1 of/de 1

PM-1 3 1/2" x 4" PHOTOGRAPHIC MICROCOPY TARGET
NBS 1010a ANSI/ISO #2 EQUIVALENT

PRECISIONSM RESOLUTION TARGETS

PIONEERS IN METHYLENE BLUE TESTING SINCE 1974

MicroD
TECHNOLOGY

1880 COUNTY ROAD 3, BURNINGWELL, WV 26037 USA
TEL: 812 488 7887 FAX: 812 488 7887 TLA: 8124887887

National Library
of Canada

Acquisitions and
Bibliographic Services Branch

395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale
du Canada

Direction des acquisitions et
des services bibliographiques

395 rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file - Votre référence

Our file - Notre référence

NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.

INVESTIGATING CONCEPTUALLY-DRIVEN MENTAL OPERATIONS WITH
CONCEPTUALLY-DRIVEN INDIRECT MEASURES OF MEMORY

by

RODERICK D. GRANT, B.A.

A thesis submitted to
the Faculty of Graduate Studies and Research
in partial fulfilment of
the requirements for the degree of

Master of Arts

Department of Psychology

Carleton University

Ottawa, Ontario

April 22, 1992

© copyright

1992, Roderick D. Grant

National Library
of Canada

Acquisitions and
Bibliographic Services Branch

395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale
du Canada

Direction des acquisitions et
des services bibliographiques

395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Author: Author's name

Editor: Author's name

The author has granted an irrevocable non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of his/her thesis by any means and in any form or format, making this thesis available to interested persons.

L'auteur a accordé une licence irrévocable et non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de sa thèse de quelque manière et sous quelque forme que ce soit pour mettre des exemplaires de cette thèse à la disposition des personnes intéressées.

The author retains ownership of the copyright in his/her thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without his/her permission.

L'auteur conserve la propriété du droit d'auteur qui protège sa thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

ISBN 0-315-79824-6

Canada

The undersigned recommend to the Faculty of Graduate Studies
and Research acceptance of the thesis

"INVESTIGATING CONCEPTUALLY-DRIVEN MENTAL OPERATIONS WITH
CONCEPTUALLY-DRIVEN INDIRECT MEASURES OF MEMORY"

submitted by Roderick D. Grant, B.A.
in partial fulfilment of the requirements for
the degree of Master of Arts

Thesis Supervisor

Chair, Department of Psychology

Carleton University

May 12, 1992

Abstract

In the present study two new conceptually-driven indirect measures of memory were used to examine Graf and Schacter's (1989) semantic unitization hypothesis (i.e., that generating a sentence with the members of a word pair induces the processing of a meaningful association). In the study phase, subjects were presented with a series of related and unrelated word pairs. Subjects were asked to generate a meaningful sentence with the two words (sentence generation) or they were asked to compare the number of vowels and consonants in the two words (letter comparison condition). In the test phase, the studied word pairs were re-presented along with some new related and unrelated word pairs. Subjects were asked to judge, as quickly as possible, whether the words were related or unrelated (semantic decision test) or they were asked to judge, as quickly as possible, how related the words were (subjective evaluation test). In general, the results with the related word pairs supported the semantic unitization hypothesis while the results with the unrelated word pairs did not. The implications of these results for the semantic unitization hypothesis, and for future studies of human learning and memory, are discussed.

Acknowledgements

I would like to thank my advisor, Mary Gick, for the tremendous amount of support and encouragement that she has given me throughout this project. I would also like to thank Kate for her love and understanding.

Table of Contents

	<u>Page</u>
Title Page	i
Acceptance Sheet	ii
Abstract	iii
Acknowledgements	iv
Table of Contents	v
List of Tables	viii
List of Figures	ix
 Introduction	 1
Indirect Measures of Memory	1
Special Populations	1
Normal Memory Processes	3
Dissociations with the levels of processing paradigm	3
Dissociations with the generation effect paradigm	5
A Theoretical Account of the Dissociations Between Direct and Indirect Measures of Memory	7
A Test of the Processing Account	9
Conceptually-driven indirect measures of memory	12
Rationale for the Present Experiment	16
Study Conditions	20
Semantic Decision Test	20
Subjective Evaluation Test	22
Word Fragment Completion Test	24
 Method	 25
Subjects	25
Design	25
Materials	25
Procedure	26

Table of Contents (continued)

	<u>Page</u>
Results	30
Study Phase	30
Response latencies	30
Word Fragment Completion Test	35
Target words produced	35
Semantic Decision Test	36
Errors	36
Response latencies	42
Subjective Evaluation Test	46
Errors	46
Response latencies	50
Subjective ratings	54
 Discussion	 58
Summary of Results	58
Response latencies in the study phase	58
Target words produced on the word fragment completion test	59
Errors on the semantic decision test and the subjective evaluation test	59
Response latencies on the semantic decision test and the subjective evaluation test	60
Ratings on the subjective evaluation test	61
Implications	61
Implications for the semantic unitization hypothesis	61
Implications for the relationship between encoding time and memory performance	64
Implications for the processing account and the memory systems account	65
Implications for future studies of human learning and memory	66

Table of Contents (continued)

	Page
Future Research	67
Investigating unrelated word pairs in an unmixed list	67
Investigating different retention intervals	68
Conclusions	69
References	70
Appendix A: Format 1	80
Appendix B: Format 2	81
Appendix C: Practice Items and Fillers	82

List of Tables

<u>Table</u>	<u>Description</u>	<u>Page</u>
1	Summary of Predictions	31
2	Analysis of Variance Summary Table for Response Latencies (in ms) in the Study Phase as a Function of Test Condition, Study Condition and Word Pair Relation.	34
3	Analysis of Variance Summary Table for the Proportion of Targets on the Word Fragment Completion Test as a Function of Study Condition, Word Pair Relation and Item Status.	38
4	Analysis of Variance Summary Table for the Proportion of Errors on the Semantic Decision Test as a Function of Study Condition, Word Pair Relation and Item Status.	41
5	Analysis of Variance Summary Table for Response Latencies (in ms) on the Semantic Decision Test as a Function of Study Condition, Word Pair Relation and Item Status.	45
6	Analysis of Variance Summary Table for the Proportion of Errors on the Subjective Evaluation Test as a Function of Study Condition, Word Pair Relation and Item Status.	49
7	Analysis of Variance Summary Table for Response Latencies (in ms) on the Subjective Evaluation Test as a Function of Study Condition, Word Pair Relation and Item Status.	53
8	Analysis of Variance Summary Table for the Ratings on the Subjective Evaluation Test as a Function of Study Condition, Word Pair Relation and Item Status.	57

List of Figures

Figure	Description	Page
1	Mean Response Latencies (in ms) in the Study Phase as a Function of Study Condition, Test Condition and Word Pair Relation	33
2	Mean Proportion of Targets on the Word Fragment Completion Test as a Function of Word Pair Relation, Study Condition and Item Status	37
3	Mean Proportion of Errors on the Semantic Decision Test as a Function of Word Pair Relation, Study Condition and Item Status	40
4	Mean Response Latencies (in ms) on the Semantic Decision Test as a Function of Word Pair Relation, Study Condition and Item Status	44
5	Mean Proportion of Errors on the Subjective Evaluation Test as a Function of Word Pair Relation, Study Condition and Item Status	48
6	Mean Response Latencies (in ms) on the Subjective Evaluation Test as a Function of Word Pair Relation, Study Condition and Item Status	52
7	Mean Ratings on the Subjective Evaluation Test as a Function of Word Pair Relation, Study Condition and Item Status.	56

Studies of memory have traditionally relied upon tests which require the deliberate retrieval of information from a prior learning episode (e.g., free recall, cued recall and recognition). In recent years, however, a growing number of studies have employed the use of tests which do not make any reference to the past (see Richardson-Klavehn & Bjork, 1988 and Schacter, 1987 for reviews). Memory is revealed on these tests by a change in performance that is attributable to the recent episode. For example, when subjects are required to complete a series of word fragments (e.g., _H_O_EM), recent exposure to the solution words in the study phase of an experiment (e.g., THEOREM) increases the probability of completing the word fragments (e.g., Tulving, Schacter & Stark, 1982). Recent encounters also facilitate the identification of briefly presented words (e.g., Jacoby & Dallas, 1981), the reading of transformed script (e.g., Kolars, 1975) and the ability to make lexical decisions (e.g., Scarborough, Cortese & Scarborough, 1977). Since these tests make no reference to an event in the subject's past they have been called indirect (Johnson & Hasher, 1987), implicit (Graf & Schacter, 1985) or incidental (Jacoby, 1984) measures of memory. In contrast, the traditional measures of memory have been called direct (Johnson & Hasher, 1987), explicit (Graf & Schacter, 1985) or intentional (Jacoby, 1984) measures of memory since they require the subject to retrieve information from a prior learning episode. The purpose of the present study was to examine certain theoretical assumptions about memory processes, using indirect measures of memory. In order to understand the rationale for the present work, a review of the relevant studies in implicit memory will be presented, followed by a discussion of a theoretical framework used to account for the research.

Indirect Measures of Memory

Special Populations

The indirect measures of memory were first used in studies with amnesic patients

(see Shimamura, 1986 and Squire & Cohen, 1984 for reviews). In a series of experiments, Warrington and Weiskrantz (1968, 1970, 1974, 1978; Weiskrantz & Warrington, 1970) compared a group of amnesic patients with a group of matched controls on a number of direct and indirect measures of memory. As expected, they found the amnesics to be impaired on the direct measures of recall and recognition when compared to the matched controls. However, they found the amnesics to be unimpaired on the indirect measures of word stem completion and word fragment completion. Other researchers have provided similar results. Amnesics show the effects of recent exposure when spelling homophones (e.g., Jacoby & Witherspoon, 1982), completing picture fragments (e.g., Milner, Corkin & Teuber, 1968), reading transformed text (e.g., Cohen & Squire, 1980), producing category exemplars (e.g., Graf, Shimamura & Squire, 1985), making lexical decisions (e.g., Moscovitch, 1984), identifying briefly presented words (e.g., Cermak, Talbot, Chandler & Wolbarst, 1985), assembling jigsaw puzzles (e.g., Brooks & Baddeley, 1976) and solving the Tower of Hanoi problem (e.g., Cohen, Eichenbaum, Deacedo & Corkin, 1985).

In recent years, the indirect measures of memory have also been used to investigate the effects of aging on memory. It is well documented that elderly adults (i.e., adults over 65 years of age) perform worse than young adults (i.e., adults between 18 and 25 years of age) on the direct measures of recall and recognition (see Burke & Light, 1981 and Craik, 1977 for reviews). Recent evidence, however, suggests that the elderly are unimpaired on the indirect measures of memory (see Graf, 1990 and Howard, 1988 for reviews). Elderly adults and young adults show the same effects of recent exposure when spelling homophones (e.g., Howard, 1988), naming pictures (e.g., Mitchell, 1989), completing word fragments (e.g., Light, Singh & Capps, 1986), finishing word stems (e.g., Howard, 1988), making lexical decisions (e.g., Moscovitch, 1982) and identifying degraded words

(e.g., Light & Singh, 1987). More recently, a number of investigators have also examined the performance of young children (i.e., children under 6 years of age) on the indirect measures of memory. While performance on the direct measures is known to develop in the early years of life (Bjorklund & Bjorklund, 1985), these studies have found no evidence for a similar trend on the indirect measures. The ability to identify fragmented pictures (e.g., Parkin & Streete, 1988) and the ability to produce category exemplars (e.g., Greenbaum & Graf, 1989) has not varied with age.

Normal Memory Processes

The indirect measures have not only been used to examine the memory performance of special populations (e.g., amnesic patients, elderly adults and young children). They have also been used to examine the memory performance of normally functioning young adults (i.e., adults between 18 and 25 years of age). In some of these studies, the direct measures of memory have been compared with the indirect measures of memory in their response to traditional encoding manipulations. The results of these studies have presented a significant challenge to memory theorists since some of the encoding manipulations have had different effects on the direct and indirect measures of memory. These "dissociations" in performance are probably best exemplified in studies which have employed the "levels of processing" paradigm and the "generation effect" paradigm.

Dissociations with the levels of processing paradigm. One very robust finding in the study of human learning and memory is that the recall and recognition of semantically processed information is better than the recall and recognition of structurally processed information (see Lockhart & Craik, 1990 for a review). One early demonstration of this phenomenon was provided by Craik and Tulving (1975). They presented subjects with a series of common nouns and had them make judgements about their semantic

characteristics (e.g., IS THE WORD AN ANIMAL NAME?) or their structural characteristics (e.g., IS THE WORD PRINTED IN CAPITAL LETTERS?). This study presentation was then followed by a recognition test. A series of old words (i.e., studied words) was intermixed with a series of new words (i.e., nonstudied words) and the subject's task was to indicate which words were in the study phase. Using this measure of memory, Craik and Tulving (1975) found that subjects were able to recognize more semantically encoded words than structurally encoded words.

In several further experiments, Craik and Tulving (1975) found that semantic encoding led to better performance than structural encoding, under both intentional and incidental learning conditions, on all of the direct measures of memory (i.e., free recall, cued recall and recognition). Over fifteen years of subsequent research in the area of cognitive psychology (see Lockhart & Craik, 1990 for a review), and in such diverse areas as psychopharmacology (e.g., Malone, Kershner & Siegel, 1988), psychophysiology (e.g., Sanquist, Rohrbaugh, Syndulko & Lindsley, 1980), biological psychology (e.g., Leong, Wong, Wong & Hiscock, 1985), abnormal psychology (e.g., Boyd & Ellis, 1986), developmental psychology (e.g., Owings & Baumeister, 1979) and educational psychology (e.g., Watkins, 1983), has generally confirmed that the recall and recognition of semantically processed information is better than the recall and recognition of structurally processed information (see Morris, Bransford & Franks, 1977 and Stein, 1978 for examples of cases in which semantic encoding does not lead to better performance than structural encoding on a direct measure of memory).

In several recent experiments, however, the advantage for semantically encoded information has not been revealed on the indirect measures of memory. In one of these experiments, Jacoby and Dallas (1981) presented subjects with a series of words and had them make semantic judgements (e.g., IS THE CENTER OF THE NERVOUS SYSTEM?)

or structural judgements (e.g., CONTAINS THE LETTER L?). This study phase was then followed by a recognition test and a perceptual identification test. On the latter test, a series of old and new words were presented for brief durations (e.g., 35 ms) and the subject's task was to make an identification. On the recognition test, Jacoby and Dallas (1981) replicated the results of Craik and Tulving (1975). They found that subjects were able to recognize more semantically encoded words than structurally encoded words. On the perceptual identification test, however, Jacoby and Dallas (1981) did not find an advantage for the semantically processed words. The facilitation in identification performance (i.e., for old words over new words) was equivalent for the semantically encoded words and the structurally encoded words. Other investigators have found similar dissociations, between the direct and indirect measures of memory, when words are studied in a levels of processing paradigm. Graf, Mandler and Haden (1982) found this dissociation using free recall and word stem completion, Graf and Mandler (1984) found this dissociation using cued recall and word stem completion and Light and Singh (1987) found this dissociation using free recall and perceptual identification.¹

Dissociations with the generation effect paradigm. Another very robust finding in the study of human learning and memory is that the recall and recognition of internally generated information is better than the recall and recognition of externally presented information. This phenomenon, commonly referred to as the generation effect, was first outlined in a series of experiments by Slamecka and Graf (1978). In one of these experiments, subjects were presented with a series of semantically related word pairs.

