND ADDRESS U.S. Army Research Institute for the Behavioral and Social Sciences 5001 Eisenhower Ave., Alexandria, VA 22333-5600	12. REPORT DATE December 1986
	13. NUMBER OF PAGES 34
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) --	15. SECURITY CLASS. (of this report) Unclassified
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE --

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

--

18. SUPPLEMENTARY NOTES

--

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Unaware memory
Recognition
Memory priming
Depth of processing
Hypothesis generation

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Memory researchers have distinguished two forms of memory: deliberate recollection of prior events versus the unaware influence of prior events on the performance of a later task. This research investigated the nature of this distinction by further delineating the conditions under which aware and unaware memory are observed to be independent.

Two experiments tested whether items presented for study would influence performance on a hypothesis generation task regardless of subjects' ability to

(Continued)

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

ARI Technical Report 731

20. (Continued)

recognize the items. Depth of processing of the study items was also manipulated to test whether this variable would have a different effect on hypothesis generation and recognition. The hypothesis generation task in Experiment 1 required subjects to generate category instances. In Experiment 2, subjects formed hypotheses about the possible use of a described land area. After the hypothesis generation phase of each experiment, primed hypotheses were tested for recognition.

Data analyses for both experiments revealed that particular hypotheses were primed by study items and that priming was unrelated to recognition performance. Level of processing of the study items influenced recognition but not priming. These results suggest that hypothesis generation, a relatively complex cognitive task, may be added to the growing list of tasks in which unaware memory is observed. The data also encourage the exploration of unaware memory in a variety of tasks, including the systematic comparison of situations in which the phenomenon does and does not occur.

Technical Report 731

Unaware Memory in Hypothesis Generation Tasks

Judith E. Brooks and Bethany H. Drum

Instructional Technology Systems Technical Area
Zita M. Simutis, Chief

Training Research Laboratory
Jack H. Hiller, Director

U.S. ARMY RESEARCH INSTITUTE FOR THE BEHAVIORAL AND SOCIAL SCIENCES
5001 Eisenhower Avenue, Alexandria, Virginia 22333-5600

Office, Deputy Chief of Staff for Personnel
Department of the Army

December 1986

Army Project Number
2Q263743A794

Education and Training

Approved for public release; distribution unlimited.

ARI Research Reports and Technical Reports are intended for sponsors of R&D tasks and for other research and military agencies. Any findings ready for implementation at the time of publication are presented in the last part of the Brief. Upon completion of a major phase of the task, formal recommendations for official action normally are conveyed to appropriate military agencies by briefing or Disposition Form.

FOREWORD

The Instructional Technology Systems Technical Area of the U.S. Army Research Institute performs research and development to improve officers' use of problem-solving and information-processing skills in planning, operations, and decision making. The identification of cognitive capabilities and limitations is fundamental to this effort, forming part of the research base from which improved instructional methods and other products can be developed for the Army training community.

This report describes an investigation of human capabilities and limitations pertaining to memory and the role that these play in tasks that vary in their demand for attempts to remember information. The findings reported here contribute in particular to our understanding of the memory processes involved in hypothesis generation, a skill that is essential to effective problem solving and decision making.

Further investigation and development of the ideas presented in this report will allow us to predict the kinds of tasks, in addition to hypothesis generation, that may be unintentionally influenced by prior information. Furthermore, it will help guide the development of improved, more sensitive techniques for assessing learning.

EDGAR M. JOHNSON
Technical Director

UNAWARE MEMORY IN HYPOTHESIS GENERATION TASKS

EXECUTIVE SUMMARY

Requirement:

The requirement was to investigate the distinction between aware and unaware forms of memory and to assess the involvement of unaware memory in the performance of hypothesis generation tasks.

Procedure:

Two experiments tested whether prior information would affect performance on a hypothesis generation task independently of whether the information could be recognized. The extent to which the information was meaningfully processed was varied to assess whether this factor would have a different effect on hypothesis generation and memory recognition performance. The basic procedure for both experiments was to present information for study and then to present a set of problems that required subjects to formulate hypotheses. In the last phase of the experiment, subjects were tested for recognition of any primed hypotheses.

Findings:

The data from both experiments indicated that prior information primed the generation of particular hypotheses and that the priming effect was independent of subjects' ability to recognize the information. The distinction between aware and unaware memory was thus supported in the context of the relatively complex task of hypothesis generation.

Utilization of Findings:

The research findings provide information in support of current efforts to improve officers' use of problem-solving and information-processing skills in planning, operations, and decision making. Specifically, knowledge of the memory processes involved in hypothesis generation supports the development of techniques for improving this essential skill and contributes to our ability to predict the kinds of situations in which task performance may be unintentionally influenced by prior exposure to information. The findings may also help guide the development of improved techniques for assessing learning.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

UNAWARE MEMORY IN HYPOTHESIS GENERATION TASKS

INTRODUCTION

This report describes two experiments that were performed to support the development of techniques to improve military officers' cognitive job performance. In particular, the research focused on hypothesis generation, a critical cognitive skill that supports problem solving and information processing activities in tasks such as planning, operations, and decision making. Current theoretical views of memory were applied to the problem of hypothesis generation in an attempt to better understand how people formulate hypotheses in response to problem situations. The research results presented in this report have implications for training hypothesis generation skills and for memory assessment. The remainder of this introductory section discusses the relevant background literature leading up to the research and provides an overview of the research design and rationale.

