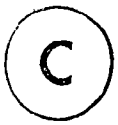
AN INTERPRETIVE FRAMEWORK FOR THE GENERATION EFFECT AND ITS GENERALIZABILITY TO RELATED MEMORY PHENOMENA

By



PETER GRAF, B.Sc.

A Thesis

Submitted to the School of Graduate Studies

in Partial Fulfilment of the Requirements

for the Degree

Doctor of Philosophy

McMaster University

August, 1981

A FRAMEWORK FOR THE GENERATION EFFECT AND RELATED MEMORY PHENOMENA

DOCTOR OF PHILOSOPHY (1981) McMASTER UNIVERSITY
(Psychology) Hamilton, Ontario

TITLE: An Interpretive Framework for the Generation Effect and its Generalizability to Related Memory Phenomena

AUTHOR: Peter Graf, B.Sc. (University of Toronto)

SUPERVISOR: Professor B.A. Levy

NUMBER OF PAGES: x, 180

ABSTRACT

The generation effect refers to the superior retention of verbal materials that were generated by the subjects rather than supplied by the experimenter and just read by the subjects. The major objective of this thesis is to present a new interpretive framework for the generation effect obtained with sentences. This framework emphasizes the distinction between two organizational processes. The first process characterizes the organization among the words of a sentence -- its interword organization. The second process characterizes the organization among the perceptual and conceptual elements within a word -- its intraword organization. The experiments presented highlight the importance of distinguishing between these two types of organization. Results from tests that are primarily sensitive to interword organization showed a generation effect with meaningful but not with anomalous sentences. Results from tests that are primarily sensitive to intraword organization showed a generation effect with both types of sentences. Results from a test that is simultaneously highly sensitive to both of these organizational dimensions revealed both an increase in interword organization along a meaning dimension and an increase in intraword organization.

The generalizability of the current interpretive framework was tested on generation effects obtained with word pairs; it was also examined on the superior retention of sentences read in geometrically transformed as compared to normal typography. The present results cannot readily be assimilated by any of the interpretive views

previously offered for the generation effect. The ability of the proposed framework to accommodate these related findings revealed it as an effective investigative tool that promises to be useful in the examination of memory phenomena related to the generation effect.

ACKNOWLEDGEMENTS

Several people contributed directly or indirectly to the completion of this thesis. While only a few are mentioned, I am deeply thankful to all of them. The support, encouragement, and camaraderie of my friends and colleagues in and out of graduate school made the last four years not just a bearable but a treasurable segment of my life. The unfailing love, confidence, and understanding from my parents and my dearest friend Kathryn Ann Bourgeault were invaluable in sustaining the effort that came to fruition in this thesis. I offer this work to them, with my heartfelt thanks for their love.

I owe a profound debt of gratitude to Betty Ann Levy who chaired the supervisory committee. Betty Ann you were more than my supervisor. You have unselfishly given me friendship, encouragement, and academic guidance during my four years at McMaster. Both your praises and criticisms contributed to my progress. You shared with me many of the trials and triumphs in your life as an academic, as a supervisor, and as a friend. I cannot easily express my appreciation for all you have done for me. My hope and goal is to repay you one day for what I have learned by being an equally considerate and understanding colleague to people who work with me.

My work and education have vastly benefitted from the co-supervisors Lee R. Brooks and Lorraine G. Allan. Lorraine's comments and criticisms on all my work were helpful in shaping this thesis and in clarifying problems of presentation. Lee has constantly encouraged me to think beyond the limits of my immediate research problem and to speculate about its implications for and relationships to other

psychological phenomena. The help of Lorraine and Lee improved the quality of this thesis.

Finally, I thank F.I.M. Craik and N.J. Slamecka who prepared me for graduate school and who assisted me in various ways throughout it. I am also thankful to I. Begg and A.B. Kristofferson who were available for consultation on many occasions.

TABLE OF CONTENTS

TABLE OF CONTENTS (cont'd)

LIST OF TABLES

LIST OF FIGURES

Chapter 1

INTRODUCTION

Verbal materials that have been independently generated by subjects are remembered better than materials that have simply been read. This retention advantage for generated materials has been labelled the generation effect (Slamecka & Graf, 1978). The central aim of this thesis is to examine the beneficial memorial consequences underlying the generation effect. Evidence related to this effect has been reported in both the educational and the psychological literature.

Numerous contemporary investigations related to the generation effect originated from a longstanding controversy in the educational literature. This controversy concerned the pedagogic, motivational, and memorial advantages of discovery learning over reception learning (e.g., Bruner, 1960; Dewey, 1910). Although a survey of the literature reveals no widely shared definition or operationalization of discovery learning, its essential feature is that the "principal content of what is to be learned is not given but must be independently discovered by the learner" (Ausubel, 1963, p. 16). In contrast, in reception learning, "the entire content of what is to be learned is presented to the learner in its final form" (Ausubel, 1963, p. 16).

At the center of this controversy were conflicting views of learning. Proponents of the discovery approach viewed learning as involving the subject's interpretation of task demands, and the active retrieval, reorganization, and reevaluation of stored information to

meet those task demands (Bruner, 1961, 1966; Dewey, 1910). Consequently, the main purpose of the discovery approach was to teach pupils the skills involved in applying their own knowledge in finding solutions to new problems (Bruner, 1961). In contrast, the view of learning underlying the reception approach was of a learner involved in passive, mechanical, and rote memorization (Bruner, 1961, 1966). The repetition of presented materials involved in rote memorization was characterized as leading to the passive cataloguing of supplied ideas and concepts within an existing cognitive structure. Thus, the essence of this educational controversy was whether the learner should be viewed as a "slate", written upon by experience, or whether he or she should be described as an active participant in the accumulation of knowledge.