¹ Since a complete review of studies involving the levels of processing paradigm is beyond the scope of this paper the interested reader is referred to Craik and Lockhart (1972), Cermak and Craik (1979), and Lockhart and Craik (1990).

One-half of these word pairs appeared in an incomplete format (e.g., RAPID-F___) and the subject was required to read the cue and generate the target. The remaining half of the word pairs were presented in a complete format (e.g., RAPID-FAST) and the subject was required to read the cue and the target. This study presentation was then followed by a recognition test. Using this measure of memory, Slamecka and Graf (1978) found that subjects were able to recognize more self-generated targets than externally-presented targets.

Subsequent research has indicated that the generation advantage can be obtained with a number of different generation tasks. An effect has been found when the generation task is defined by an associate rule (e.g., LAMP-L___ vs. LAMP-LIGHT; Rabinowitz & Craik, 1986), a category rule (e.g., RUBY-D___ vs. RUBY-DIAMOND; Slamecka & Graf, 1978) an antonym rule (e.g., LONG-S___ vs. LONG-SHORT; McElroy & Slamecka, 1982) and a rhyme rule (e.g., SAVE-C___ vs. SAVE-CAVE; Payne, Neely & Burns, 1986). The phenomenon has also been revealed when subjects are required to complete word fragments (e.g., AL_OHO_ vs. ALCOHOL; Glisky & Rabinowitz, 1985), sentence frames (e.g., DIAMONDS ARE A GIRL'S BEST ___ vs. DIAMONDS ARE A GIRL'S BEST FRIEND; Anderson, Goldberg & Hidde, 1971) and simple multiplication problems (e.g., $7 \times 4 = ___$ vs. $7 \times 4 = 28$; Gardiner & Rowley, 1984). In fact, a generation effect has even been found when subjects are required to solve anagrams (e.g., FSCAR vs. SCARF; Jacoby & Dallas, 1981) and flip inverted letters (e.g., EHAVEN vs. HEAVEN, Nairne, Pusey & Widner, 1985). Finally, it should be noted that the phenomenon has been revealed on all of the direct measures of memory. An advantage for generated items has been found using tests of free recall (e.g., Slamecka & Graf, 1978), cued recall (e.g., Jacoby, 1978) and recognition (e.g., McElroy & Slamecka, 1982) (see Begg & Snider, 1987; McElroy & Slamecka, 1982; and Slamecka & Katsaiti, 1987; however, for

examples of cases in which generation does not lead to better performance than reading on a direct measure of memory).

In several recent experiments, however, the generation effect has not been revealed on the indirect measures of memory. In the study phase of one of these experiments, Jacoby (1983a) had subjects generate target words in the context of an antonym cue (e.g., HOT-???) or read target words in the context of a neutral stimulus (e.g., XXX- COLD). This study phase was then followed by a test of recognition memory and a test of perceptual identification. Like Slamecka and Graf (1978), Jacoby (1983a) found a generation effect on the recognition memory test. Subjects recognized more self-generated targets than externally-presented targets. On the perceptual identification test, however, Jacoby (1983a) found a reverse generation effect. The facilitation in identification performance (i.e., for old words over new words) was greater for externally-presented targets than self-generated targets. Other investigators have found similar dissociations between the direct and indirect measures of memory when words are studied in a generation effect paradigm. For example, Winnick and Daniel (1970) and Clarke and Morton (1983) found this dissociation using free recall and visual duration threshold.²

A Theoretical Account of the Dissociations Between

Direct and Indirect Measures of Memory

One theoretical account of the dissociations, often referred to as the processing account (e.g., Schacter, 1987) or the non-abstractionist account of memory (e.g.,

² Since a complete review of studies involving the generation effect paradigm is beyond the scope of this paper the interested reader is referred to Slamecka and Graf (1978), Slamecka and Katsaiti (1987) and Begg and Snider (1987).

Richardson-Klavehn & Bjork, 1988), explains the dissociations by examining the interaction between study processes and test processes (e.g., Blaxton, 1989; Jacoby, 1983a, 1983b; Kolers & Roediger, 1984; Roediger & Blaxton, 1987a, 1987b; Roediger, Weldon & Challis, 1989; Srinivas & Roediger, 1990; Weldon & Roediger, 1987). This processing theory is based on the principle of encoding specificity (Tulving & Thomson, 1973), or the idea of transfer-appropriate processing (Morris, Bransford & Franks, 1977), which suggests that memory in general is defined by the extent to which the processing operations performed at study overlap with the processing operations required at test. This type of account was first used by Jacoby (1983a) to explain the dissociation he observed with the generation effect paradigm. Specifically, he explained his results by considering the processes involved in the study tasks (i.e., generating and reading) and the processes involved in the memory tests (i.e., recognition and perceptual identification). With regard to the study tasks, Jacoby (1983a) argued that generating a word in the context of an antonym cue (e.g., HOT-???) was primarily a conceptually-driven (top-down) process since the subject needed to use the antonym cue (i.e., HOT) to retrieve the target word (i.e., COLD) from memory in order to say it aloud. On the other hand, he argued that reading a word in the context of a neutral stimulus (e.g., XXX-COLD) was primarily a data-driven (bottom-up) process since the subject needed to process the physical features of the target word (i.e., COLD) in order to say it aloud. With regard to the memory tests, Jacoby (1983a) argued that recognition was primarily a conceptually-driven measure, because the subject needed to reconstruct the study episode in order to make a recognition decision. On the other hand, he argued that perceptual identification was primarily a data-driven measure, because the subject needed to rely on the analysis of physical features in order to identify the briefly presented words. On the basis of this analysis, Jacoby (1983a) argued that a generation effect was observed on the

conceptually-driven recognition test because subjects did more conceptually-driven processing in the generation task than in the reading task, while a reverse generation effect was observed on the data-driven perceptual identification test because subjects did more data-driven processing in the reading task than in the generation task.

A Test of the Processing Account

Following up on the argument made by Jacoby (1983a), Blaxton (1989) noticed that the experiments which have revealed dissociations between direct and indirect measures of memory have confounded the type of processing (i.e., conceptually-driven or data-driven) with the type of test (i.e., direct or indirect). For example, in the experiments showing a dissociation with the levels of processing paradigm (e.g., Graf et al., 1982; Graf & Mandler, 1984; Jacoby & Dallas, 1981; Light & Singh, 1987), one conceptually-driven direct measure of memory (e.g., free recall, cued recall or recognition) was compared with one data-driven indirect measure of memory (e.g., perceptual identification or word stem completion). Similarly, in the experiments showing a dissociation with the generation effect paradigm (e.g., Clarke & Morton, 1983; Jacoby, 1983a; Winnick & Daniel, 1970), one conceptually-driven direct measure of memory (e.g., free recall or recognition) was compared with one data-driven indirect measure of memory (e.g., perceptual identification or visual duration threshold). Thus, from these experiments, it is unclear whether the dissociations are the result of variations in the type of processing or the type of test.

In order to correct this confound, Blaxton (1989) conducted an experiment with the generation effect paradigm and four measures of memory: a conceptually-driven direct measure, a conceptually-driven indirect measure, a data-driven direct measure, and a data-driven indirect measure. Specifically, in the study phase of this experiment, she had subjects generate target words in the context of a semantic cue (e.g., HAWK-????) or

read target words in the context of a neutral stimulus (e.g., XXX-EAGLE). In the test phase, one group of subjects was given a semantic cued recall test (i.e., a conceptually-driven direct measure) in which they were given a series of semantic cues (e.g., FALCON) to help them retrieve the target words from the study phase (e.g., EAGLE). A second group of subjects was given a general knowledge test (i.e., a conceptually-driven indirect measure) in which they were given a series of questions (e.g., WHAT WAS THE NAME OF ARMSTRONG AND ALDRIN'S LUNAR MODULE?) with answers which corresponded to the target words from the study phase (e.g., EAGLE). A third group of subjects was given a graphemic cued recall test (i.e., a data-driven direct measure) in which they were given a series of graphemic cues (e.g., EAGER) to help them retrieve the target words from the study phase (e.g., EAGLE). A fourth group of subjects was given a word fragment completion test (i.e., a data-driven indirect measure) in which they were given a series of word fragments (E_G_E) with solutions which corresponded to the target words from the study phase (e.g., EAGLE).

Since the processing account focuses on the match in processing operations performed at study and test, it leads to the prediction that generation (i.e., a conceptually-driven study process) will produce a superior level of performance than reading (i.e., a data-driven study process) on the two conceptually-driven measures of memory (i.e., semantic cued recall and general knowledge). On the other hand, the processing account leads to the prediction that reading (i.e., a data-driven study process) will produce a superior level of performance than generation (i.e., a conceptually-driven study process) on the two data-driven measures of memory (e.g., graphemic cued recall and word fragment completion). Blaxton (1989) obtained this pattern of results. Specifically, she found a generation effect on the semantic cued recall test (i.e., generation led to a higher level of recall than reading) and the general knowledge test (i.e., generation led to

a higher probability of answering a general knowledge question with a study word than did reading) while she found a reverse generation effect on the graphemic cued recall test (i.e., reading led to a higher level of recall than generation) and the word fragment completion test (i.e., reading led to a higher probability of completing a fragment with a study word than did generation). Since this dissociation in memory performance occurred between the conceptually-driven measures (i.e., semantic cued recall and general knowledge) and the data-driven measures (i.e., graphemic cued recall and word fragment completion), and not between the direct measures (i.e., semantic cued recall and graphemic cued recall) and the indirect measures (i.e., general knowledge and word fragment completion), the results suggest that dissociations are the result of variations in the type of processing rather than variations in the type of test.

While the experiment by Blaxton (1989) provided support for the processing account, it did not provide support for another theoretical account of memory which has been used to explain the dissociations between direct and indirect measures of memory. This theory, often referred to as the memory systems account (e.g., Schacter, 1987) or the abstractionist account (e.g., Richardson-Klavehn & Bjork, 1988), suggests that memory consists of a set of memory systems, and dissociations occur because the direct measures of memory rely on a different memory representation or memory system than the indirect measures of memory. For example, Tulving (1985, 1987) has suggested that the dissociations occur because the direct measures of memory rely on an episodic memory system (i.e., a system which contains the memory for personal experiences, episodes or events) and the indirect measures of memory rely on a semantic memory system (i.e., a system which contains the memory for words, concepts, rules and other abstract ideas). In a similar manner, Squire (1986, 1987) has suggested that the dissociations occur because the direct measures of memory rely on a declarative memory

system (i.e., a system which is involved in the formation of new representations or data structures) and the indirect measures of memory rely on a procedural memory system (i.e., a system which records information about processing operations and their modifications). With these assumptions, the memory systems account would not predict dissociations between the direct measures since they rely on the same memory system (i.e., the episodic or declarative memory system) and it would not predict dissociations between the indirect measures since they rely on the same memory system (i.e., the semantic or procedural memory system). As a result, the memory systems account can not explain the dissociations that Blaxton (1989) found between the direct measures of memory (i.e., between the semantic cued recall test and the graphemic cued recall test), or the dissociations that Blaxton (1989) found between the indirect measures of memory (i.e., between the general knowledge test and the word fragment completion test).

Conceptually-driven indirect measures of memory. While previous experiments with the levels of processing paradigm (e.g., Graf et al., 1982, Graf & Mandler, 1984, Jacoby & Dallas, 1981; Light & Singh, 1987) and the generation effect paradigm (e.g., Clarke & Morton, 1983; Jacoby, 1983a; Winnick & Daniel, 1970) have indicated that the advantage of a conceptually-driven study process (e.g., a semantic orienting task or a generation task) over a data-driven study process (e.g., a structural orienting task or a reading task) could not be revealed on an indirect measure of memory, the experiment by Blaxton (1989) indicated that the advantage could be revealed on an indirect measure of memory if it was conceptually-driven (e.g., the general knowledge test). Since the present study directly addresses the processes involved in conceptually-driven indirect measures of memory, a brief review of these tests will be presented. In one experiment, Hamman (1990) examined the effect of semantic encoding on the general knowledge test. Subjects in this study were presented with a series of words (e.g., HEMLOCK) under

semantic encoding conditions (e.g., RATE THE WORD ON A 5-POINT LIKING SCALE) and structural encoding conditions (e.g., DOES THE WORD SHARE ANY VOWELS WITH THE PREVIOUS WORD). Then subjects were required to answer a series of general knowledge questions (e.g., WHAT DID SOCRATES DRINK AT HIS EXECUTION?). Using this conceptually-driven indirect measure of memory, Hamman (1990) found an advantage for semantic encoding. Studying a word under semantic encoding conditions led to a higher probability of producing that word as a response to a general knowledge question, than did studying a word under structural encoding conditions.

The advantage of conceptually-driven study processes over data-driven study processes has also been revealed on another conceptually-driven indirect measure of memory. A number of recent studies have used a category association test that was first developed by Graf et al. (1985). In one of these experiments, Srinivas and Roediger (1990) examined the effect which generation would have on this test. In the study phase, subjects were presented with a series of sentences. One-half of these were presented in an incomplete format (e.g., LIGHTNING IS ASSOCIATED WITH _____) and the subjects were required to generate the final word and then read the sentence aloud. The remaining sentences were presented in a complete format (e.g., LIGHTNING IS ASSOCIATED WITH THUNDER) and the subjects were required to read them aloud. Memory for the target words (e.g., THUNDER) was then assessed. Subjects were presented with a series of category names (e.g., WEATHER PHENOMENA) and they were required to produce as many exemplars as they could in a 30 second period. Using this conceptually-driven indirect measure of memory, Srinivas and Roediger (1990) found the typical generation effect. Generating a target word at the time of study led to a higher probability of producing that word as a response to a category name, than did reading a

target word at the time of study. These results indicate that the category association test is also useful in revealing the effects of prior conceptually-driven study processes (see Kinoshita, 1989, however, for a different result with a nonsemantic generation task).

Srinivas and Roediger (1990) have provided further evidence for this conclusion. In a second experiment, they examined the effect of semantic encoding on the category association test. In the study phase, subjects were presented with a series of words (e.g., SWEATER) under semantic encoding conditions (e.g., RATE THE WORD ON A PLEASANTNESS SCALE) and structural encoding conditions (e.g., COUNT THE NUMBER OF CONSONANTS). This presentation was then followed by a category association test. Subjects were presented with a category name (e.g., AN ARTICLE OF CLOTHING) and they were required to produce as many exemplars as they could in a 30 second period. Using this conceptually-driven indirect measure of memory, Srinivas and Roediger (1990) found an advantage for semantic encoding. Studying a word under semantic encoding conditions led to a higher probability of producing that word as a response to a category name, than did studying a word under structural encoding conditions (see Hamman, 1990 and Kinoshita, 1989 for similar results).

Finally, the classification of indirect measures of memory into data-driven tests and conceptually-driven tests can be used to explain a previous result that was obtained by Graf and Schacter (1985). In one experiment, they had subjects study a series of related word pairs (e.g., MOLD-BREAD) and unrelated word pairs (e.g., DRYER-BLOCK) under semantic encoding conditions (e.g., GENERATE A MEANINGFUL SENTENCE WITH THE TWO WORDS) or structural encoding conditions (e.g., COMPARE THE NUMBER OF VOWELS IN THE TWO WORDS). In the test phase, the subjects were given a word fragment completion test in which the fragments were presented in the same context as at study (e.g., MOLD-BRE___; DRYER-BLO___) or in a different context (e.g., RIPE-

BRE___; JAIL-BLO___). The subject's task was to complete each fragment with the first word that came to mind. Using this experimental paradigm, Graf and Schacter (1985) found a semantic encoding advantage for both types of word pairs. Studying the related and unrelated word pairs under semantic encoding conditions led to the completion of more same-context fragments than different-context fragments, whereas studying the related and unrelated word pairs under structural encoding conditions led to the completion of an equal number of same-context fragments and different-context fragments.