Background Literature

In recent years, memory researchers have made an interesting distinction between two forms of memory. One form is characterized by the conscious recollection of experienced events and is observable in traditional memory tasks such as recognition and recall. Another form of memory is sometimes observed in tasks such as lexical decision, perceptual identification, and word fragment completion. These tasks, which do not require the performer to recollect previously experienced events, nevertheless appear to be facilitated or influenced by prior exposure to relevant events. Indeed, since the prior events themselves often cannot be consciously retrieved, this latter form of memory revealed in priming effects has been referred to by some researchers as an unaware form of remembering (e.g., Jacoby & Witherspoon, 1982).

Three research approaches employing tasks that do or do not require deliberate retrieval of previously presented information have yielded evidence for a dissociation between aware and unaware forms of memory. That is, the two forms of memory appear to rely on different kinds of information or on different memory processes. First, experimental manipulations of variables such as the depth of initial processing and the time interval between exposure and test have been shown to differentially affect performance on the two types of tasks (e.g., Jacoby & Dallas, 1981; Masson, 1984; Tulving, Schacter, & Stark, 1982). A second approach has been to demonstrate that recognition performance can be statistically independent of performance on tests of unaware memory (e.g., Eich, 1984; Jacoby & Witherspoon, 1982; Tulving et al., 1982). Third, comparisons between amnesic and normal subjects have revealed performance differences in recall and recognition but not in tasks that do not require deliberate remembering (e.g., Graf & Schacter, 1985; Kihlstrom, 1980; Warrington & Weiskrantz, 1970).

It is not clear what the memory mechanisms are that underlie this observed dissociation. Possible interpretations have been offered within the context of Tulving's (1972) distinction between episodic and semantic memory systems (e.g., see Jacoby & Witherspoon, 1982). Aware memory may rely on retrieval of an episodic memory trace of an item's presentation, whereas unaware memory may occur when an item's prior presentation has lowered the activation threshold for, or has primed that item's representation in semantic memory. An alternative possibility raised by Tulving (1983) is that priming effects may reflect the existence of "free radicals", fragments that have been detached from episodic memory but that have not been incorporated into semantic memory.

Though an adequate theoretical account of the distinction between aware and unaware memory awaits considerable more study, the distinction is important for at least two reasons. For one, it suggests a need for sensitivity to the different ways in which memory can be revealed, and therefore has implications for memory assessment. The dissociation is also important for its potential to influence our understanding of memory systems and ways in which events are represented in memory. The current research was designed to further investigate the nature of the dissociation by helping to delineate the conditions under which it is obtained.

Researchers have typically studied the distinction by comparing subjects' performance on recognition, free recall, or cued recall tasks with performance on a variety of tasks that do not require deliberate retrieval. The latter tasks can be divided into two broad categories. One category consists of those tasks that require reprocessing of the same stimuli. For example, performance on word identification (Jacoby & Dallas, 1981) and lexical decision (Scarborough, Cortese, & Scarborough, 1977) tasks is facilitated following presentation of those same words in an earlier phase of the experiment. This facilitation may be due to the relative ease of reprocessing the word's features. Similarly, rereading speed for sentences and for typographically inverted text is faster even days and months after initial presentation compared to reading speed for comparable new materials (Kolers, 1976; Masson, 1984).

A second category of tasks requires subjects to generate information from memory. Within this broad category, at least two types of tasks can be further distinguished. In one type, there is a partial overlap of sensory or perceptual information between the initial item presentation and subsequent test information that can be used by subjects as a cue to generate the original item. A common task in this subcategory is word completion, whereby subjects are given word fragments and are to complete them with appropriate words, some of which are from the study list (e.g., __RIC__ for APRICOT). Memory is evidenced when subjects are more likely to complete word fragments formed from study words than word fragments formed from new words (e.g., Tulving et al., 1982). Anagram solving (e.g., IHTAB can be unscrambled to form HABIT) is another kind of generation task that involves some perceptual overlap of presentation and test items. Research has shown that anagrams for words that were presented earlier in the experiment are more likely to be solved within a time limit than are anagrams for words that were not previously encountered in the experiment (Radtke & Englert, 1983).

In the other subcategory of generation-type tasks, there is no sensory or perceptual overlap of information between items presented for study and cues for those items at test. For example, Graf, Shimamura, and Squire (1985) presented a random list of items from different conceptual categories to amnesic patients and control subjects. In the subsequent test phase, category labels were presented as cues, and subjects were to generate the first few exemplars that came to mind. Even in amnesic patients who could not later recall the exemplars, a priming effect was evidenced in that previously presented exemplars were more likely to be generated than would have been expected by chance. The authors concluded that sensory and perceptual overlap between study and test cue items is not essential for this type of memory to be observed.

The data from memory tasks that require generation processes, particularly when there is little sensory or perceptual overlap between study and test, are intriguing. In particular, they suggest to us the possibility that an unaware form of memory may be revealed in a variety of tasks that have not yet been examined. Many everyday kinds of memory activities do not involve a great deal of sensory or perceptual overlap of information between an event's occurrence and its memory cue. A comparison between aware and unaware remembering in the context of such tasks therefore may yield information of particular value to researchers concerned with the pragmatic implications of this line of research for understanding memory functioning. Enhanced understanding of the conditions under which memory without awareness is observed also contributes to the development of theories that attempt to describe and explain the nature of the differences between aware and unaware forms of memory.

Current Research

Two experiments were devised to test the possibility that items presented for study may influence performance on a relatively complex task requiring generation processes, without subjects' conscious recollection of the previously presented items. The selected task, hypothesis generation, was chosen for two reasons. First, it is a cognitive activity that people engage in frequently, and second, it seemed a natural extension of the recently studied category instance production task. The specific task used in the first experiment was a modification of the category instance production task used by Graf et al. (1985) and Kihlstrom (1980). In the second experiment, the priming task required subjects to form a hypothesis regarding the possible use of a described land area.