Similar views of learning are contained in the psychological literature. On the one hand, simple models of learning based on classical conditioning (Pavlov, 1927), or on operant conditioning (Skinner, 1938), parallelled the view of learning assumed to underlie the reception approach. To illustrate, the classical conditioning experiment involves the contiguous presentation of a neutral (conditioned) stimulus with one that elicits a certain response unconditionally. Through repeated presentation of these two stimuli, the neutral stimulus comes to have the power of eliciting the response previously produced by the latter stimulus only. According to such models, learning is portrayed as the relatively automatic outcome of appropriate experimental circumstances. On the other hand, the view of learning assumed to characterize the discovery approach reflects what Dember (1974) labelled the "cognitive revolution" in psychology.

Through this revolution, the learner has come to be modelled as an active self-determining individual who continuously seeks out and selects information from the environment. Learning is represented as consisting of active interpretation based on internalized sources of information, and as involving reorganization and recombination to bring old knowledge into agreement with new facts.

Perhaps inspired by this cognitive revolution in psychology, Ausubel (1963) and other proponents of reception learning (e.g., Cronbach, 1966; Wittrock, 1966) offered a reexamination of the cognitive activities involved in discovery and reception learning. This reexamination focused on Bruner's (1961) earlier argument implying a synonymity between discovery and active learning. Ausubel argued that discovery and cognitive activity are orthogonal dimensions, and that cognitive activity is required for effective learning in both the reception and the discovery approach. To illustrate, he asserted that in reception learning, the cataloguing of a presented concept under existing cognitive structures demands decisions about relevance and relationship, reconciliation between new and old knowledge, and sometimes even translation into a personal frame of reference. Perhaps more importantly, Ausubel also recognized that the cognitive activity involved in reception learning may differ qualitatively from that involved in integrating and reorganizing new information with existing knowledge in order to discover new relationships, as in discovery learning.

Subsequent empirical investigations inspired by this new cognitive perspective focused on the memorial consequences associated

with the different cognitive activities assumed to be involved in discovery and reception learning. Focusing on these cognitive activities marks the transition away from the earlier applied work on discovery learning to the contemporary analytic examination of the intrinsic learning advantages associated with discovering as compared to receiving. For this purpose, contemporary researchers have developed methodological approaches that offer several analytic advantages.

The most significant analytic advantage of the approaches used in recent examinations of discovery and reception learning stems from limiting the source of the learning difference to the cognitive activities involved in discovering and receiving. To illustrate, the assertion that discovery learning is more effective than reception learning permits several hypotheses. At one extreme, the differences between a pupil and a teacher in age, motivation, interest, and general background virtually guarantee that the pupil generates materials about a topic that differ from those supplied by the teacher under reception learning conditions. These material differences may account for the faster learning in the discovery as compared to the reception approach. At the other extreme, the learning difference may stem entirely from the different cognitive activities assumed to be involved in discovering and receiving. While the applied researcher need not be concerned with isolating these or any other sources contributing to improved learning, their isolation is critical for attempts to understand the intrinsic learning advantages associated with the cognitive activities involved in discovery as compared to reception learning. In order to limit the source of the learning difference to

these cognitive activities, recent methodological approaches insure that the materials studied under discovery and reception learning conditions are identical.

Experiments that constrain subjects to study the same materials under discovery and reception learning conditions focus on only one of several sources that may contribute to the superiority of discovery over reception learning. Thus, the results obtained from such studies do not permit generalizations about discovery and reception learning. For the purpose of preventing such unwarranted generalizations, in this thesis, the term generating will be used for discovery, and the term reading for receiving. However, the results from such studies do provide potentially important insights into the memorial effects stemming from the qualitatively different cognitive activities assumed to be involved in generating (discovering) and in reading (receiving).

To date, no well founded interpretive account has been offered for the superior retention of identical materials that were self-generated rather than read. In chapter 2 of this thesis I will argue that the theoretical nature of previous accounts severely limits the profitability of their application to the generation effect in future research. Existing accounts also lack explicit specifications of their implications for the generation effect, making them untestable and largely indiscriminable. The major objective of this thesis is to offer an alternative interpretive framework for the generation effect. In order to accomplish this objective, two questions must be addressed. First, is it possible to account for the generation effect in terms of psychological processes that are already well described in the

literature? Second, do these processes facilitate the formulation of testable predictions that will permit the discrimination of the proposed framework from the extant alternatives? The present series of experiments addresses these questions in order to advance our understanding of the generation effect, and of related memory phenomena.

1.1 Overview of thesis

This thesis is based on an extensive investigation of the generation effect obtained with sentences. It offers (a) an interpretive framework for this effect based on well established psychological processes, (b) empirical investigation of predictions generated from this framework, (c) extension of the framework to generation effects obtained with pairs of words, and (d) further extension to a related memory phenomenon stemming from the reading of sentences in geometrically transformed and normal typography.

Focusing on the generation effect obtained with sentences was required by research constraints. Specifically, earlier investigators (Bobrow & Bower, 1969) suggested that comprehension or an appreciation of the meaningfulness of materials played an important role in the generation effect. Pursuit of this suggestion required working with linguistic units that would be large enough to permit assessment of comprehension, and also to facilitate manipulations of meaningfulness. In addition, given the historical connections between the generation effect and learning in the classroom, the aim was to examine the

memorial effects of generating on materials that would facilitate generalization to classroom materials.

The interpretive framework offered for the generation effect emphasizes the distinction between two organizational processes. The first process characterizes the organization among the words of to-be-remembered materials; it has been labelled interword organization (Mandler, 1979). The second process characterizes the organization among the perceptual and conceptual elements within a word. This type of organization is affected by repeated exposure to a word. Each exposure to a word provides an opportunity for interrelating the visual, phonological, and the semantic features of the word to form a unified memory trace. Mandler (1980) has described this organization as intraword organization, and he suggested that it underlies the experience of word familiarity. The interpretive framework offered for the generation effect postulates that generating as compared to reading a sentence results in increased interword organization along a meaning dimension and in increased intraword organization. The lawful relationships between these organizational processes and performance on various measures of memory have been well documented (Mandler, 1979, 1980; Miller, 1956; Tulving, 1962, 1966).