This result was difficult to interpret at the time, since a number of previous studies failed to reveal a semantic encoding advantage on a word fragment completion test (e.g., Graf et al., 1982; Graf & Mandler, 1984). These discrepant findings, however, can be explained by using the processing account. In the study of Graf and Schacter (1985), the word fragments were presented in the context of another word (e.g., MOLD-BRE___). Since the completion of the fragments can be influenced by conceptual factors, this measure could be classified as a conceptually-driven word fragment completion test (see Roediger & Blaxton, 1987a, Footnote 7). In contrast, the word fragments in the study of Graf et al. (1982) and Graf and Mandler (1984) were not presented in the context of another word (e.g., BRE___). Since the completion of the fragments can not be influenced by conceptual factors, this measure could be classified as a data-driven word fragment completion test (Roediger & Blaxton, 1987a). Thus, the advantage of a conceptually-driven study process (i.e., a semantic encoding task) over a data-driven study process (i.e., a structural encoding task) in the study of Graf and Schacter (1985) could be attributed to their use of a conceptually-driven indirect measure of memory, while the failure to reveal a similar advantage in the studies of Graf et al. (1982) and Graf and Mandler (1984) could be attributed to their use of a data-driven indirect measure of

memory.

Rationale for the Present Experiment

These recent investigations which have employed conceptually-driven indirect measures of memory (e.g., Blaxton, 1989; Graf & Schacter, 1985; Hamman, 1990; Kinoshita, 1989; Srinivas & Roediger, 1990), and the previous investigations which have employed conceptually-driven direct measures of memory (e.g., Craik & Tulving, 1975; Slamecka & Graf, 1978), have both demonstrated that the mental operations which are engaged by a conceptually-driven study task (e.g., semantic encoding or generation) somehow make an item more available on a conceptually-driven measure of memory (e.g., free recall, cued recall, recognition, general knowledge, category association and conceptually-driven word fragment completion), than the mental operations which are engaged by a data-driven study task (e.g., structural encoding or reading). Since the mental operations which are engaged by a study task can not be directly observed, many theories about the nature of these operations have been developed. For example, when words are presented by themselves during study (e.g., STRIPES), the mental operation of semantic elaboration has been offered as a theoretical explanation for the advantage of semantic encoding over structural encoding (e.g., Anderson & Reder, 1979) and the advantage of generation over reading (e.g., Nairne et al., 1985). Specifically, it has been suggested that the conceptually-driven study tasks (e.g., semantic encoding or generation) activate the semantic attributes of an item to a greater extent than the data-driven study tasks (e.g., structural encoding or reading). This activation then spreads to associated items, thereby creating a network of potential retrieval routes. Similarly, when words are presented in the context of another word during study (e.g., TIGER-STRIPES, TIGER-STR_PES), the mental operation of semantic unitization has been offered as a theoretical explanation for the advantage of semantic encoding over structural encoding

(e.g., Graf & Schacter, 1989) and the advantage of generation over reading (e.g., Donaldson & Bass, 1980). Specifically, it has been suggested that the conceptually-driven study tasks (e.g., semantic encoding or generation) induce the processing of a semantic association between the words to a greater extent than the data-driven study tasks (e.g., structural encoding or reading), thereby increasing the probability that a target word will be produced when only a portion of the word pair is presented in the test phase (e.g., TIGER-____; TIGER-STR__).

While the mental operations which are engaged by a conceptually-driven study task (e.g., the activation of semantic attributes and associations) can not be directly observed, a record of their occurrence should be stored in memory along with the items that were presented in the study phase. Unfortunately, the existing conceptually-driven measures of memory (e.g., free recall, cued recall, recognition, general knowledge, category association and conceptually-driven word fragment completion) have only been designed to index the prior occurrence of items that were presented in the study phase. It is possible, however, that conceptually-driven indirect measures of memory can be designed to index the prior occurrence of specific mental operations. To demonstrate this possibility, consider the results of Graf and Schacter (1985). Studying related word pairs (e.g., MOLD-BREAD) and unrelated word pairs (e.g., DRYER-BLOCK) under semantic encoding conditions (e.g., GENERATE A MEANINGFUL SENTENCE WITH THE TWO WORDS) led to the completion of more same-context fragments (e.g., MOLD-BRE__; DRYER-BLO__) than different-context fragments (e.g., RIPE-BRE__; JAIL-BLO__). In contrast, studying word pairs under structural encoding conditions (e.g., COMPARE THE NUMBER OF VOWELS IN THE TWO WORDS) led to the completion of an equal number of same-context fragments and different-context fragments. To account for these results, Graf and Schacter (1989) suggested that the advantage of the

sentence generation task over the letter comparison task was the result of a difference in the mental operation of semantic unitization. Specifically, they argued that the sentence generation task induced the processing of a meaningful connection or association between the members of a word pair. This in turn led to the formation of a unitized representation in memory which could later be reintegrated on a word fragment completion test. On the other hand, they argued that the letter comparison task did not lead to a unitized representation in memory because it did not induce the processing of a meaningful connection or association between the members of a word pair.

While this account of the mental operations engaged by the sentence generation task and the letter comparison task is intuitively appealing, one way to empirically verify that these mental operations actually occurred in the study phase would be to re-present the studied word pairs (i.e., old related word pairs and old unrelated word pairs) along with some new related word pairs and new unrelated word pairs and ask subjects to judge, as quickly as possible, whether the words in each pair were related or unrelated. If the sentence generation task does induce the processing of a meaningful association between the members of a word pair, and the letter comparison task does not, then sentence generation subjects should be faster at making "related" decisions about old related word pairs than new related word pairs (i.e., the recent meaningful association of two related words should facilitate the ability to make a "related" decision) and letter comparison subjects should take a similar amount of time to make "related" decisions about old related word pairs and new related word pairs. In contrast, sentence generation subjects should be slower at making "unrelated" decisions about old unrelated word pairs than new unrelated word pairs (i.e., the recent meaningful association of two unrelated words should inhibit the ability to make an "unrelated" decision) and letter comparison subjects should take a similar amount of time to make "unrelated" decisions about old

unrelated word pairs and new unrelated word pairs.

Another way to verify that the sentence generation task induces the processing of a meaningful association between the members of a word pair would be to re-present the studied word pairs (i.e., old related word pairs and old unrelated word pairs) along with some new related word pairs and new unrelated word pairs and ask subjects to judge how related they felt the words were. If the sentence generation task does induce the processing of a meaningful association between the members of a word pair, and the letter comparison task does not, then sentence generation subjects should rate old related word pairs more highly than new related word pairs (i.e., the recent meaningful association of two related words should make them seem more related) and letter comparison subjects should rate old related word pairs about the same as new related word pairs. Similarly, sentence generation subjects should rate old unrelated word pairs more highly than new unrelated word pairs (i.e., the recent meaningful association of two unrelated words should also make them seem more related) and letter comparison subjects should rate old unrelated word pairs about the same as new unrelated word pairs.

The present experiment will employ these two measures of memory, which I will refer to as the semantic decision test and the subjective evaluation test, in order to investigate whether conceptually-driven indirect measures of memory can be designed to index the prior occurrence of specific conceptually-driven mental operations. Obtaining the expected results will not only provide support for Graf and Schacter's (1989) semantic unitization hypothesis but will indicate that future studies could use a similar methodology (i.e., examining the re-processing of study items) to investigate other (perhaps less intuitive) theories about the mental operations which are engaged by conceptually-driven study tasks (These implications will be described in more detail in the Discussion).

Study Conditions

In the present experiment, the study conditions were based on the experiment conducted by Graf and Schacter (1985). In the study phase, subjects were presented with a series of related (e.g., NEEDLE-THREAD) and unrelated word pairs (e.g., CELERY-MUSKET). One-half of the subjects were asked to generate a meaningful sentence using the two words (the sentence generation condition) and one-half of the subjects were asked to compare the number of vowels and consonants in the two words (the letter comparison condition). The letter comparison task was primarily a data-driven study task since it did not require the subject to go beyond the physical features of the word pair. The sentence generation task was primarily a conceptually-driven study task, since it required the subject to go beyond the physical features of the word pair and consider the meanings of each word. The subject's memory was then assessed on either the semantic decision test, the subjective evaluation test or the conceptually-driven word fragment completion test that was employed by Graf and Schacter (1985).

Semantic Decision Test

One group of subjects from each study condition received the semantic decision test. During this test, subjects were presented with a series of old related word pairs (e.g., NEEDLE-THREAD), new related word pairs (e.g., MONKEY-BANANA), old unrelated word pairs (e.g., CELERY-MUSKET) and new unrelated word pairs (e.g., POODLE-PASTRY). The subject's task was to judge, as quickly as possible, whether the words were related or unrelated. This test was primarily a conceptually-driven indirect measure of memory since it made no reference to the study phase of the experiment, and it required the subject to go beyond the physical features of the word pair and consider the semantic relationship between the members of the word pair.

The primary measure of interest on this test was the time required to make semantic

decisions. If the sentence generation task does induce the processing of a meaningful association between the members of a word pair, and the letter comparison task does not, then sentence generation subjects should be faster at making "related" decisions about old related word pairs than new related word pairs (i.e., the recent meaningful association of two related words should facilitate the ability to make a "related" decision) and letter comparison subjects should take a similar amount of time to make "related" decisions about old related word pairs and new related word pairs. In contrast, sentence generation subjects should be slower at making "unrelated" decisions about old unrelated word pairs than new unrelated word pairs (i.e., the recent meaningful association of two unrelated words should inhibit the ability to make an "unrelated" decision) and letter comparison subjects should take a similar amount of time to make "unrelated" decisions about old unrelated word pairs and new unrelated word pairs.

A second measure of interest on the semantic decision test was the number of errors made (i.e., responding that a related word pair was "unrelated" or responding that an unrelated word pair was "related"). If the sentence generation task does induce the processing of a meaningful association between the members of a word pair, and the letter comparison task does not, then sentence generation subjects should make fewer errors on old related word pairs than new related word pairs (i.e., the recent meaningful association of two related words should decrease the probability of misclassifying them as "unrelated") and letter comparison subjects should make a similar number of errors on old related word pairs and new related word pairs. In contrast, sentence generation subjects should make more errors on old unrelated word pairs than new unrelated word pairs (i.e., the recent meaningful association of two unrelated words should increase the probability of misclassifying them as "related") and letter comparison subjects should make a similar number of errors on old unrelated word pairs and new unrelated word pairs.

Subjective Evaluation Test

A second group of subjects from each study condition received the subjective evaluation test. During this test, subjects were presented with a series of old related word pairs (e.g., NEEDLE-THREAD), new related word pairs (e.g., MONKEY-BANANA), old unrelated word pairs (e.g., CELERY-MUSKET) and new unrelated word pairs (e.g., POODLE-PASTRY). The subject's task was to judge as quickly as possible how related they felt the words were on a scale from 1 (WEAKLY RELATED) to 5 (STRONGLY RELATED). This test was primarily a conceptually-driven indirect measure of memory, since it made no reference to the study phase of the experiment and it required the subject to go beyond the physical features of the word pair and consider the semantic relationship between the members of the word pair.

The primary measure of interest on this test were the subjective evaluation ratings. If the sentence generation task does induce the processing of a meaningful association between the members of a word pair, and the letter comparison task does not, then sentence generation subjects should rate old related word pairs more highly than new related word pairs (i.e., the recent meaningful association of two related words should make them seem more related) and letter comparison subjects should rate old related word pairs about the same as new related word pairs. Similarly, sentence generation subjects should rate old unrelated word pairs more highly than new unrelated word pairs (i.e., the recent meaningful association of two unrelated words should make them seem more related) and letter comparison subjects should rate old unrelated word pairs about the same as new unrelated word pairs.

A second measure of interest on the subjective evaluation test was the time required to make subjective evaluations. The predictions for this test are identical to those of the semantic decision test. On the related word pairs, sentence generation subjects should

be faster at making subjective evaluations about old related word pairs than new related word pairs (i.e., the recent meaningful association of two related words should facilitate the ability to make subjective evaluations about related word pairs) and letter comparison subjects should take a similar amount of time to make subjective evaluations about old related word pairs and new related word pairs. On the unrelated word pairs, sentence generation subjects should be slower at making subjective evaluations about old unrelated word pairs than new unrelated word pairs (i.e., the recent meaningful association of two unrelated words should inhibit the ability to make subjective evaluations about unrelated word pairs) and letter comparison subjects should take a similar amount of time to make subjective evaluations about old unrelated word pairs and new unrelated word pairs.

A third measure of interest on the subjective evaluation test was the number of errors made. On this test, an error was operationally defined as occurring when a related word pair was given a low rating on the scale (e.g., 1 or 2) or an unrelated word pair was given a high rating on the scale (e.g., 4 or 5). The predictions for this test are identical to those of the semantic decision test. On the related word pairs, sentence generation subjects should make fewer errors on old related word pairs than new related word pairs (i.e., the recent meaningful association of two related words should decrease the probability of misclassifying them as "unrelated") and letter comparison subjects should make a similar number of errors on old related word pairs and new related word pairs. On the unrelated word pairs, the sentence generation subjects should make more errors on old unrelated word pairs than new unrelated word pairs (i.e., the recent meaningful association of two unrelated words should increase the probability of misclassifying them as "related") and letter comparison subjects should make a similar number of errors on old unrelated word pairs and new unrelated word pairs.

Word Fragment Completion Test

One group of subjects from each study condition received the conceptually-driven word fragment completion test that was employed by Graf and Schacter (1985). During this test, subjects were presented with a series of incomplete old related word pairs (e.g., NEEDLE-THR__), new related word pairs (e.g., MONKEY-BAN__), old unrelated word pairs (e.g., CELERY-MUS__) and new unrelated word pairs (e.g., POODLE-PAS__). The subject's task was to read the initial word aloud and then say the first word that came to mind in response to the three-letter word stem.

The primary measure of interest on this test was the number of word pair fragments that were completed with target words. The predictions for this test were derived from the study of Graf and Schacter (1985). When they used this measure together with the study tasks employed in the present experiment, they found a semantic encoding advantage for both the related and the unrelated word pairs. Specifically, sentence generation subjects completed a greater number of old related word pair fragments than new related word pair fragments with target words than did letter comparison subjects, and they also completed a greater number of old unrelated word pair fragments than new unrelated word pair fragments with target words than did letter comparison subjects. In addition, they found that subjects in both study conditions provided more target words in the context of a related word (e.g., NEEDLE-THR__) than in the context of an unrelated word (e.g., CELERY-MUS__). If Graf and Schacter's (1985) findings are replicated in the present experiment, this will indicate that the advantage of a conceptually-driven study task (e.g., sentence generation) over a data-driven study task (e.g., letter comparison) can be revealed with the present materials and instructions. Thus, any failure to obtain the expected results with the semantic decision test or the subjective evaluation test can not be attributed to these sources.

Method

Subjects

The subjects consisted of 144 undergraduate students from Carleton University. They participated in a 30 minute session for credit in an introductory psychology course. Two subjects were replaced because they had difficulty completing the sentence generation task. One other subject was replaced because he made too many errors on the semantic decision test (i.e., he made errors on more than 40% of the word pairs).