The general procedure was the same for the two experiments. First, subjects saw a list of items and made an incidental judgment regarding each one's meaning or physical features. Following a short distractor task, subjects were asked to solve a set of problems, each of which required the generation of a hypothesis in response to some limited information. Subjects were then instructed to look back over their responses to the problems and to mark any that they recognized as having been presented in the initial part of the experiment.

The study allowed three questions to be addressed. The first was whether the presentation of items in the first phase would prime those same items as responses in the experimental task, hypothesis generation. The second question was whether primed items would be recognized as having been presented earlier. If memory for earlier presented items was revealed in priming performance but not in recognition performance, this would provide evidence for unaware remembering in the types of hypothesis generation tasks being considered here. The third question was whether item processing depth in phase one would have a differential effect on priming and recognition memory performance. An observed effect of depth on recognition but not on hypothesis priming would replicate the results of previous experiments (e.g., Jacoby & Dallas, 1981; Tulving et al., 1982) that have shown a functional dissociation between aware and unaware remembering.

EXPERIMENT 1

Method

Subjects. Subjects were 80 undergraduate students enrolled in psychology courses at George Mason University. All participated voluntarily and received a $5 payment and/or course credit.

Design. The experiment was a 2 x 2 mixed factorial design. The between-subjects factor was the type of orienting task performed by subjects as individual words were presented for study during the first phase of the experiment. Half of the subjects judged whether or not each word contained a particular letter of the alphabet (physical feature orientation), and half judged whether each word's meaning was pleasant or unpleasant (semantic orientation). The within-subjects factor was the type of problem (experimental versus control) solved by subjects in the second phase of the experiment. During the hypothesis generation (second) phase, subjects were presented both with experimental problems that could be completed with target items previously seen in the study phase, and with control problems that could not be appropriately completed using any phase one study items.

Materials. The materials designed for the hypothesis generation phase of the experiment consisted of a set of short fill-in-the-blank problems. For each problem, the blank was to be filled in with the name of an instance of a conceptual category. The cue information contained within each problem consisted of the name of the category (e.g., four-legged animal) plus other information that served to narrow the number of possible responses while still allowing more than one or two appropriate answers. The following is an example:

"X" is a four-legged animal that runs fast and
lives in the wild.
"X" is a(n)______________.

An original set of 25 problems was developed and pilot tested with 70 students enrolled in undergraduate psychology courses at local universities. Based on the pilot data, 11 of the problems were reworded and then administered to an additional group of 50 students to ensure that each problem would elicit a range of appropriate responses. Three of the problems were finally dropped from the set because, in each case, a given exemplar accounted for more than half of all the responses to the problem. The remaining 22 problems comprised the set of problems used in the experiment.

For each of these problems, a target response (hypothesis) was selected based on the frequency with which it had been given as a response during pilot testing. Responses that occurred with relatively low frequency were selected to serve as target solutions. On the average, each target accounted for only 14% of all the responses that had been given to the problem (the range was 9%-28%), and a target was never the most frequent response. For the example problem above, lion, which was the fourth most popular answer and which accounted for 12% of the responses, was designated as the target. Care was taken to ensure that target items were not homonyms and that a given target would be an appropriate response to no more than one problem.

The 22 problems and their corresponding target solutions were equally divided into two sets, A and B. The assignment of problems to sets was accomplished by matching the target responses on frequency and dominance values taken from the pilot data. Thus, in terms of the popularity of the designated targets as solutions to the problems, the two sets were nearly equivalent. For half of the subjects, the Set A problems were experimental problems, and the Set B problems were controls. The reverse was true for the other half of the subjects.

Two study lists to be used during the first phase of the experiment were constructed using target items. Each list consisted of the 11 items that were the designated targets for the set of 11 experimental problems (Set A or Set B). Three buffers appeared at the beginning and at the end of the list for a total list length of 17 items. A slide was made of each list item printed in capital letters, and two random presentation orders of each list were developed.

Booklets for the second, hypothesis generation phase of the experiment were constructed by randomly ordering the 11 experimental and 11 control problems. Each problem appeared on a separate page of the 22-page booklet, and 2 different random orders were employed. Additional materials and equipment included the forms used by subjects for recording yes/no judgments during the phase one orienting task; booklets containing arithmetic problems for use during a distractor interval; a Kodak carousel slide projector with a timer; and a stopwatch.

Procedure. Testing was done in groups of 10 or fewer subjects. The first part of the experiment involved presentation of a 17-item study list of words via slides at an 8-second rate. Subjects were instructed to make a yes/no judgment for each item as it appeared, by circling YES or NO on their answer sheet. Subjects in the letter search condition judged whether the word contained a particular letter of the alphabet which was announced out loud by the experimenter as the word appeared on the screen. In the semantic judgment condition, subjects indicated whether or not the meaning of each word was pleasant. At the end of this task, answer sheets were collected, and each subject was given a set of arithmetic problems to work on for 5 minutes. These were collected at the end of the 5-minute period.

The next part of the experiment was the critical, hypothesis generation phase in which subjects received a booklet of 22 fill-in-the-blank problems. Subjects were instructed to write down the first response that came to mind for each problem. No mention was made of the connection between the earlier presentation list and the set of problems. This task was timed such that subjects were given 10 seconds to read and respond to each problem in the booklet. The experimenter used a stopwatch to time this task and to pace subjects through the booklet.

The final part was a recognition task in which subjects were to look over the hypotheses they had just generated as answers to the problems, and to circle any that they recognized as having been presented at the beginning of the experiment. The recognition task was self paced and took approximately 2 minutes. Booklets were then collected, and subjects were debriefed.