The involvement of these organizational processes in the generation effect is investigated in two series of experiments, one focusing on interword organization, and the other on intraword organization. The combined evidence from these experiments establishes and strengthens the proposed framework as a whole, permitting the formulation of a strong prediction. This prediction, tested in a

separate series of experiments, concerns performance on a new measure of retention which is sensitive to both interword and intraword organization.

The generalizability of the proposed framework is tested in two ways. First, the framework is applied to the interpretation of generation effects obtained with pairs of related words. Since considerable recent research on the generation effect has focused on words and on pairs of related words, an attempt to formulate a comprehensive account of the generation effect necessitates generalization to this research. Second, the framework is used in the interpretation of the retention superiority typically observed for materials read in geometrically transformed as compared to normal typography (Kolers, 1973, 1975). This retention difference will be referred to as the transformation effect. Although reading in transformed typography does not involve generating verbal materials, it will be argued that the two tasks seem to command similar cognitive interactions with the to-be-remembered materials. Consequently, the application of the framework to the transformation effect represents an attempt to generalize it to a related memory phenomenon.

1.2 Outline of Thesis

This thesis includes seven sections, each of which constitutes a separate chapter. In chapter 2 the theoretical arguments that previous investigators have provided in interpreting the generation effect are reviewed. These arguments contributed significantly to the formulation of the interpretive framework proposed for the generation effect. The

framework is detailed at the end of the chapter. The intent of the chapter is to set the stage for the theoretical and empirical work included in this dissertation.

Chapters 3 and 4 describe a series of experiments that provide an empirical demonstration of the importance of distinguishing between the two organizational processes specified as components of the framework. The third chapter includes six experiments examining the effects of generating and reading on interword organization. The fourth chapter consists of two experiments focusing on intraword organization. Since the general method employed in these experiments is identical, it will be detailed once at the beginning of chapter 3, with only deviations from this general method provided for each experiment.

The fifth chapter includes three experiments that test a prediction derived from the proposed framework as a whole. Although the previous experiments provide strong support for the two organizational processes specified as components of the framework, they have considered these processes in isolation. That is, the previous experiments focused either on interword organization or on intraword organization, but they never considered the effects of generating and reading on both of these organizational dimensions simultaneously. The framework as a whole offers a specific prediction about the combined effects of generating on inter- and intraword organization. The experiments included in this section examined this prediction, using a test that is highly sensitive to both of these organizational dimensions. This test, which was inspired by the work of Humphreys (1976, 1978), and by Rabinowitz,

Mandler, and Barsalou (1977), will be referred to as a varied-context word recognition test.

This varied-context word recognition test is employed in chapters 6 and 7, which represent attempts to generalize the framework. The sixth chapter includes one experiment concerned with the generation effect obtained with pairs of words that are phonemically related (rhyme). The seventh chapter presents two experiments which focus on the transformation effect that involves the superior retention of sentences read in geometrically transformed rather than normal typography. These two sets of experiments indicate that the proposed framework provides a promising investigative tool not only for the generation effect, but also for related memory phenomena.

The final chapter will summarize the empirical work and discuss theoretical conclusions. It will consider other extensions of the framework, indicate unresolved issues, and outline directions for future research.

Chapter 2

REVIEW OF EMPIRICAL AND THEORETICAL WORK AND AN INTERPRETIVE FRAMEWORK

This chapter sets the stage for the theoretical and empirical work included in this dissertation. Numerous experiments have compared memory for materials that were subject-generated with memory for materials that were just read. They employed different paradigms and different verbal materials. Various theoretical arguments have been proposed in interpretation of the findings. Since these arguments contributed significantly to the formulation of an alternative interpretive framework for the generation effect, this review will focus on the existing interpretive views.

2.1 Review of empirical work

The central focus of this thesis is confined to the intrinsic learning advantages of generating as compared to reading. Within those confines, two criteria must be met to demonstrate a generation effect: first, retention performance on self-generated materials must exceed performance on materials that were just read, and second the materials studied in the Generate and in the Read condition must not differ in any way that could in itself account for the observed retention difference. The first criterion is self evident; the second requires elaboration. It is well documented that some verbal materials are better remembered than others. To give a few examples, meaningful sentences are learned faster than random lists of grammatically ordered words (Marks & Miller,

1964; Schulman, 1974), words are learned faster than nonsense syllables (Noble, 1952), concrete words are learned faster than abstract words (Borkowski & Eisner, 1968; Paivio, Yuille, & Madigan, 1968), and personally relevant information is acquired more readily than nonrelevant information (Keenan, MacWhinney, & Mayhew, 1977; Rogers, Kuiper, & Kirker, 1977). Consequently, if the materials studied in the Generate and Read conditions differ in any potentially interpretable respect, this difference may in itself partially or totally account for the observed difference in retention performance. In other words, the difference due to the study materials would be completely confounded with the difference due to generating and reading. In order to limit the sources that contribute to the generation effect to the cognitive activities involved in generating and in reading, it is easiest to present identical materials for study in the Generate and Read condition.