Design

A 3 x 2 x 2 x 2 mixed factorial design was used with test condition (word fragment completion, semantic decision or subjective evaluation) and study condition (sentence generation or letter comparison) as between-subjects factors and word pair relation (related or unrelated) and item status (old or new) as within-subjects factors. Twenty-four subjects were randomly assigned to each of the six experimental conditions formed by the factorial combination of test condition and study condition.

Materials

A set of 48 related word pairs (e.g., NEEDLE-THREAD) and a set of 48 unrelated word pairs (e.g., CELERY-MUSKET) were constructed by pairing low frequency nouns (1-20 per million) from the Kucera and Francis (1967) word norms. The 192 words selected were all between 4 and 6 letters in length and they all had unique three-letter word stems (e.g., NEE, THR, CEL, MUS). This ensured that items on the word fragment completion test (e.g., NEEDLE-THR__) could be completed with only one item from the total set of 192 words (e.g., THREAD). In addition, for the 96 target words selected, an English dictionary listed at least 8 other words with the same three-letter word stem (e.g., THREAD, THREAT, THRIFT, THROW, THRILL, THRUSH, THRONG, THRUST, THROB). This ensured that the target word was only one of many possible solutions for the

corresponding item on the word fragment completion test. The set of related word pairs, and the set of unrelated word pairs, were divided into two pools that were matched for word length (first related word pair pool: $M = 5.0$, $SD = 0.8$; second related word pair pool: $M = 5.0$, $SD = 0.8$; first unrelated word pair pool: $M = 5.0$, $SD = 0.8$; second unrelated word pair pool: $M = 5.0$, $SD = 0.8$) and word frequency (first related word pair pool: $M = 9.7$, $SD = 6.8$; second related word pair pool: $M = 9.7$, $SD = 7.1$; first unrelated word pair pool: $M = 9.9$, $SD = 6.8$; second unrelated word pair pool: $M = 9.9$, $SD = 6.7$). One format was then constructed by randomly assigning a related word pair pool and an unrelated word pair pool to the old condition and a related word pair pool and an unrelated word pair pool to the new condition (see Appendix A). A second format was then constructed by rotating the word pairs through the condition of item status such that old word pairs in the first format were new word pairs in the second format (see Appendix B). One-half of the subjects within each of the six experimental conditions formed by the factorial combination of test condition and study condition received the first format while the remaining half received the second format. An additional 8 related word pairs and an additional 8 unrelated word pairs were selected for use as study practice items, primacy fillers (i.e., items at the beginning of the study list), recency fillers (i.e., items at the end of the study list) and test practice items (see Appendix C).

Procedure

The subjects were tested in individual sessions. Upon their arrival in the laboratory subjects were informed of the procedure involved in the study phase; however, they were not informed of the procedure involved in the test phase. Subjects in the letter comparison study condition were told that they would be presented with a series of word pairs (e.g., BOWL-SPOON; PENNY-BLADE) and that their task would be to compare the number of vowels and consonants in the two words. Specifically, they were instructed to

say the word VOWELS and then indicate (e.g., YES or NO) whether the two words had the same number of vowels. Then they were to say the word CONSONANTS and indicate (e.g., YES or NO) whether the two words had the same number of consonants. Subjects in the sentence generation study condition were told that they would be presented with a series of word pairs (e.g., BOWL-SPOON; PENNY-BLADE) and that their task would be to generate a meaningful sentence using the two words. The subjects in both study conditions were also informed that the computer would be recording response latencies on each trial.

In the study phase, each subject was presented with a series of 56 word pairs. The first four items and the last four items each consisted of 2 related word pairs and 2 unrelated word pairs. These items were presented in the same order for all subjects. They were added to minimize any primacy and recency effects that may occur in the test phase (i.e., the tendency to remember items at the beginning and end of a study list). The remaining items consisted of 24 related word pairs and 24 unrelated word pairs. These items were presented in a new random order for each subject. Each trial began when a word pair appeared in an uppercase format on the center line of the computer monitor. When it appeared the subject was required to perform their given task (i.e., letter comparison or sentence generation). After giving a response the subject was required to press the spacebar on the computer keyboard. This was then followed by a 1000 ms blank interval. This sequence of events was repeated until the entire list of 56 word pairs had been presented. During the study phase, a tape-recorder recorded the subject's oral responses while the computer recorded latencies on each trial from the onset of a word pair to the spacebar press of the subject. In order to acquaint the subjects with this procedure, the presentation of study items was preceded by a practice phase which consisted of 2 related word pairs and 2 unrelated word pairs.

The test phase immediately followed the study phase, and again, the subjects were not informed of the relationship between the test phase and the study phase. In all three test conditions, subjects were presented with a random series of 96 word pairs. These items consisted of the 48 word pairs presented in the study phase plus 48 new word pairs (24 related and 24 unrelated). Subjects in the word fragment completion test condition were told that they would be presented with a series of word pairs in an incomplete format (e.g., PARCEL-STA__; ARROW- PER__) and that their task would be to complete the fragment with the first word that comes to mind. Each trial began when an incomplete word pair was presented in an uppercase format on the center line of the computer monitor. The subject's task was to read the initial word aloud and then say the first word that came to mind in response to the three-letter word stem. The subjects were told that the word they produced did not have to be related to the initial word. After giving a response the subject was required to press the spacebar on the computer keyboard. This was then followed by a 750 ms blank interval. This sequence of events was repeated until the entire series of 96 word pairs had been presented. A tape-recorder recorded the subject's oral responses.

Subjects in the semantic decision test condition were told that they would be presented with a series of word pairs (e.g., PARCEL-STAMP; ARROW-PERCH) and that their task would be to judge whether the words were related or unrelated. Each trial began when the message "PRESS THE MIDDLE KEY WHEN READY" appeared on the bottom line of the computer monitor. When the subject initiated the trial, by pressing a designated key on a response box, the message was removed and the screen remained blank for a period of 750 ms. A fixation display (e.g., +) was then presented in the center of the monitor for a period of 750 ms. This was followed by another blank interval of 750 ms. The word pair was then presented in an uppercase format on the center line of the

monitor. The subject's task was to indicate, as quickly as they could, whether the words were related or unrelated by using the index finger of one hand to press a key labelled "R" on the response box or by using the index finger of the other hand to press a key labelled "U" on the response box. One-half of the subjects made the "R" response with their right hand and one-half of the subjects made the "R" response with their left hand. After a response was made, the word pair was removed and the screen remained blank for a period of 750 ms. This sequence of events was repeated until the entire series of 96 word pairs had been presented. The computer recorded the responses selected (e.g., "R" or "U") and calculated latencies, on each trial, from the onset of a word pair to the key press of the subject.

Subjects in the subjective evaluation test condition were told that they would be presented with a series of word pairs (e.g., PARCEL-STAMP; ARROW-PERCH) and that their task would be to judge how related they felt the words were. Each trial began when a word pair was presented, in an uppercase format, on the center line of the computer monitor. The subject's task was to rate the relationship between the words on a scale from 1 (WEAK) to 5 (STRONG) by pressing a key, from 1 to 5, on the top row of the computer keyboard. The subjective nature of this test was pointed out to subjects and they were encouraged to use the full range of values when making their judgements. In addition, they were encouraged to select their ratings, as quickly as possible, by going with their first instinct about the relationship between the words. After a response was made, the word pair was removed and the screen remained blank for a period of 750 ms. This sequence of events was repeated until the entire series of 96 word pairs had been presented. The computer recorded the responses selected (e.g., "1", "2", "3", "4" or "5") and calculated latencies on each trial from the onset of a word pair to the key press of the subject.

In order to acquaint subjects with the procedure that was involved in their particular test condition, the presentation of test items was preceded by a practice phase which consisted of 2 related word pairs and 2 unrelated word pairs.

Results

The data obtained from the study phase (e.g., response latencies), the word fragment completion test (e.g., target words produced), the semantic decision test (e.g., errors and response latencies) and the subjective evaluation test (e.g., errors, response latencies and subjective ratings) were analyzed separately. The results of these analyses are outlined in the following sections. Each section (1) describes the data involved, (2) specifies the analyses conducted, (3) outlines the predictions made (these predictions are also summarized in Table 1), (4) provides a figure of the data obtained, (5) provides an analysis of variance summary table, and (6) highlights the important results. Since the predictions involve a number of higher-order interactions, main effects and interactions will only be discussed if they do not enter into any higher-order interactions (an alpha level of .05 was used for all statistical analyses, unless stated otherwise).

Study Phase

Response latencies. The median number of milliseconds required to complete the study tasks were calculated for each subject. The data were analyzed with a $3 \times 2 \times 2$ mixed-design analysis of variance with test condition (word fragment completion, semantic decision or subjective evaluation) and study condition (sentence generation or letter comparison) as between-subjects factors and word pair relation (related or unrelated) as a within-subjects factor. The analysis was conducted as a check on two aspects of the experimental design. First, it would confirm that the sentence generation task was a conceptually-driven study task and that the letter comparison task was a data-

Table 1

Summary of Predictions.Study Phase: Response Latencies

SG faster latencies on RWP than UWP
 LC similar latencies on RWP and UWP

Word Fragment Completion Test: Target Words Produced

SG more targets on ORWP than NRWP more targets on OUWP than NUWP
 LC similar targets on ORWP and NRWP similar targets on OUWP and NUWP

Semantic Decision Test: Errors

SG fewer errors on ORWP than NRWP more errors on OUWP than NUWP
 LC similar errors on ORWP and NRWP similar errors on OUWP and NUWP

Semantic Decision Test: Response Latencies

SG faster latencies on ORWP than NRWP slower latencies on OUWP than NUWP
 LC similar latencies on ORWP and NRWP similar latencies on OUWP and NUWP

Subjective Evaluation Test: Errors

SG fewer errors on ORWP than NRWP more errors on OUWP than NUWP
 LC similar errors on ORWP and NRWP similar errors on OUWP and NUWP

Subjective Evaluation Test: Response Latencies

SG faster latencies on ORWP than NRWP slower latencies on OUWP than NUWP
 LC similar latencies on ORWP and NRWP similar latencies on OUWP and NUWP

Subjective Evaluation Test: Ratings

SG higher ratings on ORWP than NRWP higher ratings on OUWP than NUWP
 LC similar ratings on ORWP and NRWP similar ratings on OUWP and NUWP

SG = Sentence Generation, LC = Letter Comparison, RWP = Related Word Pairs, UWP = Unrelated Word Pairs, ORWP = Old Related Word Pairs, NRWP = New Related Word Pairs, OUWP = Old Unrelated Word Pairs, NUWP = New Unrelated Word Pairs

driven study task. If the sentence generation task was a conceptually-driven study task, subjects should complete the sentence generation task with the related word pairs more quickly than with the unrelated word pairs (i.e., it should be easier to generate a sentence with two related words than two unrelated words). If the letter comparison task was a data-driven study task, subjects should take a similar amount of time to complete the letter comparison task with the related word pairs and the unrelated word pairs. This pattern of results would be revealed in an interaction between study condition and word pair relation. Second, the analysis would confirm that subjects in the different test conditions were treated similarly in the study phase. If they did have similar experiences in the study phase, then there should be no main effect of test condition, and test condition should not enter into any higher-order interactions. Obtaining this result will permit comparisons between the word fragment completion test, the semantic decision test and the subjective evaluation test.

The mean response latencies in the study phase are displayed in Figure 1 as a function of study condition, test condition and word pair relation, and the results of the analysis of variance are presented in Table 2. The analysis revealed the expected interaction between study condition and word pair relation, $F(1,138) = 149.619$, $MSe = 943207.140$. A simple effects analysis revealed that the sentence generation subjects completed their study task with the related word pairs ($M = 5221$, $SD = 1613$) more quickly than with the unrelated word pairs ($M = 8077$, $SD = 2681$), $F(1,138) = 311.220$, $MSe = 943207.140$, while the letter comparison subjects took a similar amount of time to complete their study task with the related word pairs ($M = 6346$, $SD = 2161$) and the unrelated word pairs ($M = 6402$, $SD = 2164$), $F(1,138) < 1$. This result indicates that "the sentence generation task was a conceptually-driven study task and that the letter comparison task was a data-driven study task. The analysis also failed to reveal a main

Figure 1. Mean Response Latencies (in ms) in the Study Phase as a Function of Study Condition, Test Condition (WFC = Word Fragment Completion, SD = Semantic Decision, SE = Subjective Evaluation) and Word Pair Relation.

Table 2

Analysis of Variance Summary Table for Response Latencies (in ms) in the Study Phase
as a Function of Test Condition, Study Condition and Word Pair Relation.

Source of Variance	SS	DF	MS	F	
Test Condition	5941508.007	2	2970754.003	0.338	
Study Condition	5452977.920	1	5452977.920	0.620	
Test Condition by Study Condition	7312841.340	2	3656420.670	0.416	
Error	1213660143.646	138	878638.722		
Word Pair Relation	152534944.531	1	152534944.531	161.719	***
Test Condition by Word Pair Relation	639536.271	2	319768.135	0.339	
Study Condition by Word Pair Relation	141121400.003	1	141121400.003	149.619	***
Test Condition by Study Condition by Word Pair	1410614.382	2	705307.191	0.748	
Error	130162585.313	138	943207.140		

* $p < .05$, ** $p < .01$, *** $p < .001$

effect of test condition and test condition did not enter into any higher-order interactions. This result indicates that the subjects in the different test conditions had similar experiences in the study phase. In turn, it suggests that it is legitimate to make comparisons in performance between the word fragment completion test, the semantic decision test and the subjective evaluation test.

Word Fragment Completion Test

Target words produced. The proportion of fragments completed with target words was calculated for each subject. The data were analyzed with a 2 x 2 x 2 mixed-design analysis of variance with study condition (sentence generation or letter comparison) as a between-subjects factor and word pair relation (related or unrelated) and item status (old or new) as within-subjects factors. The predictions for this analysis were derived from Graf and Schacter (1985). When they used this word fragment completion test and the study tasks employed in the present experiment, they found a semantic encoding advantage for both the related and the unrelated word pairs. Specifically, sentence generation subjects completed a greater number of old related word pair fragments than new related word pair fragments with target words than did letter comparison subjects, and they also completed a greater number of old unrelated word pair fragments than new unrelated word pair fragments with target words than did letter comparison subjects. This pattern of results would be revealed in the present experiment in an interaction between study condition and item status. Graf and Schacter (1985) also found that subjects in both study conditions provided more target words in the context of a related word (e.g., NEEDLE-THR__) than in the context of an unrelated word (e.g., CELERY-MUS__). In the present experiment, this would be revealed in a main effect of word pair relation.

The mean proportion of targets provided on the word fragment completion test are displayed in Figure 2 as a function of word pair relation, study condition and item status,

and the results of the analysis of variance are presented in Table 3. The analysis revealed the expected interaction between study condition and item status, $F(1,46) = 57.499$, $MSe = 0.010$. A simple effects analysis revealed that the sentence generation subjects completed many more old word pair fragments ($M = .61$, $SD = .28$) than new word pair fragments ($M = .32$, $SD = .24$), $F(1,46) = 195.280$, $MSe = 0.010$, while the letter comparison subjects completed a few more old word pair fragments ($M = .39$, $SD = .25$) than new word pair fragments ($M = .32$, $SD = .25$), $F(1,46) = 10.566$, $MSe = 0.010$. The analysis also revealed the expected main effect of word pair relation, $F(1,46) = 672.993$, $MSe = 0.014$. Specifically, subjects provided more target words in the context of a related word ($M = .63$, $SD = .18$) than in the context of an unrelated word ($M = .18$, $SD = .16$). These findings replicate the results obtained by Graf and Schacter (1985) and they indicate that the advantage of a conceptually-driven study task (e.g., sentence generation) over a data-driven study task (e.g., letter comparison) can be revealed on a conceptually-driven indirect measure of memory (e.g., conceptually-driven word fragment completion test) with the present materials and instructions. Thus, a failure to obtain the expected results with the semantic decision test or the subjective evaluation can not be attributed to these sources.