Results

Hypothesis generation. The first set of data to be analyzed were the hypotheses that subjects generated as responses to the 22 fill-in-the-blank problems. The goals of analyzing the generated hypotheses were, first, to assess whether subjects indeed exhibited priming by generating more previously-presented targets than not-previously-presented targets as responses to the problems. Second, it was of interest to determine whether the degree of any such priming effect was related to the way subjects had been induced to process previously presented target items in phase one of the experiment.

As a preliminary step, an analysis of variance (ANOVA) was performed to verify that there were no effects due to the particular ordering of study items or problems, nor effects due to any peculiar differences between problem sets A and B. The dependent variable in this analysis was the total number of designated target hypotheses that were generated to the problems. The independent between-subject variables were type of judgment (pleasantness versus letter), random study and test order (Order 1 versus Order 2), and problem set designated as experimental (Set A versus Set B). The independent within-subject variable was problem type (experimental problems whose targets were presented in phase one versus control problems whose targets were not presented in phase one). For the evaluation of effects in this and in all subsequent analyses, a .05 significance level was adopted. No main effects or interactions involving order or assignment of problems to condition approached significance, allowing us to collapse the data across these variables.

The main analysis, performed on the collapsed data, was a 2 x 2 mixed ANOVA, with type of judgment during phase one (pleasantness versus letter) as the between-subject variable and type of hypothesis generation problem (experimental versus control) as the within-subject variable. Means and standard deviations are shown in Table 1.

The main effect of type of problem was significant, $F(1,78) = 14.45$, $MSe = 1.56$, $p<.01$. Overall, subjects generated significantly more of the designated target hypotheses for the experimental problems (whose target responses had been previously presented) than they did for the control problems (whose designated targets had not been previously presented). The Table 1 means

Table 1

Number of Targets Generated as a Function of Type of Judgment and Type of Problem (Experiment 1)

	Type of problem		
Judgment	Experimental	Control	Mean
Pleasantness			
Mean	2.25	1.53	1.89
SD	1.21	1.01	
Letter			
Mean	1.60	0.83	1.21
SD	1.28	1.01	
Mean	1.93	1.18	

reveal that both the pleasantness and the letter judgment groups demonstrated this priming effect, and that the size of the effect was nearly equivalent for the two groups. Moreover, the equivalence of the priming effect for the two groups is supported by the absence of a statistically significant judgment condition x problem type interaction.

The main effect of judgment type on target generation was also statistically significant, $F(1,78) = 17.91$, MSe = 1.02, $p < .01$. Subjects who judged the pleasantness of items during the target presentation phase subsequently generated more targets to both experimental and control problems than did subjects who had made letter judgments.

Recognition. Recognition data were analyzed for all subjects who had generated at least one of the previously presented targets as a response to an experimental problem during the hypothesis generation phase. The probability of recognizing a target, given that a target had been generated, was calculated for each subject by dividing the total number of recognized targets by the total number of generated targets. In addition, the value of the probability was corrected for the individual's false alarm rate (the general tendency to incorrectly recognize items that in fact had not been seen before in the experiment), using the standard correction formula, p(recognized generated) - p(false alarm) / 1 - p(false alarm).

The corrected, conditional probabilities were compared for subjects in the two judgment conditions to assess whether recognition of targets was statistically higher for subjects in the pleasantness judgment group than for subjects in the letter judgment group. The mean conditional probability was .82 for the pleasantness judgment group (N = 38) and .38 for the letter judgment group (N = 32). A t-test indicated that this difference in recognition performance was statistically significant, $t(68) = 5.31$, $p < .001$.

As an alternative way to examine the recognition data, we calculated overall recognition failure rates for the two groups. The result was that the pleasantness judgment group failed to recognize only 12% of the targets that they had generated. The letter judgment group, on the other hand, failed to recognize 53% of the targets that they had generated. The results of both recognition data analyses therefore show that semantic processing of the target items during the initial presentation phase enhanced subjects' ability to later recognize those targets. A more superficial type of target processing that focused on physical features, rather than on meaning, led to relatively poor recognition, even though subjects had just generated those targets as responses to the hypothesis generation problems.

Discussion

An important Experiment 1 outcome was the evidence for a priming effect in the hypothesis generation task that we employed. With respect to analysis of the hypotheses that subjects generated in phase two, the significant main effect of problem type showed that subjects were more likely to generate the designated target responses for experimental problems, whose targets had been presented in phase one, than they were to generate the designated targets for control problems whose targets had not been previously studied.

While this observed priming of relatively uncommon hypotheses is an interesting finding in itself, it is also important to note that the priming effect was nearly equivalent for subjects in the different target encoding conditions. The difference between the number of control problem targets and the number of experimental problem targets that subjects generated during phase two was as great for subjects who had made only a letter search of the phase one targets as it was for subjects who had initially processed the targets more deeply by making a pleasantness judgment. This conclusion is supported statistically by the absence of a significant judgment type x problem type interaction.

The recognition data, on the other hand, revealed that depth of initial target processing had a substantial effect on subjects' ability to later recognize those targets. Target recognition was significantly higher for subjects who had focused on the meaning of the targets than it was for subjects who had focused on the targets' physical features. This result is consistent with a large body of research that relates processing depth to recognition and other forms of deliberate remembering (e.g., Craik & Tulving, 1975; Lockhart, Craik, & Jacoby, 1976). The most striking aspect of the recognition data is that subjects failed to recognize more than half of the shallowly encoded targets that they themselves had just generated as responses to the hypothesis generation problems.

When considered simultaneously, the hypothesis generation data and the recognition data together provide evidence for remembering without awareness in the type of hypothesis generation task that was used in this experiment. The fact that an earlier presentation of a target increased the chances of that target being later generated as a hypothesis suggests that subjects were remembering the earlier presentation. However, whereas subjects demonstrated memory for targets by generating them as hypotheses, subjects also failed to recognize many of those same targets when given a recognition test. Thus it appears that subjects may not have been aware of the influence of prior presentations on the hypotheses that they generated.