The confounding of materials between the Generate and Read conditions presents a major problem in the interpretation of the early research. The impetus for this early research was evidence that randomly paired words were easier to remember as a pair if they were linked to each other by another word or short phrase (Montague, Adams, & Kiess, 1966; Adams & Montague, 1967). The effectiveness of such mediator words or phrases was extensively documented. In addition, it was also noted that adults learned pairs of words more quickly if they were required to independently generate a sentence containing the two words of a pair, as compared to when they just read such a sentence supplied by the experimenter (Bobrow & Bower, 1969; Bower & Winzenz,

1970). However, in these studies, there was no control over the mediator sentences that the subjects generated, and as a result, these sentences probably differed from those supplied by the experimenter. Therefore, the observed retention levels may reflect either a difference in memorability for the generated and supplied mediator sentences, a difference in retention due to generating and reading, or a combination of both. In short, in these studies differences in the study materials were completely confounded with differences due to study conditions, making interpretation impossible.

With a growing awareness of this type of confounding, investigators turned to various approaches designed to insure that exactly the same materials were studied in the Generate and Read condition. Among these, perhaps the most promising, albeit deceptive, was the yoking procedure. The essential element in a yoking procedure is that the materials generated by one subject are supplied to another subject in the Read condition. By yoking subjects in this manner exactly the same materials can be studied in both the Generate and the Read condition. Both Pelton (1969) and Bruner (1961) relied on such yoking procedures to insure comparability of materials studied in the Generate and Read condition. Bobrow and Bower (1969) used a more indirect yoking approach. Searching through their records of the sentences that subjects had generated, they identified a fair number that were identical to sentences that had been supplied to other subjects in the Read condition. A comparison of retention of these sentences revealed better performance in the Generate condition than in the Read condition. In summary, relying on some form of a yoking

procedure, many investigators reported differences in retention that they attributed to the difference between generating and reading. These researchers were confident that the observed generation effects were not confounded with material differences.

However, the yoking procedure introduces a confounding that is subtle and considerably more difficult to identify than a direct material difference. This difference stems from the selection of idiosyncratically effective mediators. To illustrate, the subjects who are allowed to generate sentences containing two words may construct sentences that are subjectively meaningful, vivid, and closely tied to personal experience. Such sentences are highly memorable (Rogers et al., 1977; Raye, Johnson, & Taylor, 1980), but uniquely memorable to the person generating them. For a yoked subject who is given these sentences, but who does not share the same experiential background as the person who generated them, the same sentences may be considerably more difficult to remember. Thus, even when nominally identical items are studied in the Generate and Read condition, the difference in retention performance may still be confounded with idiosyncratic material differences.

A superiority in retention between the Generate and Read condition can be attributed to a difference in the cognitive activities involved in generating and reading only when both actual material differences and idiosyncratic material differences are avoided. Meeting this requirement demands that the materials that will be generated can be predicted. With such predictability, exactly the same materials can be prescribed for study in both the Generate and the Read condition, and

they will not be idiosyncratically more memorable in either condition. The literature presents two approaches that have been used for achieving high predictability of materials to be generated. The first approach developed from the work of Bobrow and Bower (1969), and focused on the effects of generation on words within sentences. The second approach was pioneered by Schwartz (1971) and by Schwartz and Walsh (1974) and examined generation effects with word pairs. These two approaches are reviewed separately.

Generating words within sentences. Anderson and his colleagues (Anderson, 1970; Anderson, Goldberg, & Hidde, 1971; Anderson & Kulhavy, 1972; Anderson, Royer, Kulhavy, Thornburg, & Klemt, 1971) employed a generation manipulation that greatly increased the likelihood that the materials supplied in one condition would be exactly the same as those produced in the Generate condition. Previously, Anderson, Royer, Kulhavy, Thornburg, and Klemt (1971) obtained some evidence suggesting that paired-associate learning was facilitated when the pair (e.g., SIG-YELLOW) was presented in a thematic sentence. In paired-associate learning, subjects are learning to give the response term (e.g., YELLOW) when the stimulus term (e.g., SIG) is presented. The facilitated learning reported by these authors occurred only when the thematic sentence containing the words was first briefly exposed with a blank inserted in place of the response term of the pair (e.g., Before turning red, traffic SIG are _ _ _ _ _ _). This outcome led Anderson, Goldberg, and Hidde (1971) to present sentences constructed such that the last word was determined or constrained by the context formed by the rest of

the sentences (e.g., Mothers bake delicious apple _ _ _). Since all subjects completed these sentences in the expected way, the superior retention reported for words that were generated, rather than read, provided a convincing demonstration of the beneficial memorial consequences accompanying the activity of generating.

Supportive results have also been reported by Kane and Anderson (1978) who introduced an interesting procedural variation. These authors provided subjects with incomplete sentences in the Generate condition, some of which were highly constrained (e.g., The nurse lost the notes from the clip _ _ _ _ _), and others that allowed more than one completion (e..g, Most people like coffee with _ _ _ _ _). For the imperfectly constrained sentences, Kane and Anderson arbitrarily designated one response as correct (e.g., SUGAR) and the other(s) as incorrect (e.g., CREAM). During study, when a subject produced an incorrect completion, the experimenter offered the correct one to the subject who simply repeated it. Consequently, when an incorrect completion was given, subjects were not required to generate the correct completion but simply to speak it aloud just as in the Read condition. Thus, these authors were able to compare not only retention of words that were read with words that were correctly generated, but also with retention of words that were incorrectly generated, as well as with retention of experimenter-supplied corrected words.

The results from a later retention test revealed that the words from correctly completed sentences were remembered better than the words that were only read. However, performance on the incorrectly generated words and on the experimenter-supplied corrected words also exceeded the

level of retention observed in the Read condition. These surprising results suggest that explicit generation of an item may not be necessary for that item to show a generation effect at recall; it may be sufficient that an attempt was made to generate that word. However, whether or not attempted generation is sufficient for the appearance of a generation effect remains to be demonstrated, since the procedure employed by Kane and Anderson (1971) introduced idiosyncratic item selection effects. Specifically, subjects who completed the weakly constrained sentences probably responded with completion words that were personally appropriate or relevant. Such idiosyncratically appropriate words are well remembered (Rogers et al., 1977; Taylor et al., 1980). The experimenter-supplied corrected words may have enjoyed a retention advantage by virtue of being closely related to these idiosyncractically effective words, rather than by being related to words that were self-generated.