Semantic Decision Test

Errors. The proportion of errors made on the semantic decision test (i.e., responding that a related word pair was "unrelated" or responding that an unrelated word pair was "related") was calculated for each subject. The data were analyzed with a $2 \times 2 \times 2$ mixed-design analysis of variance with study condition (sentence generation or letter comparison) as a between-subjects factor and word pair relation (related or unrelated) and item status (old or new) as within-subjects factors. If the sentence generation task does induce the processing of a meaningful association between the members of a word pair,

Figure 2. Mean Proportion of Targets on the Word Fragment Completion Test as a Function of Word Pair Relation, Study Condition and Item Status.

Table 3

Analysis of Variance Summary Table for the Proportion of Targets on the Word Fragment Completion Test as a Function of Study Condition, Word Pair Relation and Item Status.

Source of Variance	SS	DF	MS	F	
Study Condition	0.547	1	0.547	22.034	***
Error	1.142	46	0.025		
Word Pair Relation	9.368	1	9.368	672.993	***
Study Condition by Word Pair Relation	0.003	1	0.003	0.189	
Error	0.640	46	0.014		
Item Status	1.507	1	1.507	148.347	***
Study Condition by Item Status	0.584	1	0.584	57.499	***
Error	0.467	46	0.010		
Word Pair Relation by Item Status	0.001	1	0.001	0.109	
Study Condition by Word Pair Relation by Item Status	0.009	1	0.009	0.869	
Error	0.488	46	0.011		

* $p < .05$, ** $p < .01$, *** $p < .001$

and the letter comparison task does not, then sentence generation subjects should make fewer errors on old related word pairs than new related word pairs (i.e., the recent meaningful association of two related words should decrease the probability of misclassifying them as "unrelated") and letter comparison subjects should make a similar number of errors on old related word pairs and new related word pairs. In contrast, sentence generation subjects should make more errors on old unrelated word pairs than new unrelated word pairs (i.e., the recent meaningful association of two unrelated words should increase the probability of misclassifying them as "related") and letter comparison subjects should make a similar number of errors on old unrelated word pairs and new unrelated word pairs. This pattern of results would be revealed in an interaction between study condition and item status for the related word pairs that was different from an interaction between study condition and item status for the unrelated word pairs. Overall, this pattern of results would be revealed in a three-way interaction between word pair relation, study condition and item status.

The mean proportion of errors made on the semantic decision test are displayed in Figure 3 as a function of word pair relation, study condition and item status, and the results of the analysis of variance are presented in Table 4. The analysis revealed the expected interaction between study condition, word pair relation and item status, $F(1,46) = 6.001$, $MSe = 0.003$. This three-way interaction was investigated in two separate 2 (study condition) $\times$ 2 (item status) analyses of variance for the related word pairs and the unrelated word pairs. The first analysis with the related word pairs revealed a marginally significant interaction between study condition and item status, $F(1,46) = 3.191$, $MSe = .003$, $p < .08$. A simple effects analysis revealed that the sentence generation subjects did make fewer errors on the old related word pairs ($M = .05$, $SD = .05$) than the new related word pairs ($M = .12$, $SD = .08$), $F(1,46) = 17.302$, $MSe = .003$, while the letter

Figure 3. Mean Proportion of Errors on the Semantic Decision Test as a Function of Word Pair Relation, Study Condition and Item Status.

Table 4

Analysis of Variance Summary Table for the Proportion of Errors on the Semantic Decision Test as a Function of Study Condition, Word Pair Relation and Item Status.

Source of Variance	SS	DF	MS	F
Study Condition	0.005	1	0.005	0.951
Error	0.218	46	0.005	
Word Pair Relation	0.091	1	0.091	16.578 ***
Study Condition by Word Pair Relation	0.021	1	0.021	3.740
Error	0.254	46	0.006	
Item Status	0.014	1	0.014	7.255 **
Study Condition by Item Status	0.001	1	0.001	0.204
Error	0.086	46	0.002	
Word Pair Relation by Item Status	0.045	1	0.045	17.620 ***
Study Condition by Word Pair Relation by Item Status	0.015	1	0.015	6.001 *
Error	0.117	46	0.003	

* $p < .05$, ** $p < .01$, *** $p < .001$

comparison subjects made a similar number of errors on the old related word pairs ($M = .08$, $SD = .07$) and the new unrelated word pairs ($M = .11$, $SD = .06$), $F(1,46) < 2.7$. The second analysis with the unrelated word pairs also revealed the expected interaction between study condition and item status, $F(1,46) = 4.503$, $MSe = .001$. A simple effects analysis revealed that the sentence generation subjects did make more errors on the old unrelated word pairs ($M = .07$, $SD = .08$) than the new unrelated word pairs ($M = .04$, $SD = .05$), $F(1,46) = 8.271$, $MSe = .001$, while the letter comparison subjects made a similar number of errors on the old unrelated word pairs ($M = .03$, $SD = .05$) and the new unrelated word pairs ($M = .03$, $SD = .04$), $F(1,46) < 1$.

Response latencies. The median number of milliseconds required to make correct semantic decisions (i.e., responding that a related word pair was "related" or responding that an unrelated word pair was "unrelated") were calculated for each subject. The data were analyzed with a $2 \times 2 \times 2$ mixed-design analysis of variance with study condition (sentence generation or letter comparison) as a between-subjects factor and word pair relation (related or unrelated) and item status (old or new) as within-subjects factors. If the sentence generation task does induce the processing of a meaningful association between the members of a word pair, and the letter comparison task does not, then sentence generation subjects should be faster at making "related" decisions about old related word pairs than new related word pairs (i.e., the recent meaningful association of two related words should facilitate the ability to make a "related" decision) and letter comparison subjects should take a similar amount of time to make "related" decisions about old related word pairs and new related word pairs. In contrast, sentence generation subjects should be slower at making "unrelated" decisions about old unrelated word pairs than new unrelated word pairs (i.e., the recent meaningful association of two unrelated words should inhibit the ability to make an "unrelated" decision) and letter comparison

subjects should take a similar amount of time to make "unrelated" decisions about old unrelated word pairs and new unrelated word pairs. This pattern of results would be revealed in an interaction between study condition and item status for the related word pairs that was different from an interaction between study condition and item status for the unrelated word pairs. Overall, this pattern of results would be revealed in a three-way interaction between word pair relation, study condition and item status.

The mean response latencies on the semantic decision test are displayed in Figure 4 as a function of word pair relation, study condition and item status, and the results of the analysis of variance are presented in Table 5. The analysis failed to reveal the expected interaction between word pair relation, study condition and item status, $F(1,46) < 1$, however, the analysis did reveal an interaction between study condition and item status, $F(1,46) = 14.798$, $MSe = 9414.902$. A simple effects analysis revealed that the sentence generation subjects were much faster at making semantic decisions about old word pairs ($M = 1027$, $SD = 247$) than new word pairs ($M = 1202$, $SD = 326$), $F(1,46) = 78.738$, $MSe = 9414.902$, as compared with the letter comparison subjects (old word pairs, $M = 1062$, $SD = 215$; new word pairs, $M = 1130$, $SD = 244$, $F(1,46) = 11.787$, $MSe = 9414.902$). Thus, the sentence generation task decreased the time required to make semantic decisions (e.g., by 175 ms) to a greater extent than the letter comparison task (e.g., by 68 ms). This result was expected for the related word pairs but it was not expected for the unrelated word pairs. For these items, it was anticipated that the sentence generation subjects would be slower at making semantic decisions about old unrelated word pairs than new unrelated word pairs.

The analysis also revealed an interaction between word pair relation and item status, $F(1,46) = 7.379$, $MSe = 4099.586$. A simple effects analysis revealed that subjects were faster at making semantic decisions about old unrelated word pairs ($M = 1136$, $SD = 235$)

Figure 4. Mean Response Latencies (in ms) on the Semantic Decision Test as a Function of Word Pair Relation, Study Condition and Item Status.

Table 5

Analysis of Variance Summary Table for Response Latencies (in ms) on the Semantic Decision Test as a Function of Study Condition, Word Pair Relation and Item Status.

Source of Variance	SS	DF	MS	F	
Study Condition	15805.021	1	15805.021	0.074	
Error	9782138.792	46	212655.191		
Word Pair Relation	1197640.083	1	1197640.083	47.501	***
Study Condition by Word Pair Relation	36300.000	1	36300.000	1.440	
Error	1159793.917	46	25212.911		
Item Status	712968.750	1	712968.750	75.728	***
Study Condition by Item Status	139320.750	1	139320.750	14.798	***
Error	433085.500	46	9414.902		
Word Pair Relation by Item Status	30250.521	1	30250.521	7.379	**
Study Condition by Word Pair Relation by Item Status	93.521	1	93.521	0.023	
Error	188580.958	46	4099.586		

* $p < .05$, ** $p < .01$, *** $p < .001$

than new unrelated word pairs ($M = 1233$, $SD = 323$), $F(1,46) = 26.817$, $MSe = 8380.755$, however, subjects were much faster at making semantic decisions about old related word pairs ($M = 953$, $SD = 189$) than new related word pairs ($M = 1100$, $SD = 234$), $F(1,46) = 100.993$, $MSe = 5133.733$. Thus, the recent exposure to related word pairs in the study phase of the experiment decreased the time required to make semantic decisions (e.g., by 147 ms) to a greater extent than the recent exposure to unrelated word pairs (e.g., by 97 ms).

Subjective Evaluation Test

Errors. The proportion of errors made on the subjective evaluation test (i.e., responding 1 or 2 for a related word pair or responding 4 or 5 for an unrelated word pair) were calculated for each subject. The data were analyzed with a $2 \times 2 \times 2$ mixed-design analysis of variance with study condition (sentence generation or letter comparison) as a between-subjects factor and word pair relation (related or unrelated) and item status (old or new) as within-subjects factors. The predictions for this analysis are identical to those of the semantic decision test. On the related word pairs, sentence generation subjects should make fewer errors on old related word pairs than new related word pairs (i.e., the recent meaningful association of two related words should decrease the probability of misclassifying them as "unrelated") and letter comparison subjects should make a similar number of errors on old related word pairs and new related word pairs. On the unrelated word pairs, sentence generation subjects should make more errors on old unrelated word pairs than new unrelated word pairs (i.e., the recent meaningful association of two unrelated words should increase the probability of misclassifying them as "related") and letter comparison subjects should make a similar number of errors on old unrelated word pairs and new unrelated word pairs. This pattern of results would be revealed in an interaction between study condition and item status for the related word pairs that was

different from an interaction between study condition and item status for the unrelated word pairs. Overall, this pattern of results would be revealed in a three-way interaction between word pair relation, study condition and item status.

The mean proportion of errors made on the subjective evaluation test are displayed in Figure 5 as a function of word pair relation, study condition and item status, and the results of the analysis of variance are presented in Table 6. The analysis failed to reveal the expected interaction between word pair relation, study condition and item status, $F(1,46) < 1$, however, the analysis did reveal an interaction between word pair relation and item status, $F(1,46) = 11.973$, $MSe = 0.001$. A simple effects analysis revealed that subjects made fewer errors on old related word pairs ($M = .04$, $SD = .04$) than new related word pairs ($M = .07$, $SD = .07$), $F(1,46) = 11.457$, $MSe = 0.002$, however, subjects made a similar number of errors on old unrelated word pairs ($M = .02$, $SD = .04$) and new unrelated word pairs ($M = .02$, $SD = .04$), $F(1,46) < 1$. Unlike the analysis of errors on the semantic decision test, these results do not indicate that the sentence generation task decreases the number of errors on related word pairs (i.e., fewer errors on old related word pairs than new related word pairs) to a greater extent than the letter comparison task and they do not indicate that the sentence generation task increases the number of errors on unrelated word pairs (i.e., more errors on old unrelated word pairs than new unrelated word pairs) to a greater extent than the letter comparison task. One possible reason for this discrepancy will be considered in the Discussion.

The analysis also revealed an interaction between study condition and word pair relation, $F(1,46) = 9.812$, $MSe = 0.002$. A simple effects analysis revealed that the sentence generation subjects ($M = .03$, $SD = .05$) made fewer errors than the letter comparison subjects ($M = .07$, $SD = .06$) on the related word pairs, $F(1,46) = 14.744$, $MSe = 0.003$, however, the sentence generation subjects ($M = .02$, $SD = .04$) made a similar

Figure 5. Mean Proportion of Errors on the Subjective Evaluation Test as a Function of Word Pair Relation, Study Condition and Item Status

Table 6

Analysis of Variance Summary Table for the Proportion of Errors on the Subjective
Evaluation Test as a Function of Study Condition, Word Pair Relation and Item Status.

Source of Variance	SS	DF	MS	F
Study Condition	0.016	1	0.016	5.442 *
Error	0.132	46	0.003	
Word Pair Relation	0.045	1	0.045	18.472 ***
Study Condition by Word Pair Relation	0.024	1	0.024	9.812 **
Error	0.113	46	0.002	
Item Status	0.009	1	0.009	4.477 *
Study Condition by Item Status	0.001	1	0.001	0.545
Error	0.089	46	0.002	
Word Pair Relation by Item Status	0.013	1	0.013	11.973 ***
Study Condition by Word Pair Relation by Item Status	0.001	1	0.001	0.084
Error	0.048	46	0.001	

* $p < .05$, ** $p < .01$, *** $p < .001$

number of errors as the letter comparison subjects ($M = .02$, $SD = .04$) on the unrelated word pairs, $F(1,46) < 1$. Again, these results are somewhat inconsistent with the results obtained on the semantic decision test. The analysis of errors on the semantic decision test (see Table 4) did reveal a marginally significant interaction between study condition and word pair relation, $F(1,46) = 3.740$, $MSe = 0.006$, $p < .06$, however, a simple effects analysis revealed a pattern opposite to the present results. Specifically, on the semantic decision test, the sentence generation subjects ($M = .08$, $SD = .07$) made a similar number of errors as the letter comparison subjects ($M = .09$, $SD = .07$) on the related word pairs, $F(1,46) < 1$, however, the sentence generation subjects ($M = .06$, $SD = .07$) made more errors than the letter comparison subjects ($M = .03$, $SD = .05$) on the unrelated word pairs, $F(1,46) = 4.495$, $MSe = 0.005$. Although the interactions between study condition and word pair relation were different on the semantic decision test and the subjective evaluation test, a comparison of Figures 3 and 5 indicate that the pattern of old-new differences (which reveal the effects of a recent exposure) were similar on the semantic decision test and the subjective evaluation test.