The distinction between the kind of remembering reflected in the hypothesis generation task and the aware form of remembering reflected in the recognition task is further supported by the different depth of processing effects observed in the two tasks. With respect to hypothesis generation, the size of the priming effect (that is, the difference between the number of target responses generated to the control and experimental problems) was as large for subjects who had judged the physical characteristics of targets as it was for subjects who had judged the targets' semantic characteristics. With respect to recognition of targets, subjects in the pleasantness judgment group clearly outperformed subjects in the letter judgment group. The finding that depth of processing affected recognition performance but not the presence or degree of priming lends support to the current view that the two forms of memory rely on different types of information or processes. The poor recognition of targets by subjects who revealed memory for those same targets in a hypothesis generation task also reinforces the notion that the act of deliberate retrieving is an important key to understanding the nature of the differences between the two forms of memory (Jacoby & Witherspoon, 1982).

One result observed in the hypothesis generation task data was unexpected. The unanticipated finding was that subjects in the semantic judgment group generated more targets overall. Although both judgment groups generated more experimental than control targets, and therefore exhibited priming, the semantic judgment group tended to generate more target items to both experimental and control problems. The difficulty in explaining this effect lies particularly with the control problems. There is no readily apparent explanation as to why subjects in the semantic processing condition generated a relatively large number of unprimed, normatively uncommon responses to the control problems. However, to the extent that the various categories and target items overlapped in meaning, the control targets may possibly have been elicited as associations. While this particular result is discrepant with the results of previous experiments in this area, it does not seem to us to detract from the priming effect of fundamental interest. That is, the direction and size of the difference between the number of experimental and control targets generated was essentially equivalent for the two judgment groups.

In summary, the overall pattern of Experiment 1 results seems to us to provide evidence for an unaware form of remembering in the context of a hypothesis generation task and offers further support for the dissociation between aware and unaware memory. Prior information that primed hypothesis generation was not necessarily recognized, and the level of processing of the prior information influenced recognition performance but not the degree of priming. This finding encouraged us to explore unaware memory in one other type of hypothesis generation task, one that was somewhat more representative of real world problems in which hypotheses must be generated.

EXPERIMENT 2

Method

Subjects. Subjects were 80 undergraduate student volunteers enrolled in psychology courses at George Mason University. All received a $5 payment and/or course credit for participating.

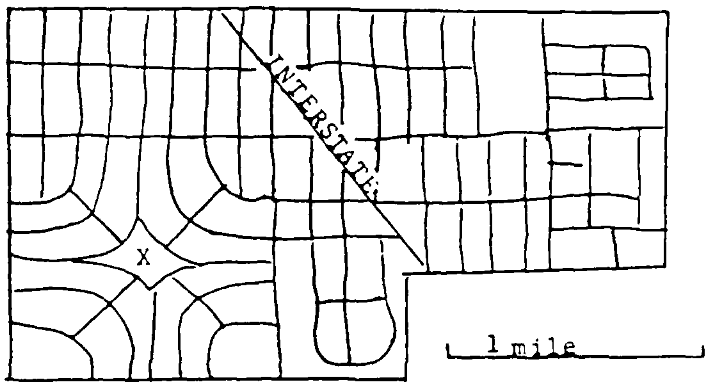
Design. The experiment was a 2 x 2 between-subjects design. Four groups were formed by the factorial combination of the type of information that was presented to subjects in an initial presentation phase (target items versus control items) and the type of processing subjects were instructed to engage in as each item was presented (pleasantness judgment versus letter search).

Materials. The critical materials were a set of hypothesis generation problems. Each problem consisted of a hand drawn map of an area marked "X" with its surrounding environment, three statements situated below the map describing area X, and a blank line for writing a hypothesis as to how area X might be used. The problems were originally developed and described by Gettys, Manning, Mehle, and Fisher (1980) and were modified slightly for use in this experiment. Each problem appeared on a separate page of a test booklet. A sample problem is shown in Figure 1.

Nine map problems of this type were pilot tested on 50 undergraduate students at local universities. Results of the pilot data were used to select target solutions or hypotheses that were relatively infrequent, yet subjectively judged to be reasonable. One problem was dropped, because a single hypothesis accounted for more than half of all the responses that pilot subjects had given to that problem. The remaining eight problems comprised the experimental set. For each problem, a particular hypothesis was selected and designated as the target solution. On the average, the target accounted for 14% of pilot subjects'responses to the problem (range was 6%-23%) and was never the most frequent response. The following eight items were the designated target solutions: amusement park, Indian reservation, school, prison, cemetery, airport, football stadium, nuclear waste dump.

Two presentation lists were constructed using the 8 target items (above) and 16 control items. The control items were selected such that, to the extent possible, they were high in frequency and concreteness ratings (Paivio, Yuille, & Madigan, 1968), unrelated to target items in meaning, and matched to target items in word length and item length. Moreover, none of the control items would have been suitable responses to any of the land use problems. An experimental list was formed consisting of the 8 target items intermixed with 8 of the control items, and a control list was formed using all 16 control items. Buffers appeared at the beginning and end of each list for a total list length of 22 items. A slide was made of each item printed in capital letters, and two random presentation orders of each list were developed. The remaining materials and equipment were the same as those used in Experiment 1.