Nevertheless, as will be further illustrated below, the act of generating seems to confer a retention advantage not only on the item that is produced but also on the material that guides or constrains the production. For example, Anderson, Goldberg, and Hidde (1971) required subjects to reproduce the sentences that had been used to constrain the production of words. Subjects were asked to reproduce the constraining sentence (e.g., Mothers bake delicious apple _ _ _) when given the completion word (e.g., PIE) as a cue for recall. The authors reported a generation effect not only on the words that were used to complete the sentences, but when given the completion words as recall cues, the constraining sentences also showed a generation effect.

In summary, the research on the generation of words within sentences offers (a) a convenient paradigm to insure comparability of materials in the Generate and Read condition, (b) a convincing demonstration of the beneficial memorial consequences of generating over reading, and (c) an indication that the activity of generating confers a memorial advantage over materials that are not actually generated. This last observation raises questions about the appropriateness of the label "generation effect." However, since the phrase "generation effect" serves merely as a label for the memorial effects associated with generation, whether they are a direct or indirect consequence of it, its applicability remains defensible.

In concluding this section, it is perhaps worth pointing out that the work on the generation of words within sentences has an interesting parallel in research examining the effectiveness of self-instructional programs that require filling in blanks. A single frame from such a program bears a strong resemblance to the stimuli used by Anderson and his colleagues. Yet in the light of the obvious resemblance between the activity of generating and that of filling in the blanks in the frame of a self-instructional program, it is rather surprising that students who write answers to the frames in a program often learn no more than students who read the frames with the blanks already filled in (e.g., Anderson, Faust, & Roderick, 1968; Della-Piana, 1962; Stolurow & Walker, 1962). A reexamination of these programs, however, revealed that many instructional programs systematically oversupply the learner with the requested information (Kemp & Holland, 1966); the frames to be filled in present the required

information in a prominent place to the learner who simply inserts it in the available blanks. Kemp and Holland (1966) selected numerous self-instructional programs that did not permit such systematic filling in of answers, but that required the learner to interpret and to rearrange the material supplied in order to use it in the production of an answer. These programs conferred substantial learning advantages. It is worth noting that these task requirements resemble the manipulations used in the experiments described in this thesis.

Generating word pairs. A different approach to the problem of separating the beneficial memorial effects of generating from the effects due to material and idiosyncratic item selection was taken by Schwartz (1971) and by Schwartz and Walsh (1974). Schwartz's subjects worked with letter-word pairs (e.g., A-PIE; D-CAT). Subjects in the Read condition were given complete pairs to read (e.g., APPLE-PIE; DOG-CAT), while the Generate condition required completion of each pair with a word that (a) began with the supplied letter and (b) was commonly associated with the other word. With this procedure, 86 per cent of the words were identical in the Generate and in the Read condition. Schwartz (1971) reported a small but significant recall advantage in the Generate as compared to the Read condition.

Even higher predictability of materials studied in the Generate condition was achieved by Schwartz and Walsh (1974), but then the advantage due to generating all but disappeared. Subjects in the Read condition were supplied with two letter bi-grams linked to one another with a single mediator-letter (e.g., CR - o - WD; EV - e - NT). The

generate subjects had to produce their own mediator-letter to link the two bi-grams (e.g., CR - _ - WD; EV - _ - NT); they were instructed to choose mediators such that the linked bi-grams would form words. With this paradigm, the authors reported achieving 94 per cent predictability of generated mediators in one experiment and 98 per cent in another, but the usual retention advantage found with generating as compared to reading disappeared. Therefore, Schwartz and Walsh (1974) were forced to conclude that the typical generation effects reported by Bobrow and Bower (1969), Bower and Winzenz (1970), Pelton (1969), and by Schwartz (1971), were obtained because subjects in the Read condition were unfairly disadvantaged. They explained that the read subjects were in some instances prevented from using mediators that might have been idiosyncratically more effective. In short, they concluded that there was no retention benefit associated with the activity of generating as compared to reading.

However, the results reported by Schwartz and Walsh (1974) may have stemmed from a procedural peculiarity that prevented the appearance of a generation effect. Unlike other experimenters, Schwartz and Walsh instructed subjects in one experiment (Experiment 1) to spell out the words supplied in the Read condition and those constructed in the Generate condition, and they limited the written recall of 10 words to 45 seconds in another experiment (Experiment 2). Although it is not immediately apparent how the spelling task could have prevented the appearance of a generation effect, the speeded recall test may simply not have given subjects sufficient time to reproduce all the words that they remembered. Consequently, while these authors pioneered a

promising paradigm for achieving comparability of materials studied in the Read and Generate condition, their procedures differed considerably from previous investigations, making comparisons of results difficult.

Slamecka and Graf (1978) combined a variant of the paradigm developed by Schwartz and Walsh (1974) with the traditional study procedures used in paired-associate learning, and observed substantial generation effects. These authors provided subjects with a rule for relating two words, such as rhyme or opposite, the left-hand member of a word pair and the first letter of the right-hand member (e.g., HALL-B _ _ _; YOUNG-O _ _). Subjects in the Generate condition were instructed to complete each pair using a word that was related to the left-hand member according to the specified relational rule; the to-be-generated word also had to begin with the initial letter provided for the right-hand member. For example, using the relational rule rhyme, the left-hand word HALL, and the initial letter B, the word BALL was generated, and then the complete pair spoken aloud. Subjects in the Read condition were also informed of the relational rule, yet their task was simply to read the supplied words of each pair aloud. With the production constraints imposed by this paradigm, for some experiments, over 99 per cent of the words generated were also supplied by the experimenter in the Read condition. Slamecka and Graf reported that with a variety of relational rules (rhymes, synonyms, opposites, category, associates), designs (within and between subjects), study procedures (paced and unpaced), and test procedures (recognition, free recall, and cued recall), there was a consistent large memorial

advantage associated with having generated, rather than only read, the right-hand member of the word pair.