Response latencies. The median number of milliseconds required to make correct subjective evaluations (i.e., responding 3, 4 or 5 to related word pairs and responding 1, 2 or 3 to unrelated word pairs) were calculated for each subject. The data were analyzed with a $2 \times 2 \times 2$ mixed-design analysis of variance with study condition (sentence generation or letter comparison) as a between-subjects factor and word pair relation (related or unrelated) and item status (old or new) as within-subjects factors. The predictions for this analysis are identical to those of the semantic decision test. On the related word pairs, sentence generation subjects should be faster at making subjective evaluations about old related word pairs than new related word pairs (i.e., the recent meaningful association of two related words should facilitate the ability to make subjective

evaluations about related word pairs) and letter comparison subjects should take a similar amount of time to make subjective evaluations about old related word pairs and new related word pairs. On the unrelated word pairs, sentence generation subjects should be slower at making subjective evaluations about old unrelated word pairs than new unrelated word pairs (i.e., the recent meaningful association of two unrelated words should inhibit the ability to make subjective evaluations about unrelated word pairs) and letter comparison subjects should take a similar amount of time to make subjective evaluations about old unrelated word pairs and new unrelated word pairs. This pattern of results would be revealed in an interaction between study condition and item status for the related word pairs that was different from an interaction between study condition and item status for the unrelated word pairs. Overall, this pattern of results would be revealed in a three-way interaction between word pair relation, study condition and item status.

The mean response latencies on the subjective evaluation test are displayed in Figure 6 as a function of word pair relation, study condition and item status, and the results of the analysis of variance are presented in Table 7. The analysis failed to reveal the expected interaction between word pair relation, study condition and item status, $F(1,46) < 1$, however, the analysis did reveal an interaction between study condition and item status, $F(1,46) = 29.728$, $MSe = 107493.376$. A simple effects analysis revealed that the sentence generation subjects were faster at making subjective evaluations about old word pairs ($M = 2010$, $SD = 800$) than new word pairs ($M = 2579$, $SD = 1124$), $F(1,46) = 72.148$, $MSe = 107493.376$, while the letter comparison subjects took a similar amount of time making subjective evaluations for old word pairs ($M = 2120$, $SD = 567$) and new word pairs ($M = 2172$, $SD = 618$), $F(1,46) < 1$. Similar to the results on the semantic decision test, the sentence generation task decreased the time required to make subjective evaluations (e.g., by 569 ms) to a greater extent than the letter comparison

Figure 6. Mean Response Latencies (in ms) on the Subjective Evaluation Test as a Function of Word Pair Relation, Study Condition and Item Status.

Table 7

Analysis of Variance Summary Table for Response Latencies (in ms) on the Subjective Evaluation Test as a Function of Study Condition, Word Pair Relation and Item Status.

Source of Variance	SS	DF	MS	F	
Study Condition	1059696.333	1	1059696.333	0.598	
Error	81558635.083	46	1773013.806		
Word Pair Relation	8040944.083	1	8040944.083	16.186	***
Study Condition by Word Pair Relation	751000.333	1	751000.333	1.512	
Error	22851786.083	46	496777.958		
Item Status	4625829.188	1	4625829.188	43.034	***
Study Condition by Item Status	3195588.021	1	3195588.021	29.728	***
Error	4944695.292	46	107493.376		
Word Pair Relation by Item Status	107825.521	1	107825.521	1.164	
Study Condition by Word Pair Relation by Item Status	20791.687	1	20791.687	0.225	
Error	4259612.292	46	92600.267		

* $p < .05$, ** $p < .01$, *** $p < .001$

task (e.g., by 52 ms). Again, this result was expected for the related word pairs but it was not expected for the unrelated word pairs. For these items, it was anticipated that the sentence generation subjects would be slower at making subjective evaluations about old unrelated word pairs than new unrelated word pairs. Finally, the analysis also revealed a main effect of word pair relation, $F(1,46) = 16.186$, $MSe = 496777.958$. This result indicated that the subjects were faster at making subjective evaluations for the related word pairs ($M = 2015$, $SD = 788$) than the unrelated word pairs ($M = 2425$, $SD = 823$).

Subjective ratings. The mean ratings on the subjective evaluation test were calculated for each subject. The data were analyzed with a $2 \times 2 \times 2$ mixed-design analysis of variance with study condition (sentence generation or letter comparison) as a between-subjects factor and word pair relation (related or unrelated) and item status (old or new) as within-subjects factors.³ If the sentence generation task does induce the processing of a meaningful association between the members of a word pair, and the letter comparison task does not, then sentence generation subjects should rate old related word pairs more highly than new related word pairs (i.e., the recent meaningful association of two related words should make them seem more related) and letter comparison subjects should rate old related word pairs about the same as new related word pairs. Similarly, sentence generation subjects should rate old unrelated word pairs more highly than new unrelated word pairs (i.e., the recent meaningful association of two unrelated words should also make them seem more related) and letter comparison subjects should rate old unrelated word pairs about the same as new unrelated word pairs. This pattern of results would be

³ An analysis of variance was used since an interaction was predicted and I know of no suitable nonparametric test. A similar strategy has been followed by recent investigators (e.g., Jacoby, Allan, Collins & Larwill, 1988; Johnson, Kim & Risse, 1985).

revealed in an interaction between study condition and item status.

The mean ratings on the subjective evaluation test are displayed in Figure 7 as a function of word pair relation, study condition and item status, and the results of the analysis of variance are presented in Table 8. The analysis revealed an unexpected interaction between study condition, word pair relation and item status, $F(1,46) = 10.984$, $MSe = 0.008$. This three-way interaction was investigated in two separate 2 (study condition) $\times$ 2 (item status) analyses of variance for the related word pairs and the unrelated word pairs. The first analysis with the related word pairs revealed a marginally significant interaction between study condition and item status, $F(1,46) = 3.466$, $MSe = 0.010$, $p < .07$. A simple effects analysis revealed that the sentence generation subjects rated old related word pairs ($M = 4.73$, $SD = .27$) higher than new related word pairs ($M = 4.63$, $SD = .25$), $F(1,46) = 9.803$, $MSe = 0.010$, while the letter comparison subjects rated old related word pairs ($M = 4.52$, $SD = .26$) the same as new related word pairs ($M = 4.51$, $SD = .31$), $F(1,46) < 1$. Even though this difference for the sentence generation subjects was very small, there was a great deal of consistency in the pattern of results obtained. Nineteen out of 24 sentence generation subjects provided higher ratings for the old related word pairs than the new related word pairs (4 people showed the opposite result and there was 1 tie). The second analysis with the unrelated word pairs also revealed an interaction between study condition and item status, $F(1,46) = 7.560$, $MSe = .006$. A simple effects analysis revealed that the sentence generation subjects rated old unrelated word pairs ($M = 1.31$, $SD = .23$) lower than new unrelated word pairs ($M = 1.40$, $SD = .25$), $F(1,46) = 11.961$, $MSe = 0.006$, while the letter comparison subjects rated old unrelated word pairs ($M = 1.22$, $SD = .22$) the same as new unrelated word pairs ($M = 1.21$, $SD = .18$), $F(1,46) < 1$. Again, even though this difference for the sentence generation subjects was very small, there was a great deal of consistency in the pattern of results

Figure 7. Mean Ratings on the Subjective Evaluation Test as a Function of Word Pair Relation, Study Condition and Item Status.

Table 8

Analysis of Variance Summary Table for the Ratings on the Subjective Evaluation Test as a Function of Study Condition, Word Pair Relation and Item Status.

Source of Variance	SS	DF	MS	F	
Study Condition	1.116	1	1.116	8.272	**
Error	6.208	46	0.135		
Word Pair Relation	526.953	1	526.953	5458.663	***
Study Condition by Word Pair Relation	0.006	1	0.006	0.061	
Error	4.441	46	0.097		
Item Status	0.004	1	0.004	0.423	
Study Condition by Item Status	0.001	1	0.001	0.059	
Error	0.419	46	0.009		
Word Pair Relation by Item Status	0.094	1	0.094	12.219	***
Study Condition by Word Pair Relation by Item Status	0.084	1	0.084	10.984	**
Error	0.352	46	0.008		

* $p < .05$, ** $p < .01$, *** $p < .001$

obtained. Eighteen out of 24 sentence generation subjects provided lower ratings for the old unrelated word pairs than the new unrelated word pairs (4 people showed the opposite result and there were 2 ties). This result was not expected. It indicates that the sentence generation task decreased the subjective ratings of unrelated word pairs (i.e., made unrelated word pairs seem more unrelated). The prediction was that the sentence generation task would increase the subjective ratings of unrelated word pairs (i.e., make unrelated word pairs seem more related).

Discussion

Summary of Results

Response latencies in the study phase. The sentence generation subjects were faster at completing their study task with the related word pairs than the unrelated word pairs ($M = 5221$ ms and $M = 8077$ ms respectively) while the letter comparison subjects took a similar amount of time to complete their study task with the related word pairs and the unrelated word pairs ($M = 6346$ ms and $M = 6402$ ms respectively). Since the performance of the sentence generation subjects was affected by the variation in word pair relation, and the performance of the letter comparison subjects was not, this result confirms that the sentence generation task was primarily a conceptually-driven study task (i.e., one which required subjects to go beyond the physical features of the word pair) and it confirms that the letter comparison task was primarily a data-driven study task (i.e., one which did not require subjects to go beyond the physical features of the word pair). The response latencies in the study phase also revealed that subjects in the different test conditions (i.e., word fragment completion, semantic decision and subjective evaluation) took a similar amount of time to complete their study tasks. This result indicates that it is legitimate to make comparisons in performance between the word fragment completion

test, the semantic decision test and the subjective evaluation test.

Target words produced on the word fragment completion test. The sentence generation subjects completed more old word pair fragments than new word pair fragments with target words ($M = .61$ and $M = .32$ respectively) as compared with the letter comparison subjects ($M = .39$ and $M = .32$ respectively). This result indicates that the advantage of a conceptually-driven study task (i.e., sentence generation) over a data-driven study task (i.e., letter comparison) can be revealed on a conceptually-driven indirect measure of memory (i.e., conceptually-driven word fragment completion test) with the present materials and instructions. Thus, any failure to obtain the expected results with the conceptually-driven semantic decision test or the conceptually-driven subjective evaluation test can not be attributed to these sources.

Errors on the semantic decision test and the subjective evaluation test. On the semantic decision test the sentence generation subjects made fewer errors on the old related word pairs than the new related word pairs ($M = .05$ and $M = .12$ respectively) as compared with the letter comparison subjects ($M = .08$ and $M = .11$ respectively) while the sentence generation subjects made more errors on the old unrelated word pairs than the new unrelated word pairs ($M = .07$ and $M = .04$ respectively) as compared with the letter comparison subjects ($M = .03$ and $M = .03$ respectively). These results were anticipated. Recently processing a meaningful association between the members of a related word pair was expected to decrease the probability of misclassifying them as "unrelated" whereas recently processing a meaningful association between the members of an unrelated word pair was expected to increase the probability of misclassifying them as "related". Thus, these results appear to support the semantic unitization hypothesis.

The same pattern of results, however, was not observed on the subjective evaluation test. The sentence generation subjects did not make fewer errors on the old

related word pairs than the new related word pairs ($M = .02$ and $M = .05$ respectively) as compared with the letter comparison subjects ($M = .06$ and $M = .09$ respectively) and the sentence generation subjects did not make more errors on the old unrelated word pairs than the new unrelated word pairs ($M = .03$ and $M = .02$ respectively) as compared with the letter comparison subjects ($M = .02$ and $M = .02$ respectively). Although it is unclear what caused this discrepancy between the semantic decision test and the subjective evaluation test, one possible reason may lie in the different number of alternatives that were available on these two measures. On the semantic decision test subjects could choose between two alternatives (i.e., related or unrelated). An error was made if subjects judged a related word pair as "unrelated" or an unrelated word pair as "related". On the subjective evaluation test subjects could choose between five alternatives (i.e., "1", "2", "3", "4" or "5"). An error was made if a related word pair was given a low rating on the scale (i.e., "1" or "2") or an unrelated word pair was given a high rating on the scale (i.e., "4" or "5"). A rating in the middle of the scale (i.e., "3"), however, was not considered as an error. The availability of this "middle category" on the subjective evaluation test may explain the discrepancy in error performance between the semantic decision test and the subjective evaluation test.

Response latencies on the semantic decision test and the subjective evaluation test. On the semantic decision test the sentence generation subjects were much faster at making semantic decisions about the old word pairs than the new word pairs (i.e., 175 ms faster) as compared with the letter comparison subjects (i.e., 68 ms faster). Similarly, on the subjective evaluation test the sentence generation subjects were much faster at making subjective evaluations about the old word pairs than the new word pairs (i.e., 569 ms faster) as compared with the letter comparison subjects (i.e., 52 ms faster). These results were anticipated for the related word pairs. Recently processing a meaningful

association between the members of a related word pair was expected to facilitate the ability to make semantic decisions and subjective evaluations about related word pairs. These results, however, were not anticipated for the unrelated word pairs. Recently processing a meaningful association between the members of an unrelated word pair was expected to inhibit the ability to make semantic decisions and subjective evaluations about unrelated word pairs. Thus, the results with the related word pairs appear to support the semantic unitization hypothesis while the results with the unrelated word pairs do not appear to support the semantic unitization hypothesis

Ratings on the subjective evaluation test. The sentence generation subjects gave the old related word pairs higher ratings than the new related word pairs (i.e., the old related word pairs seemed more related than the new related word pairs) while the letter comparison subjects gave similar ratings to the old related word pairs and the new related word pairs. This result was anticipated. Recently processing a meaningful association between the members of a related word pair was expected to make the words seem more related. On the other hand, the sentence generation subjects gave the old unrelated word pairs lower ratings than the new unrelated word pairs (i.e., the old unrelated word pairs seemed more unrelated than the new unrelated word pairs) while the letter comparison subjects gave similar ratings to the old unrelated word pairs and the new unrelated word pairs. This result was not anticipated. Recently processing a meaningful association between the members of an unrelated word pair was expected to make the words seem more related. Thus, the results with the related word pairs appear to support the semantic unitization hypothesis while the results with the unrelated word pairs do not appear to support the semantic unitization hypothesis.

Implications

Implications for the semantic unitization hypothesis It is clear that the results of the

present experiment have implications for Graf and Schacter's (1989) semantic unitization hypothesis. In general, the results obtained with the related word pairs appear to support the hypothesis. When compared with the effects of the letter comparison task, the sentence generation task (1) decreased the number of related word pair errors on the semantic decision test (i.e., fewer errors on old related word pairs than new related word pairs), (2) decreased the time required to make semantic decisions about related word pairs (i.e., faster semantic decisions about old related word pairs than new related word pairs), (3) decreased the time required to make subjective evaluations about related word pairs (i.e., faster subjective evaluations about old related word pairs than new related word pairs) and (4) increased the subjective rating of related word pairs (i.e., old related word pairs seemed more related than new related word pairs). All of these results were predicted from the semantic unitization hypothesis. The only result which did not support the hypothesis was the finding that the sentence generation task did not decrease the number of related word pair errors on the subjective evaluation test to a greater extent than the letter comparison task.

In contrast, the results obtained with the unrelated word pairs do not appear to support the semantic unitization hypothesis. When compared with the effects of the letter comparison task, the sentence generation task did not increase the number of unrelated word pair errors on the subjective evaluation test (i.e., more errors were not made on old unrelated word pairs than new unrelated word pairs). In addition, the sentence generation task (1) decreased the time required to make semantic decisions about unrelated word pairs (i.e., faster semantic decisions about old unrelated word pairs than new unrelated word pairs), (2) decreased the time required to make subjective evaluations about unrelated word pairs (i.e., faster subjective evaluations about old unrelated word pairs than new unrelated word pairs) and (3) decreased the subjective

rating of unrelated word pairs (i.e., old unrelated word pairs seemed more unrelated than new unrelated word pairs). None of these results were predicted from the semantic unitization hypothesis because it explicitly states that the sentence generation task induces the processing of a new meaningful association between the members of an unrelated word pair (e.g., Graf & Schacter, 1989). The only result which did support the hypothesis was the finding that the sentence generation task increased the number of unrelated word pair errors on the semantic decision test (i.e., more errors on old unrelated word pairs than new unrelated word pairs) to a greater extent than the letter comparison task.