Procedure. Subjects were tested in small groups, using a procedure similar to that of Experiment 1. Subjects initially viewed a 22-item word list on slides at an 8-second rate, and made a pleasantness or letter judgment as each item appeared. A 5-minute distractor task of solving arithmetic problems immediately followed. Subjects were then given a booklet of eight map problems. For each map problem, the task was to generate one hypothesis as to the possible use of land area X. Subjects were instructed to report the first hypothesis that came



ADDITIONAL INFORMATION ABOUT AREA X

1. Area shown is a suburb of a large city.
2. Area X should be centrally located.
3. Easy access to area X is desirable.

Area X serves a definite purpose. Think of 1 possible use for area X that is consistent with all the information provided (including the map) and write that use below.

Possible Use: ______________________________

Figure 1. Sample land use problem used in Experiment 2.

to mind and were not informed as to the connection between the initial list and the land use problems. The experimenter used a stopwatch to pace subjects through this task, allowing 45 seconds per problem.

For subjects in the control condition, the session was completed immediately following the hypothesis generation task. However, subjects in the experimental condition performed a recognition test after they had completed the land use problems. Experimental subjects were asked to look over the eight hypotheses they had just generated for the land use problems and to circle any that they recognized as having been presented at the beginning of the session.

Results

Hypothesis generation. The hypothesis generation data were analyzed to determine whether prior exposure to the designated target hypotheses increased the probability of those targets being generated as responses to the land use problems and, if so, whether such a priming effect was related to the target encoding condition. The dependent variable used in all ANOVA's was the number of target hypotheses that subjects generated to the land use problems. More than one approach, however, was used to count the number of targets generated by each subject. Initially, a strict scoring scheme was employed to determine whether or not a given response would be scored as a target. Under strict scoring, subjects were judged to have generated a target if the response exactly matched the wording of the earlier presented target item, and if the target response was generated to the intended problem, not to another problem for which it could also be a reasonable answer.

The data derived from strict scoring were submitted to a 2 x 2 x 2 between-subjects ANOVA. The independent variables were target exposure condition (experimental condition with prior exposure to targets versus control condition with no prior exposure), phase one encoding condition (pleasantness versus letter judgment), and phase one list presentation order (Order 1 versus Order 2). Since no effects involving order were of interest, nor were they significant, the data were collapsed across this variable and then submitted to a 2 (exposure condition) x 2 (judgment type) between-subjects ANOVA. The result of this analysis was that no effects were significant.

Two observations of the data, however, prompted us to reanalyze the data using more liberal scoring schemes. First, for any given problem, subjects frequently generated an item that was not a designated target for that problem, but that was a target for some other problem. This outcome had not been anticipated, since the pilot data indicated that each of the designated targets was a very unpopular answer (accounting for 0% to 4% of all responses) to all but one problem. However, since several generated hypotheses were target responses, it seemed appropriate to count them as such, even if they were not matched to the intended problem. The first liberal scoring procedure therefore ignored target-to-problem matching.

Second, several items were generated that were synonyms for, or that partially overlapped with, target items. A second scoring scheme was therefore devised in which these responses would be counted as targets. For example, graveyard (for cemetery); jail (for prison); stadium and football field (for football stadium); reservation (for Indian Reservation); and waste dump, waste dumping ground, and nuclear facility (for nuclear waste dump) were accepted as targets. Finally, a third liberal scoring scheme was used in which both unmatched targets and semantically similar responses were judged as targets.

Each of the three sets of data resulting from the three liberal scoring procedures was submitted to a 2 (study list presentation order) x 2 (exposure condition) x 2 (judgment type) between-subjects ANOVA. As expected, the results of all three analyses showed that there were no significant effects involving the random presentation orders. The data were collapsed across this variable and reanalyzed as a 2 (exposure condition) x 2 (judgment type) design.

With respect to the data derived from the first two liberal scoring schemes, the results of this analysis revealed no significant effects. That is, under scoring which allowed either unmatched targets or semantically similar responses to be counted as target answers, there were no significant differences in the total number of generated targets between subjects who were exposed to the targets and subjects who were not exposed to the targets during the presentation phase. Also, the type of judgment made on items during the presentation phase had no effect on target generation.

With respect to the data derived from the third and most liberal scoring scheme in which both unmatched targets and semantically similar respnses were counted as targets, the analysis resulted in a significant main effect of exposure condition, F (1,76) = 5.01, MSe = 1.32, $p<.03$. Relative to the control group that had no prior exposure to targets, subjects in the experimental group that had seen the targets during phase one generated significantly more of those targets as responses to the land use problems. The absence of a significant judgment x exposure condition interaction shows that this observed priming effect was as great for subjects who had judged the physical properties of targets as it was for subjects who had judged the semantic properties of targets in phase one. The group means and standard deviations are shown in Table 2.

Recognition. The target recognition data obtained from subjects in the experimental condition were analyzed with two goals in mind. One was to determine whether subjects were able to recognize any targets that they had generated during the hypothesis generation phase, and the other was to determine whether recognition performance was related to the way in which targets had been processed during phase one. In the recognition analyses, only experimental condition subjects who had generated at least one exactly-worded target as a land use hypothesis were considered. For each subject, the probability of recognizing a target, given that a target had been generated, was calculated and adjusted for that individual's false alarm rate. These conditional probabilities were then compared for subjects in the pleasantness (N=14) versus letter (N-12) judgment groups.

The mean probability of target recognition was found to be .88 for the semantic judgment group and .26 for the letter judgment group. This difference in recognition was statistically significant, t (24) = 5.17, $p<.01$. This result shows that subjects who had focused on target meaning during the target presentation phase were better able to recognize those targets later on than were subjects who had focused on the targets' physical characteristics. The data were also examined in terms of recognition failure rates. These data revealed that subjects who had made pleasantness judgments failed to recognize only 7% of the targets that they had just generated as hypotheses for the land use problems. In contrast, subjects who had made letter judgments on the targets failed to recognize 56% of the targets that they themselves had just generated as hypotheses.