Similar results were reported by Jacoby (1978) who required subjects to insert missing letters in the right-hand member of associated word pairs. To examine whether the difficulty of the generation task influenced retention performance, Jacoby supplied word pairs that had either two missing letters (e.g., FOOT-S _ _ E), or only one (e.g., CHEQUE-M_NEY). Memory performance was not affected by this manipulation. These results were replicated by Donaldson and Bass (1980) who also asked subjects to insert a single letter in the right-hand member of word pairs (e.g., TIGER-STRIP_S). However, these authors made the generation task even more trivial by (a) insuring that the missing letter was always an E, (b) informing their subjects of this condition, and (c) always underlining the corresponding supplied letter in the Read condition (e.g., TIGER-STRIPES). Nevertheless, there was a substantial memory advantage due to generating as compared to reading that was not significantly smaller than that observed with a more demanding generation task.

In summary, the research examining the generation effect with word pairs provides (a) a convenient and flexible paradigm, and (b) a convincing demonstration of the robustness of the generation phenomenon over a wide variety of materials and conditions. In addition, the results reported by Jacoby (1978) and by Donaldson and Bass (1980) imply that the magnitude of the memorial benefit of generating as compared to reading is not closely related to the difficulty of the processing task.

Evidence from related phenomena. Numerous investigators have employed procedures and reported results that bear a close resemblance to work on the generation phenomenon. To illustrate, Gardiner, Craik, and Bleasdale (1973) examined whether the difficult initial retrieval of a word conferred a subsequent recall advantage upon it. They presented subjects with definitions of common and rare words and measured the time required to retrieve (generate) these words from memory. Subsequent recall performance for words that were difficult to retrieve (requiring more than 15 seconds) was better than recall of words that were easy to retrieve (requiring less than 15 seconds). A similar demonstration of the effects of the difficulty of an initial task on later retention was offered by Auble and Franks (1978). They presented subjects with ambiguous sentences (e.g., The notes were sour because the seam split) together with disambiguating clues (e.g., bagpipe). Auble and Franks varied the time interval between the presentation of the sentence and the disambiguating clue, and instructed subjects to attempt a meaningful interpretation of the sentence during the interval. Retention was higher when the clue followed the sentence after a time interval than when the sentence and the clue were presented closely together. The subjects' effort toward comprehending the sentences was positively related to recall.

In the light of the findings indicating that the difficulty of an initial task is positively related to retention performance (Auble & Franks, 1978; Gardiner et al., 1973), it is perhaps surprising that neither Jacoby (1978) nor Donaldson and Bass (1980) observed a similar relationship between the difficulty of their generation tasks and

recall. However, since neither Jacoby nor Donaldson and Bass offer an independent index of the difficulty of their generation tasks, the relationship between their data and those of Auble and Franks (1978) and Gardiner et al. (1973) remains to be determined.

Also bearing a close resemblance to work on the generation effect is the research of Anderson and Kulhavy (1972) on the learning of words from definitions. These authors demonstrated that the learning of a word-meaning is facilitated if the learner is forced to apply the word in a sentence (generate a sentence), as compared to when such a sentence is simply read by the subject. However, a similar procedure yielded slightly ambiguous results in the work of Erdelyi, Buschke, and Finkelstein (1977). Erdelyi et al. offered subjects riddles that permitted a single word solution. A comparison of retention performance between subjects who generated the solutions and those who simply read them showed differences only after the initial test trials. The absence of a generation effect on the initial test trials remains unexplained.

More consistent patterns of results resembling the generation effect are obtained by researchers who have followed up another aspect of the work reported by Bobrow and Bower (1969). In one experiment, Bobrow and Bower supplied subjects with sentences (e.g., The farmer discovered a diamond) and instructed them to produce a sensible continuation for each (e.g., He sold it to a jeweler and used the money to buy a tractor). Retention of the supplied sentences was considerably better in the sentence continuation condition than in a condition that required repeating each study sentence three times. These results have been replicated by Griffith (1976) and by Masson and Sala (1978).

Mistler-Lachman (1974) added an interesting variation by demonstrating comparable memorial benefits when subjects were required to generate (think of) a meaningful relationship between two supplied sentences and when they generated a continuation sentence. The results obtained in these sentence continuation experiments, and by Mistler-Lachman, are particularly interesting because they demonstrate most clearly the indirect beneficial memorial consequences conferred upon material that constrains the generation of a response.

The indirect beneficial memorial consequences of generating have been harnessed for educational purposes by Wittrock and his colleagues (e.g., Wittrock, 1974, 1977; Wittrock & Carter, 1975; Doctorow, Wittrock, & Marks, 1978). To illustrate, Doctorow et al. (1978) reported that requiring students to generate sentences about story paragraphs after reading each resulted in substantially better comprehension than instructions to just read the story. In other experiments with children, generating sentences to summarize paragraph-meanings nearly doubled reading comprehension of commercially available reading materials (Marks, 1975, cited in Doctorow et al., 1978). Finally, Rickards and August (1975) showed beneficial memorial consequences for students who were instructed to find and underline the most important sentence in each paragraph of a text. These students achieved higher comprehension scores than their colleagues who read a text with experimenter underlined sentences. It is noteworthy, however, that instructions to underline the least important sentence of each paragraph produced a decrease in comprehension, when compared to the experimenter supplied group.