A consideration of the subject's reactions while completing the sentence generation task, however, may provide some insight into why the expected results were obtained with the related word pairs but they were not obtained with the unrelated word pairs. Specifically, when the sentence generation subjects were faced with a related word pair they often expressed a sense of relief because it was very easy to produce a sentence ($M = 5221$ ms). When they were faced with an unrelated word pair, however, they often expressed a sense of distress because it was very difficult to produce a sentence ($M = 8077$ ms). This observation indicates that the sentence generation subjects may have been directly encoding the related and unrelated status of the word pairs as they completed their study task. If this was the case, then it is understandable why the sentence generation subjects would be faster at making semantic decisions and subjective evaluations for the old related word pairs than the new related word pairs (i.e., because they just processed the related status of the old related word pairs) and moreover, it is understandable why the sentence generation subjects would be faster at making semantic decisions and subjective evaluations for the old unrelated word pairs than the new unrelated word pairs (i.e., because they just processed the unrelated status

of the old unrelated word pairs). Similarly, it is understandable why the sentence generation subjects would provide higher subjective evaluation ratings for the old related word pairs than the new related word pairs, and lower subjective evaluation ratings for the old unrelated word pairs than the new unrelated word pairs. This explanation can also explain performance on the word fragment completion test (i.e., the completion of more old word pair fragments than new word pair fragments for both related and unrelated word pairs) if it is assumed that the processing of word pair status (i.e., related or unrelated) can also lead to the formation of a unitized representation in memory. The nature of the link between the words, however, is one of being related or unrelated. In a sense, this would be similar to the links which exist between the members of a synonym word pair and the members of an antonym word pair. Since this hypothesis can account for the data, it is unclear whether we need the semantic unitization hypothesis that sentence generation subjects form new meaningful associations between the members of unrelated word pairs. As a result, the present experiment has direct implications for studies which have examined "implicit memory for new associations" (e.g., Graf & Schacter, 1985, 1987, 1989; Schacter & Graf, 1986, 1989).

Implications for the relationship between encoding time and memory performance

The results of the present experiment also have implications for the relationship between encoding time and memory performance. It could be suggested that the sentence generation subjects showed larger effects of previous exposure (i.e., larger differences between old and new items) than the letter comparison subjects on the word fragment completion test, the semantic decision test and the subjective evaluation test because the sentence generation task is more time consuming, and thus provides more encoding time, than the letter comparison task. In fact, this was the case for the unrelated word pairs. The sentence generation subjects ($M = 8077$, $SD = 2681$) took longer to complete

their study task than the letter comparison subjects ($M = 6402$, $SD = 2164$), $E(1,138) = 16.713$, $MSe = 6044878.314$. This was not the case, however, for the related word pairs. The sentence generation subjects ($M = 5221$, $SD = 1613$) took less time to complete their study task than the letter comparison subjects ($M = 6346$, $SD = 2161$), $E(1,138) = 12.333$, $MSe = 3692967.548$. Therefore, if encoding time had a direct impact upon memory performance in the present experiment, the sentence generation subjects would have shown larger effects of previous exposure than the letter comparison subjects on the unrelated word pairs and the letter comparison subjects would have shown larger effects of previous exposure than the sentence generation subjects on the related word pairs. An examination of Figures 2 through 6, however, clearly indicates that the sentence generation subjects showed larger effects of previous exposure than the letter comparison subjects on both the related and the unrelated word pairs. Thus, the differences observed in the memory performance of sentence generation subjects and letter comparison subjects in the present experiment can not be directly attributed to variations in encoding time.

Implications for the processing account and the memory systems account. Although the present experiment was not designed as a test of the processing account (e.g., Blaxton, 1989; Jacoby, 1983a, 1983b; Kolers & Roediger, 1984; Roediger & Blaxton, 1987a, 1987b; Roediger et al., 1989; Srinivas & Roediger, 1990; Weldon & Roediger, 1987) or the memory systems account (e.g., Squire, 1986, 1987; Tulving 1985, 1987), the results of the present experiment do have implications for these theories of memory. First, the results can easily be accommodated by the processing account. Since this theory of memory focuses on the match in processing operations performed at study and test, it would have predicted that the conceptually-driven study task (i.e., sentence generation) would lead to larger effects of previous exposure (i.e., larger differences between old and

new items) on the conceptually-driven indirect measures of memory (i.e., conceptually-driven word fragment completion, semantic decision and subjective evaluation) than the data-driven study task (i.e., letter comparison). On the other hand, the results can not be easily accommodated by the memory systems account. This theory of memory claims that all of the indirect measures of memory rely on the same memory system (e.g., the semantic or procedural system). Since previous studies have shown that conceptually-driven study tasks do not produce larger effects than data-driven study tasks on the indirect measures of perceptual identification (e.g., Jacoby, 1983a; Jacoby & Dallas, 1981; Light & Singh, 1987) and word stem completion (e.g., Graf et al., 1982; Graf & Mandler, 1984), the memory systems account would not have predicted that sentence generation (i.e., a conceptually-driven study task) would lead to larger effects of previous exposure than letter comparison (i.e., a data-driven study task) on the indirect measures of memory used in the present study (i.e., conceptually-driven word fragment completion, semantic decision and subjective evaluation).

Implications for future studies of human learning and memory. The results of the present experiment have implications for future studies of human learning and memory. The experiment demonstrated that conceptually-driven indirect measures of memory can be designed to index the prior occurrence of specific conceptually-driven mental operations. Although the present experiment examined a fairly intuitive hypothesis about the mental operations which are engaged by a conceptually-driven study task (i.e., the semantic unitization hypothesis), there are many less intuitive hypotheses which can be examined. For example, in one study Nairne et al. (1985) offered a semantic elaboration hypothesis to explain why generating a word according to a letter transposition rule (e.g., EHAVEN) led to better recognition performance than reading (e.g., HEAVEN). Specifically, they suggested that the advantage for generated items occurred because

the generation task activated the semantic attributes of study items to a greater extent than did the reading task. This activation then spreads to associated items, thereby creating a network of potential retrieval routes. This explanation seems counterintuitive. Could flipping the first two letters of a word really engage the mental operation of semantic elaboration? Fortunately, this hypothesis can be tested with a modified version of the semantic decision test that was used in the present experiment. To demonstrate this possibility, consider the following experiment. In the study phase, subjects are presented with a series of words which represent living things (e.g., EAGLE) and non-living things (e.g., TRAIN) and they are asked to study these words under generation conditions (e.g., AEGLE, BTRAIN) and reading conditions (e.g., EAGLE, TRAIN). In the test phase, the studied word pairs are re-presented along with a number of new words which represent living things (e.g., MOUSE) and non-living things (e.g., SCARF). The subject's task is to decide, as quickly as they can, whether the words represent living or non-living things. If generating a word by a letter transposition rule does activate the semantic attributes and associates of an item to a greater extent than reading, then semantic decisions (e.g., "living" or "non-living") should be faster for words that were generated at the time of study than words that were read at the time of study.

Future Research

Although the semantic decision test and the subjective evaluation test could be easily modified to examine many other hypotheses about the mental operations engaged by conceptually-driven study tasks, there are some other avenues of research which directly follow from the results of the present experiment.

Investigating unrelated word pairs in an unmixed list. One possible avenue for future research would be to explore the hypothesis that the sentence generation subjects directly encoded the related and unrelated status of the word pairs in the study phase.

This may have occurred because the related and unrelated word pairs were presented in a mixed list (i.e., within-subjects). In this situation the unrelated status of a word pair may be more salient in a background of related word pairs. Perhaps the predictions that were made for the unrelated word pairs in the present experiment would be supported if the related and unrelated word pairs were presented in an unmixed list (i.e., between-subjects). Under these circumstances subjects may not pay explicit attention to the unrelated nature of the word pairs since all of the word pairs in the list would be unrelated. Previous research has indicated that the results obtained with mixed lists can be very different than the results obtained with unmixed lists. For example, Hunt and Elliot (1980) found that the recall of orthographically distinctive words (e.g., lymph, khaki, afghan) was higher than the recall of orthographically common words (e.g., leaky, kennel, airway) when the words were previously presented in a mixed list. When these same words were presented in unmixed lists, however, Hunt and Elliot (1980) found that the recall of orthographically distinctive words was equivalent to the recall of orthographically common words (see Begg & Snider, 1987 and Slamecka & Katsaiti, 1988 for similar findings with the generation and reading tasks).

Investigating different retention intervals. A second possible avenue for future research would be to examine performance on the semantic decision test and the subjective evaluation test over different retention intervals (e.g., immediately after study, 2 hours after study, 24 hours after study). Studies which have examined the effect of various retention intervals on the data-driven word fragment completion test (e.g., Komatsu & Ohta, 1984; Roediger & Blaxton, 1987a, 1987b; Tulving et al., 1982) and the data-driven perceptual identification test (e.g., Jacoby & Dallas, 1981; Jacoby, 1983b) have revealed that the effects of a recent exposure can last several days or weeks. On the other hand, studies which have examined the effect of various retention intervals on

the data-driven word stem completion test (e.g., Graf & Mandler, 1984; Graf, Squire & Mandler, 1984) have revealed that the effects of a recent exposure can be absent within 2 hours of study. Since the conceptually-driven indirect measures of memory are a recent development, only three studies have examined the effect of different retention intervals on these measures. In one of these studies, Shimamura and Squire (1984) found that the effects of a recent exposure on a conceptually-driven free association test were absent within 2 hours of study. Similarly, Hamman (1990) found that the effects of a previous exposure on a conceptually-driven category association test were absent within 90 minutes of study. On the other hand, Schacter and Graf (1986) found that the effects of a recent presentation on a conceptually-driven word fragment completion test were still present after a 24 hour retention interval. Since the semantic decision test and the subjective evaluation test do not require subjects to produce the study items (i.e., the items are re-presented and the subjects response to them is measured), it is unclear how long these effects might last.

Conclusions

The present study has made two contributions. First, it has provided some evidence against the semantic unitization hypothesis. Specifically, it has been demonstrated that the requirement to generate a meaningful sentence with two unrelated words does not necessarily lead to the formation of a new association between the members of an unrelated word pair. Second, and perhaps more important, this study has demonstrated that conceptually-driven indirect measures of memory can be designed to index the prior occurrence of specific mental operations. I believe that the re-processing methodology used in the present study can be a useful tool for exploring the mental operations which are engaged by conceptually-driven study tasks.

References

- Anderson, J. R., & Reder, L. M. (1979). An elaborative processing explanation of depth of processing. In L. S. Cermak & F. I. M. Craik (Eds.), Levels of processing in human memory (pp. 385-404). Hillsdale, NJ: Erlbaum.
- Anderson, R. C., Goldberg, S. R., & Hidde, J. L. (1971). Meaningful processing of sentences. Journal of Educational Psychology, 62, 395-399.
- Begg, I., & Snider, A. (1987). The generation effect: Evidence for generalized inhibition. Journal of Experimental Psychology: Learning, Memory and Cognition, 13, 553-563.
- Bjorklund, D. F., & Bjorklund, B. R. (1985). Organization versus item effects of an elaborated knowledge base on children's memory. Developmental Psychology, 21, 1120-1131.
- Blaxton, T. A. (1989). Investigating dissociations among memory measures: Support for a transfer-appropriate processing framework. Journal of Experimental Psychology: Learning, Memory and Cognition, 15, 657-668.
- Boyd, B. D., & Ellis, N. R. (1986). Levels of processing and memory in mentally retarded and nonretarded persons. Intelligence, 10, 1-8.
- Brooks, D. N., & Baddeley, A. D. (1976). What can amnesic patients learn? Neuropsychologia, 14, 111-122.
- Burke, D. M., & Light, L. L. (1981). Memory and aging. The role of retrieval processes. Psychological Bulletin, 90, 513-546.
- Cermak, L. S., & Craik, F. I. M. (Eds.). (1979). Levels of processing in human memory. Hillsdale, NJ: Erlbaum.
- Cermak, L. S., Talbot, N., Chandler, K., & Wolbarst, L. R. (1985). The perceptual priming phenomenon in amnesia. Neuropsychologia, 23, 615-622.

- Clarke, R., & Morton, J. (1983). Cross modality facilitation in tachistoscopic word recognition. Quarterly Journal of Experimental Psychology, 35A, 79-96.
- Cohen, N. J., Eichenbaum, H., Deacedo, B., & Corkin, S. (1985). Different memory systems underlying acquisition of procedural and declarative knowledge. Annual of the New York Academy of Science, 44, 54-71.
- Cohen, N. J., & Squire, L. R. (1980). Preserved learning and retention of pattern analyzing skill in amnesics: Dissociation of knowing how and knowing that. Science, 210, 207-210.
- Craik, F. I. M. (1977). Age differences in human memory. In J. E. Birren & K. W. Schaie (Eds.), Handbook of the psychology of aging (pp. 384-420). New York: Van Nostrand Reinhold.
- Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. Journal of Verbal Learning and Verbal Behavior, 11, 671-684.
- 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.
- Donaldson, W., & Bass, M. (1980). Relational information and memory for problem solutions. Journal of Verbal Learning and Verbal Behavior, 19, 26-35.
- Gardiner, J. M., & Rowley, J. M. C. (1984). A generation effect with numbers rather than words. Memory & Cognition, 12, 443-445.
- Glisky, E. L., & Rabinowitz, J. C. (1985). Enhancing the generation effect through repetition of operations. Journal of Experimental Psychology: Learning, Memory and Cognition, 11, 193-205.
- Graf, P. (1990). Life-span changes in implicit and explicit memory. Bulletin of the Psychonomic Society, 28, 353-358.

- Graf, P., & Mandler, G. (1984). Activation makes words more accessible, but not necessarily more retrievable. Journal of Verbal Learning and Verbal Behaviour, 23, 553-568.
- Graf, P., Mandler, G., & Haden, P. (1982). Simulating amnesic symptoms in normal subjects. Science, 218, 1243-1244.
- 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., & Schacter, D. L. (1987). Selective effects of interference on implicit and explicit memory for new associations. Journal of Experimental Psychology: Learning, Memory and Cognition, 13, 45-53.
- Graf, P., & Schacter, D. L. (1989). Unitization and grouping mediate dissociations in memory for new associations. Journal of Experimental Psychology: Learning, Memory and Cognition, 15, 930-940.
- Graf, P., Shimamura, A. P., & Squire, L. R. (1985). Priming across modalities and priming across category levels: Extending the domain of preserved function in amnesia. Journal of Experimental Psychology: Learning, Memory and Cognition, 11, 386-396.
- Graf, P., Squire, L. R., & Mandler, G. (1984). The information that amnesic patients do not forget. Journal of Experimental Psychology: Learning, Memory and Cognition, 10, 164-178.
- Greenbaum, J. L., & Graf, P. (1989). Preschool period development of implicit and explicit remembering. Bulletin of the Psychonomic Society, 27, 417-420

- Hamman, S. B. (1990). Level-of-processing effects in conceptually driven implicit tasks. Journal of Experimental Psychology: Learning, Memory and Cognition, 16, 970-977.
- Howard, D. V. (1988). Implicit and explicit assessment of cognitive aging. In M. L. Howe & C. J. Brainerd (Eds.), Cognitive development in adulthood: Progress in cognitive developmental research (pp. 3-37). New York: Springer-Verlag.
- Hunt, R. R., & Elliot, J. M. (1980). The role of nonsemantic information in memory: Orthographic distinctiveness effects on retention. Journal of Experimental Psychology: General, 109, 49-74.
- Jacoby, L. L. (1978). On interpreting the effects of repetition: Solving a problem versus remembering a solution. Journal of Verbal Learning and Verbal Behavior, 17, 649-667.
- Jacoby, L. L. (1983a). Remembering the data: Analyzing interactive processes in reading. Journal of Verbal Learning and Verbal Behaviour, 22, 485-508.
- Jacoby, L. L. (1983b). Perceptual enhancement: Persistent effects of an experience. Journal of Experimental Psychology: Learning, Memory and Cognition, 9, 21-38.
- Jacoby, L. L. (1984). Incidental versus intentional retrieval: Remembering and awareness as separate issues. In L. R. Squire & N. Butters (Eds.), Neuropsychology of memory (pp. 145-156). New York: Guilford Press.
- Jacoby, L. L., Allan, L. G., Collins, J. C., & Larwill, L. K. (1988). Memory influences subjective experience: Noise judgements. Journal of Experimental Psychology: Learning, Memory and Cognition, 14, 240-247.