Table 2

Number of Targets Generated as a Function of Type of Judgment and Condition (Experiment 2)

	Condition		
Judgment	Experimental	Control	Mean
Pleasantness			
Mean	2.05	1.15	1.60
SD	1.28	.81	
Letter			
Mean	1.50	1.25	1.38
SD	1.40	1.02	
Mean	1.78	1.20	

Discussion

The results of Experiment 2, in which land use problems were used to elicit hypotheses, provided only weak evidence for memory in the form of primed hypotheses. Some priming was revealed in the significant main effect of exposure condition on the number of targets that were generated as hypotheses. That is, experimental subjects who had been exposed to the targets in phase one generated more of those targets as responses to the land use problems than did control subjects. This difference was significant, however, only when very liberal criteria were applied to determine how many targets had been generated. Importantly, the size of the priming effect, however weak, was equivalent for subjects in the two judgment conditions. Since subjects in the letter judgment condition generated as many targets as subjects who had judged target meaning, it appears that even a relatively superficial type of encoding was sufficient to produce memory for those targets in the form of priming.

The target recognition data showed the typical effect of processing depth on recognition memory. Recognition of generated targets was significantly better for subjects who had previously judged the meaning of the targets compared to subjects who had made a letter search of the targets. The fact that the depth of processing variable had no effect on the degree of target priming during hypothesis generation (unaware memory) but a considerable effect on target recognition (aware memory) is consistent with previously reported findings in this area (e.g., Jacoby & Dallas, 1981; Tulving et al., 1982) and suggests that the two forms of memory are in some way functionally distinct. The recognition data also provide supporting evidence for unaware remembering in a task involving hypothesis generation. Although subjects demonstrated some degree of memory for previously presented targets by generating those targets as hypotheses, subjects failed to consistently remember those same targets in a recognition situation that required deliberate remembering.

GENERAL DISCUSSION

Research to date suggests that prior events can have a priming effect on a number of tasks, such as those that require subjects to reprocess the same or similar information, or to generate instances of conceptual categories. Such priming effects appear to represent an unaware form of remembering, distinct from deliberate retrieval in two respects. First, level of performance on aware and unaware memory tests of the same information is often independent. Second, certain variables such as level of processing have been shown to affect aware but not unaware memory performance.

The experiments reported here were conducted to extend the task conditions in which remembering without awareness might be investigated. Tasks involving hypothesis generation were chosen because of the importance of this skill to Army training in tasks requiring planning, problem solving, and decision making. The results of Experiment 1, which used a modified category instance production task, showed that normatively uncommon hypotheses could indeed be primed and that primed hypotheses were often not recognized. The hypothesis generation problems employed in Experiment 2 more nearly resembled a real world kind of task, but presented difficulties from the standpoint of experimental control. Even so, priming of land use hypotheses was evident when a liberal scoring scheme was applied to the data.

In Experiment 2, the depth of processing manipulation had a clearly different effect on the two tests of memory. The typical superiority of semantic processing was reflected in recognition test performance, whereas type of processing had no observable effect on memory in the form of target priming. The Experiment 1 results showed the typical depth effect in recognition, and also unexpectedly showed a depth of processing effect on target generation. However, even though subjects who had judged target meaning generated more of the primed targets as hypotheses (compared to subjects who had made letter judgments), they also generated more unprimed targets. The critical difference between the number of primed and unprimed targets generated by subjects in the two depth conditions was the same, leading us to tentatively conclude that the depth manipulation had little or no overall effect on priming. Additional research is needed to either confirm or revise this interpretation. If confirmed, this finding would lend considerable support to the possibility that unaware memory can play a role in hypothesis generation tasks and that the two forms of remembering in this situation rely on different processes or different kinds of information.

The recognition results of both experiments, although quite consistent with previously reported findings, might also be reinterpreted in view of the nature of the recognition test itself. The recognition failure of targets that had just been generated might be attributed to the change in context between initial target presentation and the recognition test phase. A standard recognition test consisting of a list of all of the targets with distractors is an alternative test form that will be used in our own future research.

When taken as a whole, the current results appear to us to coincide with the reported findings of researchers who have observed memory without awareness in other tasks (e.g., word identification, lexical decision, word completion, anagram solving, and category instance production). The unique contribution of our research is the information it provides concerning the type of task in which memory without awareness can be observed. The phenomenon is probably not limited to reprocessing tasks and generation tasks in which there is a partial or complete overlap of perceptual or sensory information between the item's first occurrence and the priming test situation. Also to be considered now are tasks requiring a memory search and generation of a hypothesis that is consistent with a given set of information, even when the information does not overlap in a perceptual or sensory way with the hypothesis.

It should also be noted that a priming effect was evident despite the fact that the designated target hypotheses in both experiments were quite uncommon. Jacoby and Witherspoon (1982) have similarly demonstrated remembering without awareness by priming relatively uncommon interpretations of ambiguous stimuli. These researchers biased the less common meaning of homophones during a presentation phase, then subsequently asked subjects for homophone spellings. Memory revealed by spelling performance was found to be independent of recognition memory. Their data and the present results encourage exploration of both the variety of circumstances in which unaware memory influences task performance and the extent to which performance can be so influenced.

A practical implication of the research results with respect to hypothesis generation is that unaware memory may play an important role in generating the alternative that is eventually selected in a decision making or problem solving situation. When given a problem, people tend not to generate a wide variety of alternative solutions (see Gettys, 1983 for a review). Several decision making researchers have reported that people claim to run out of ideas after generating only a few alternatives (Mehle, 1982; Gettys, Manning, & Casey, 1981; Pitz, Sachs, & Heerboth, 1980). Thus, the first few hypotheses or alternative solutions to be generated can be very important because of their relatively high probability of being chosen and acted upon. Moreover, if the retrieval of such alternatives from memory is biased by a prior event that cannot be consciously remembered, decision makers or problem solvers may be hampered in their ability to monitor and evaluate the quality of the alternatives they are generating.