Finally, this section on evidence related to the generation effect would be incomplete without mention of Kolers' extensive work (e.g., Kolers, 1973, 1975, 1976; Kolers & Ostry, 1974) on the transformation effect. The transformation effect concerns the superior retention of sentences read in geometrically transformed as compared to normal typography. In a typical experiment, Kolers instructs subjects to read aloud two sets of sentences, a read or study set, and a recognize set. Each set contains some sentences in normal typography and some in geometrically transformed typography (e.g., upside down or mirror transformed). The recognize set includes all of the sentences contained in the read set as well as some new ones. While reading the sentences in the recognize set, subjects must classify each sentence either as totally new (not included in the read set), as an old sentence in the same typography as on its appearance in the read set, or as an old sentence in a new typography. Kolers has reported that sentences were recognized substantially better if they were initially read in transformed rather than in normal typography. Furthermore, subjects retained information about the meaning of these sentences, as well as about their appearance, for a period of more than a year (Kolers, 1976).

Although an analysis of the superficial demands involved in the generation of a sentence and in reading one in transformed typography reveals no similarities, the latter task also seems to include some generation of verbal materials. Specifically, a sentence displayed in a transformed typography, such as mirror transformed, represents an unfamiliar stimulus that it is difficult to read. Instead of reading such a sentence letter by letter, or word by word, a subject may read

only some of its parts and try to generate (guess at) the rest of the sentence; he or she may generate words that could follow from what was read and verify whether these generated words are similar to those shown in the transformed sentence. Based on this argument, the transformation effect seems to bear a strong resemblance to the generation effect.

Summary of empirical work. It is apparent that the generation effect is not an isolated phenomenon with closely delimited boundary conditions. The literature contains numerous related paradigms that require some degree of generation of the materials to be remembered. Generation effects are also reported for materials that are used to guide or to constrain the generation of verbal materials. It is advisable to be cautious in interpreting many of the findings from studies related to the generation phenomenon since the effects of material and idiosyncratic item selection are often hopelessly confounded among the various experimental conditions. The main reason for describing this related research was to demonstrate the extensiveness of work on this effect. In the light of this related work, it seems apparent that in developing an interpretation for the generation phenomenon, its generalizability to this related work should be a consideration. The next section will describe existing interpretive views and examine how they generalize to this related work.

2.2 Existing interpretive views

Several interpretive views have been advanced in explanation of the generation phenomenon. This section outlines these views and examines how well they can accommodate both the evidence available for the generation effect and from related work, and the additional evidence that is presented in this thesis. To anticipate this additional evidence, several experiments included in this thesis show that generating as compared to reading a sentence has two effects. First, it increases the memorability of the words of the sentence. Second, it also seems to foster the forming of meaning based relationships among these words.

The most prominent views that have been offered in interpretation of the generation effect are: (a) the comprehension view, (b) the levels of processing view, (c) the distinctiveness view, and (d) the retrieval practice view. These views are mainly exploratory frameworks that have been successfully applied to other memory phenomena. However, the profitability of their application to the generation effect is severely limited since they do not distinguish between effects of generating as compared to reading on the memorability of individual words and effects on the relationships among the words. Consequently, the evidence offered in this thesis showing effects of generating as compared to reading a sentence at both of these loci was not anticipated and cannot readily be accommmodated by the existing interpretive views. The alternative interpretive framework detailed at the end of this chapter anticipated both effects of generating as compared to reading.

Comprehension view. The oldest and perhaps most popular interpretation offered in explanation of the beneficial memorial effects of generating as compared to reading was articulated by Bobrow and Bower (1969). These authors offered the intuitively appealing observation that retention is a positive function of comprehension. In addition, they suggested that generating as compared to reading sentences insures more reliable comprehension. In short, generating produces better comprehension which in turn facilitates retention performance. Anderson (1970) supported this notion by arguing that the requirement to generate a sentence demands "full processing from the learner" (p. 364). He stated that learning is facilitated because the generate task requires meaningful processing of the study materials whereas a person "can read ... without bringing to mind the meaning of the words he is speaking" (p. 364). Numerous authors adhere to this view (e.g., Griffith, 1976; Mistler-Lachman, 1974; Wittrock, 1974; Doctorow et al., 1978).

However, the comprehension view is not well articulated and it is not clear what data could be mustered against it. This lack of specificity makes the comprehension view able to accommodate almost any available data, but eliminates it as a serious interpretive candidate for the generation effect, because it is largely untestable. It is difficult to imagine how one would examine notions of comprehension (Anderson, 1972) particularly using the materials that are often employed in generation experiments (words and word pairs). A more significant obstacle is to separate the direct effects of generating on comprehension from indirect comprehension test effects mediated by

better memory for the study materials. That is, the increased comprehension may stem from better memory for the generated materials.

Levels of processing view. The comprehension view has obvious parallels with the initial conceptualization of the levels of processing framework (Craik, 1973; Craik & Lockhart, 1972; Craik & Tulving, 1975). This framework portrayed memory as a by-product of the perceptual and conceptual activities that were required for the initial analysis of to-be-remembered materials. The original levels of processing framework consisted of a continuum of processing levels or domains (Lockhart, Craik, & Jacoby, 1976). One end of the continuum was marked by shallow processing of the superficial aspects of the to-be-remembered materials, such as graphemic and phonemic features, while the other end was delineated by deep processes involving the analysis of the semantic properties of the study materials. The basic postulate of the framework was that a deeper, more meaningful analysis of the study materials leads to more durable memory traces than an analysis of the sound and other superficial properties of the materials. In short, it was proposed that retention performance was a positive function of the depth to which the study materials had been processed.

The levels of processing framework was not presented as a theory of memory that is amenable to hypothesis testing (Craik & Lockhart, 1972; Lockhart & Craik, 1978). The intent was to provide a heuristic for describing and for advancing understanding of the relationships between the cognitive activities assumed to be engaged in during study of to-be-remembered materials and subsequent retention performance.