- 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.
- Johnson, M. K., & Hasher, L. (1987). Human learning and memory. Annual Review of Psychology, 38, 631-668.
- Johnson, M. K., Kim, J. K., & Risse, G. (1985). Do alcoholic Korsakoff's syndrome patients acquire affective reactions? Journal of Experimental Psychology: Human Learning, Memory and Cognition, 11, 22-36.
- Kinoshita, S. (1989). Generation enhances semantic processing? The role of distinctiveness in the generation effect. Memory & Cognition, 17, 563-571.
- Kolers, P. A. (1975). Memorial consequences of automatized encoding. Journal of Experimental Psychology: Human Learning and Memory, 1, 689-701.
- Kolers, P. A., & Roediger, H. L. (1984). Procedures of mind. Journal of Verbal Learning and Verbal Behavior, 23, 425-449.
- Komatsu, S. I., & Ohta, N. (1984). Priming effects in word-fragment completion for short- and long-term retention intervals. Japanese Psychological Research, 26, 194-200.
- Kucera, M., & Francis, W. (1967). Computational analysis of present day American English. Providence, RI: Brown University Press.
- Leong, C. K., Wong, S., Wong, A., & Hiscock, M. (1985). Differential cerebral involvement in perceiving Chinese characters. Levels of processing approach. Brain and Language, 26, 131-145.

- Light, L. L., & Singh, A. (1987). Implicit and explicit memory in young and older adults. Journal of Experimental Psychology: Learning, Memory and Cognition, 13, 531-541.
- Light, L. L., Singh, A., & Capps, J. L. (1986). Dissociation of memory and awareness in young and older adults. Journal of Clinical and Experimental Neuropsychology, 8, 62-74.
- Lockhart, R. S., & Craik, F. I. M. (1990). Levels of processing: A retrospective commentary on a framework for memory research. Canadian Journal of Psychology, 44, 87-112.
- Malone, M. A., Kershner, J. R., & Siegel, L. (1988). The effects of methylphenidate on levels of processing and laterality in children with attention deficit disorder. Journal of Abnormal Child Psychology, 16, 379-395.
- McElroy, L. A., & Slamecka, N. J. (1982). Memorial consequences of generating nonwords: Implications for semantic-memory interpretations of the generation effect. Journal of Verbal Learning and Verbal Behavior, 21, 249-259.
- Milner, B., Corkin, S., & Teuber, H. L. (1968). Further analysis of the hippocampal amnesia syndrome. Neuropsychologia, 6, 215-234.
- Mitchell, D. B. (1989). How many memory systems? Evidence from aging. Journal of Experimental Psychology: Learning, Memory and Cognition, 15, 31-49.
- Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. Journal of Verbal Learning and Verbal Behavior, 16, 519-533.
- Moscovitch, M. (1982). A neuropsychological approach to perception and memory in normal and pathological aging. In F. I. M. Craik & S. Trehub (Eds.), Aging and cognitive processes (pp. 55-78). New York: Plenum Press.

- Moscovitch, M. (1984). The sufficient conditions for demonstrating preserved memory in amnesia: A task analysis. In L. R. Squire & N. Butters (Eds.) The neuropsychology of memory (pp. 104-114). New York: Guilford.
- Nairne, J. S., Pusey, C., & Widner, R. L., Jr. (1985). Representation in the mental lexicon: Implications for theories of the generation effect. Memory & Cognition, 13, 183-191.
- Owings, R. A., & Baumeister, A. A. (1979). Levels of processing, encoding strategies and memory development. Journal of Experimental Child Psychology, 28, 100-118.
- Parkin, A. J., & Streeter, S. (1988). Implicit and explicit memory in young children and adults. British Journal of Psychology, 79, 361-369.
- Payne, D. G., Neely, J. H., & Burns, D. J. (1986). The generation effect: Further tests of the lexical activation hypothesis. Memory & Cognition, 14, 246-252.
- Rabinowitz, J. C., & Craik, F. I. M. (1986). Specific enhancement effects associated with word generation. Journal of Memory and Language, 25, 226-237.
- Richardson-Klavehn, A., & Bjork, R. A. (1988). Measures of memory. Annual Review of Psychology, 39, 475-543.
- Roediger, H. L., & Blaxton, T. A. (1987a). Effects of varying modality, surface features, and retention interval on priming in word fragment completion. Memory & Cognition, 15, 379-388.
- Roediger, H. L., & Blaxton, T. A. (1987b). Retrieval modes produce dissociations in memory for surface information. In D. S. Gorfein & R. R. Hoffman (Eds.), Memory and learning: The Ebbinghaus Centennial Conference (pp. 349-379). Hillsdale, NJ: Erlbaum.

- Roediger, H. L., Weldon, M. S., & Challis, B. A. (1989). Explaining dissociations between implicit and explicit measures of retention: A processing account. In H. L. Roediger III & F. I. M. Craik (Eds.) Varieties of memory and consciousness: Essays in honor of Endel Tulving (pp. 3-41). Hillsdale, NJ: Erlbaum.
- Sanquist, T. F., Rohrbaugh, J. W., Syndulko, K., & Lindsley, D. B. (1980). Electrocortical signs of levels of processing: Perceptual analysis and recognition memory. Psychophysiology, 17, 568-576.
- Scarborough, D. L., Cortese, C., & Scarborough, H. S. (1977). Frequency and repetition effects in lexical memory. Journal of Experimental Psychology: Human Perception and Performance, 3, 1-17.
- Schacter, D. L. (1987). Implicit memory: History and current status. Journal of Experimental Psychology: Learning, Memory and Cognition, 13, 501-518.
- Schacter, D. L., & Graf, P. (1986). Effects of elaborative processing on implicit and explicit memory for new associations. Journal of Experimental Psychology: Learning, Memory and Cognition, 12, 432-444.
- Schacter, D. L., & Graf, P. (1989). Modality specificity of implicit memory for new associations. Journal of Experimental Psychology: Learning, Memory, and Cognition, 15, 3-12.
- Shimamura, A. P. (1986). Priming effects in amnesia: Evidence for a dissociable memory function. Quarterly Journal of Experimental Psychology, 38A, 619-644.
- Shimamura, A. P., & Squire, L. R. (1984). Paired-associate learning and priming effects in amnesia: A neuropsychological study. Journal of Experimental Psychology: General, 113, 556-570.
- Slamecka, N. J., & Graf, P. (1973). The generation effect: Delineation of a phenomenon. Journal of Experimental Psychology: Human Learning and Memory, 4, 592-604.

- Slamecka, N. J., & Katsaiti, L. T. (1987). The generation effect as an artifact of selective displaced rehearsal. Journal of Memory and Language, 26, 589-607.
- Squire, L. R. (1986). Mechanisms of memory. Science, 232, 1612-1619
- Squire, L. R. (1987). Memory and brain. New York: Oxford University Press
- Squire, L. R., & Cohen, N. J. (1984). Human memory and amnesia. In G. Lynch, J. L. McGaugh & N. Weinberger (Eds.), Neurobiology of learning and memory (pp. 3-64). New York: Guilford.
- Srinivas, K., & Roediger, H. L. (1990). Classifying implicit memory tests: Category association and anagram solution. Journal of Memory and Language, 29, 389-412
- Stein, B. S. (1978). Depth of processing reexamined: The effects of precision of encoding and test appropriateness. Journal of Verbal Learning and Verbal Behavior, 17, 165-174.
- Tulving, E. (1985). How many memory systems are there? American Psychologist, 40, 385-398.
- Tulving, E. (1987). Multiple memory systems and consciousness. Human Neurobiology, 6, 67-80.
- 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
- Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. Psychological Review, 80, 352-373
- Warrington, E. K., & Weiskrantz, L. (1968). New method of testing long-term retention with special reference to amnesic patients. Nature, 217, 972-974
- Warrington, E. K., & Weiskrantz, L. (1970). Amnesic syndrome: Consolidation or retrieval? Nature, 228, 628-630.

- Warrington, E. K., & Weiskrantz, L. (1974). The effect of prior learning on subsequent retention in amnesic patients. Neuropsychologia, 12, 419-428.
- Warrington, E. K., & Weiskrantz, L. (1978). Further analysis of the prior learning effect in amnesic patients. Neuropsychologia, 16, 169-177.
- Watkins, D. (1983). Depth of processing and the quality of learning outcomes. Instructional Science, 12, 49-58.
- Weiskrantz, L., & Warrington, E. K. (1970). A study of forgetting in amnesic patients. Neuropsychologia, 17, 187-194.
- Weldon, M. S., & Roediger, H. L. (1987). Altering retrieval demands reverses the picture superiority effect. Memory & Cognition, 15, 269-280.
- Winnick, W. A., & Daniel, S. A. (1970). Two kinds of response priming in tachistoscopic recognition. Journal of Experimental Psychology, 84, 74-81.

Appendix A

Related-Old

NEEDLE-THREAD	BACON-GREASE	PALM-SWEAT	LADLE-SOUP
RABBIT-CARROT	SKIRT-BLOUSE	GLOVE-SCARF	ONION-TEAR
COFFIN-CORPSE	TAILOR-PANTS	SNAIL-SHELL	POTS-STOVE
GUITAR-SINGER	PUDDLE-BOOT	PONY-STRAW	NEST-CROW
KITTEN-CREAM	BIKE-HELMET	PAIL-SHOVEL	RAKE-LEAF
PADDLE-CANOE	TACK-POSTER	EAGLE-CLIFF	POLE-FLAG

Related-New

MONKEY-BANANA	CEMENT-BRICK	RAZOR-BEARD	SALAD-FORK
HAMMER-CHISEL	SWAMP-INSECT	HAWK-MOUSE	PLUG-DRAIN
MINNOW-HERON	ALTAR-PRIEST	ORGAN-CHOIR	VINE-GRAPE
WIENER-RELISH	TROUT-BROOK	ESKIMO-SEAL	CRAB-CLAW
KNIGHT-CASTLE	COCOON-MOTH	INFANT-CRIB	SOAP-BATH
PURSE-WALLET	VASE-FLOWER	APRON-CHEF	SAIL-MAST

Unrelated-Old

CELERY-MUSKET	ANCHOR-GRILL	DRUM-SHRUB	OLIVE-ATTIC
KETTLE-MARBLE	TIGER-SCOTCH	PICKLE-COMB	TAXI-RANCH
RIBBON-PEANUT	SWORD-PATIO	GOOSE-COUCH	NAIL-CHALK
ALBUM-BARREL	NAPKIN-CURB	CAKE-SCREW	DOCK-BUSH
BUTTON-STEAK	WRIST-SHACK	RAISIN-TENT	ROPE-LIME
VIOLIN-R JIPE	DONKEY-LOCK	DEER-SPADE	LION-HARP

Unrelated-New

POODLE-PASTRY	PUPPET-PLATE	DUCK-SLEEVE	MULE-SHIRT
PILLOW-RADISH	CIGAR-TURKEY	CUBE-SPIDER	MOSS-THIEF
BUCKET-RETINA	NURSE-WHEAT	APPLE-FLUTE	DOVE-BELLY
ELBOW-SPROUT	THUMB-ALLEY	JUICE-LATCH	HOSE-CONE
OYSTER-RATTLE	FIDDLE-WHIP	LAMP-SPEAR	CAPE-DISH
CEREAL-STUMP	CALF-BRANDY	PIPE-MEDAL	TIRE-TRAY

Appendix B

Related-Old

MONKEY-BANANA	CEMENT-BRICK	RAZOR-BEARD	SALAD-FORK
HAMMER-CHISEL	SWAMP-INSECT	HAWK-MOUSE	PLUG-DRAIN
MINNOW-HERON	ALTAR-PRIEST	ORGAN-CHOIR	VINE-GRAPE
WIENER-RELISH	TROUT-BROOK	ESKIMO-SEAL	CRAB-CLAW
KNIGHT-CASTLE	COCOON-MOTH	INFANT-CRIB	SOAP-BATH
PURSE-WALLET	VASE-FLOWER	APRON-CHEF	SAIL-MAST

Related-New

NEEDLE-THREAD	BACON-GREASE	PALM-SWEAT	LADLE-SOUP
RABBIT-CARROT	SKIRT-BLOUSE	GLOVE-SCARF	ONION-TEAR
COFFIN-CORPSE	TAILOR-PANTS	SNAIL-SHELL	POTS-STOVE
GUITAR-SINGER	PUDDLE-BOOT	PONY-STRAW	NEST-CROW
KITTEN-CREAM	BIKE-HELMET	PAIL-SHOVEL	RAKE-LEAF
PADDLE-CANOE	TACK-POSTER	EAGLE-CLIFF	POLE-FLAG

Unrelated-Old

POODLE-PASTRY	PUPPET-PLATE	DUCK-SLEEVE	MULE-SHIRT
PILLOW-RADISH	CIGAR-TURKEY	CUBE-SPIDER	MOSS-THIEF
BUCKET-RETINA	NURSE-WHEAT	APPLE-FLUTE	DOVE-BELLY
ELBOW-SPROUT	THUMB-ALLEY	JUICE-LATCH	HOSE-CONE
OYSTER-RATTLE	FIDDLE-WHIP	LAMP-SPEAR	CAPE-DISH
CEREAL-STUMP	CALF-BRANDY	PIPE-MEDAL	TIRE-TRAY

Unrelated-New

CELERY-MUSKET	ANCHOR-GRILL	DRUM-SHRUB	OLIVE-ATTIC
KETTLE-MARBLE	TIGER-SCOTCH	PICKLE-COMB	TAXI-RANCH
RIBBON-PEANUT	SWORD-PATIO	GOOSE-COUCH	NAIL-CHALK
ALBUM-BARREL	NAPKIN-CURB	CAKE-SCREW	DOCK-BUSH
BUTTON-STEAK	WRIST-SHACK	RAISIN-TENT	ROPE-LIME
VIOLIN-RECIPE	DONKEY-LOCK	DEER-SPADE	LION-HARP

Appendix C

Study Practice Items

BOWL-SPOON
PENNY-BLADE
PIRATE-BEETLE
TOAST-CRUMB

Primacy Fillers

RASH-HOOK
PISTOL-BULLET
JUNGLE-DECK
ROBIN-WORM

Recency Fillers

SADDLE-COLT
CABIN-MEAL
WOMB-EMBRYO
MOVIE BASKET

Test Practice Items

PUZZLE-MIST
HANGER-CLOSET
ARROW-PERCH
PARCEL-STAMP

END

2|5| 0|5| 93|

FIN