Although the focus of this research has been to identify conditions in which we observe unaware memory in the form of priming effects, it is also instructive to directly compare situations in which priming is and is not expected to occur. For example, prior exposure to information that could be used to solve insight problems does not necessarily facilitate or influence problem solution. Research reported by Perfetto, Bransford, and Franks (1983) showed that unless subjects were explicitly informed of the relationship between the presentation of obviously relevant clues and a later problem solving task, they did not benefit from clue information.

One difference between the Perfetto et al. (1983) task condition and the current tasks is that the prior information was an obvious though indirect clue, rather than a direct answer. Perhaps a more important distinction lies in the nature of the problems themselves. Insight problems typically have one single acceptable solution, whereas the types of generation problems used in the present research may be solved with any of several appropriate solutions. A systematic analysis of the differences among tasks in which we do and do not observe priming (in particular, priming effects without awareness) may be a fruitful research approach for further understanding the nature of unaware remembering.

Finally, although the current research supports and helps to define the distinction between aware and unaware remembering, it does not support any one particular theoretical view of systems of memory organization and memory processing. Future research directions will likely be toward seeking to explain differences between aware and unaware forms of remembering in the context of existing or perhaps new theoretical approaches to memory. Current efforts in this direction (see Tulving, 1985) raise the possibility that neither episodic nor semantic memory systems can account for the type of unaware learning that is reflected in tasks such as those being investigated here. The exact nature of the underlying memory system is likely to be the object of a great deal of future research.

REFERENCES

Craik, F. I. M., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. Journal of Experimental Psychology: General, 104, 268-294.

ich, E. (1984). Memory for unattended events: Remembering with and without awareness. Memory & Cognition, 12, 105-111.

Gettys, C. F. (1983). Research and theory on predecision processes (Report No. 11-30-83). Norman, OK: University of Oklahoma Decision Processes Laboratory.

Gettys, C., Manning, C., & Casey, J. (1981). An evaluation of human act generation performance (Report No. 15-8-81). Norman, OK: University of Oklahoma Decision Processes Laboratory.

Gettys, C. F., Manning, C., Mehle, T., & Fisher, S. (1980). Hypothesis generation: A final report of three years of research (Report No. TR 15-10-80). Arlington, VA: Office of Naval Research.

Graf, P. & Schacter, D. L. (1985). Implicit and explicit memory for new associations in normal and amnesic subjects. Journal of Experimental Psychology: Learning, Memory, and Cognition, 11, 501-518.

Graf, P., Shimamura, A. P., & Squire, L. R. (1985). Priming across category levels: Extending the domain of preserved function in amnesia. Journal of Experimental Psychology: Learning, Memory, and Cognition, 11, 386-396.

Jacoby, L. L., & Dallas, M. (1981). On the relationship between autobiographical memory and perceptual learning. Journal of Experimental Psychology: General, 110, 306-340.

Jacoby, L. L., & Witherspoon, D. (1982). Remembering without awareness. Canadian Journal of Psychology, 36, 300-324.

Kihlstrom, J. F. (1980). Posthypnotic amnesia for recently learned material: Interactions with "episodic" and "semantic" memory. Cognitive Psychology, 12, 227-251.

Kolers, P. A. (1976). Reading a year later. Journal of Experimental Psychology: Human Learning and Memory, 2, 554-565.

Lockhart, R. S., Craik, F. I. M., & Jacoby, L. L. (1976). Depth of processing, recognition and recall: Some aspects of a general memory system. In J. Brown (Ed.), Recall and Recognition. London: Wiley.

Masson, M. E. J. (1984). Memory for the surface structure of sentences: Remembering with and without awareness. Journal of Verbal Learning and Verbal Behavior, 23, 579-592.

Mehle, T. (1982). Hypothesis generation in an automobile malfunction inference task. Acta Psychologica, 52, 87-106.

Paivio, A., Yuille, J. C., & Madigan, S. (1968). Concreteness, imagery, and meaningfulness values for 925 nouns. Journal of Experimental Psychology Mongraph Supplement, 76, (1, Pt. 2).

Perfetto, G. A., Bransford, J. D., & Franks, J. J. (1983). Constraints on access in a problem solving context. Memory & Cognition, 11, 24-31.

Pitz, G. F., Sachs, N., & Heerboth, J. (1980). Procedures for eliciting choices in the analysis of individual decisions. Organizational Behavior and Human Performance, 26, 396-408.

Radtke, R. C., & Englert, J. A. (1983, May). Memory without awareness in problem solving tasks. Paper presented at the meeting of the Midwestern Psychological Association, Chicago.

Scarborough, D., Cortese, C., & Scarborough, H. S. (1977). Frequency and repetition effects in lexical memory. Journal of Experimental Psychology: Human Perception and Performance, 3, 1-17.

Tulving, E. (1972). Episodic and semantic memory. In E. Tulving and W. Donaldson (Eds.), Organization of memory. New York: Academic Press.

Tulving, E. (1983). Elements of episodic memory. New York: Oxford University Press.

Tulving, E. (1985). How many memory systems are there? American Psychologist, 40, 385-398.

Tulving, E., Schacter, D. L., & Stark, H. A. (1982). Priming effects in word-fragment completion are independent of recognition memory. Journal of Experimental Psychology: Learning, Memory, and Cognition, 8, 336-342.

Warrington, E. K., & Weiskrantz, L. (1970). Amnesic syndrome: Consolidation or retrieval? Nature, 228, 628-630.