Although this framework has been criticized on several grounds (e.g., Eysenck, 1978; Nelson, 1977; Postman, 1975), there is a multitude of research supporting the proposed relation between various processing tasks and retention (e.g., Craik, 1979; Cermack & Craik, 1979).

Since the levels of processing framework was advanced as a heuristic for characterizing the relationships between various processing tasks and retention performance, the central issue is whether this view offers an adequate description of the evidence available for the generation effect. The adequacy of this description has been seriously questioned by three recent studies and by the additional evidence presented in this thesis. To illustrate, Mistler-Lachman (1974) compared retention performance in a condition where subjects generated continuation sentences, to a condition where they judged the meaningfulness of the study sentences. Performance in the continuation condition was substantially better than performance in the semantic judgement condition. McFarland, Frey, and Rhodes (1980) replicated this finding reporting that subjects who generated a word to fit a specified context remembered more than subjects who decided whether the experimenter-supplied word fit the same context. The original levels of processing framework offers no basis for predicting and for describing memory differences between two study tasks that require semantic processing. Since the present studies indicate that generating is more beneficial to memory performance than a deep semantic processing task, it appears that generating does not just require more semantic processing than reading, it seems to require different processing.

An additional shortcoming of a levels of processing description for the generation effect is revealed by the data reported by Slamecka and Graf (1978). These authors used various relational rules to constrain the generation of words. Some rules were closer to the shallow end of the processing continuum of levels of processing (such as rhyme) while others were closer to the deep end (such as synonym). Since these diverse generation rules probably required different levels of processing, it seems to follow that the magnitude of the generation effect should have varied with the depth of processing required by the rules. However, the results reported by Slamecka and Graf show no evidence supporting such an argument. These results also indicate that generating is not just deeper processing than reading; it's memorial benefits appear to stem from other sources.

More important, however, the levels of processing framework does not afford an adequate description of the findings reported in this thesis. These findings reveal that generating as compared to reading a sentence has two effects, one on the memorability of the words of the sentence and one on the formation of relationships among these words. The levels of processing framework offers no basis for anticipating the possibility of effects at more than one locus. Consequently, the profitability of its application to the generation effect in future research remains questionable.

Distinctiveness view. Contemporary reports (e.g., Donaldson & Bass, 1980; McFarland et al., 1980; Jacoby, 1978; Slamecka & Graf, 1978) have favoured an alternative interpretive view of the generation effect.

This view emerged from recent revisions and amplifications of the original levels of processing framework (e.g., Craik, 1977; Jacoby, 1974; Lockhart & Craik, 1978; Moscovitch & Craik, 1976). The original framework emphasized the meaningfulness of particular processing activities in determining the retention of a word, but it treated meaning as a fixed entity that was assumed to have been either completely encoded during study or not encoded at all. The notion of distinctiveness was advanced in order to overcome the difficulties resulting from treating meaning as a fixed entity.

Several authors (Garner, 1974; Harris, Begg, & Upfold, 1980; Olson, 1970) have pointed out that the description of an object depends on what the object must be discriminated from, and that the meaning of a word is determined by the distinctions that it conveys in a certain context. Recognition of the importance of context and task requirements for determining the description or meaning of an object is central to the notion of distinctiveness (Craik & Jacoby, 1979). According to this notion memory is a by-product of the perceptual and conceptual activities required for the initial analysis of to-be-remembered materials (e.g., Craik & Lockhart, 1972; Craik & Tulving, 1975). The phenomenal experience of memory occurs when the rememberer reengages in activities performed previously (Kolers, 1973, 1975; Moscovitch & Craik, 1972; Restle, 1974). There is disagreement about the exact psychological processes that are assumed to underlie the experience of recognition. According to some authors (Kolers, 1973, 1975; Restle, 1974), the initial performance of an activity facilitates its subsequent performance, and it is this facilitation that underlies memory. Other

authors (Bower, 1967; Craik & Lockhart, 1972; Moscovitch & Craik, 1976) assume that subjects retain a record of the perceptual and conceptual operations involved in analyzing an item when it is presented for study as well as when it appears on the test. Recognition occurs when the study record matches the test record. Inspite of this important disagreement, there is a consensus that memory depends critically on the quality and quantity of the cognitive activities involved in perception and comprehension that distinctively characterize or describe the to-be-remembered material.

Kolers offered such a view of memory in interpretation of the memorial consequences of skilled and unskilled reading. He suggested that the unskilled reader retains a richer and more distinctive memory from an encounter with a sentence than the skilled reader (Kolers, 1975; Kolers & Ostry, 1974). The skilled reader knows from experience where to attend and where not to attend. He or she may have to focus on a few critical words; the rest of the sentence can be constructed from knowledge gathered in previous encounters with text (Kolers, 1975). The unskilled reader does not know where to focus, nor has he or she learned to separate the important from the irrelevant information. In addition, the whole sentence may have to be analyzed word by word, or even letter by letter, because there is less knowledge accumulated from previous encounters with text that can be relied upon in constructing the words or parts of the sentence from memory. Thus, more relevant and irrelevant analyzing operations are involved in unskilled reading than in skilled reading. As a consequence, when the same sentence is

encountered after unskilled reading, there is a greater chance that the same mental operations will be reinstated.

Perhaps encouraged by the success of the notion of distinctiveness in describing the memorial consequences of skilled and unskilled reading (Kolers, 1975; Kolers & Ostry, 1974) as well as other memory phenomena (e.g., Eysenck, 1979; Jacoby, Craik, & Begg, 1979), two subtly different distinctiveness views have been offered in interpretation of the generation effect. The first view was presented by Slamecka and Graf (1978) and Donaldson and Bass (1980). They specified that the relationship between the words of a pair may be more distinctively encoded in the Generate than in the Read condition, since reading does not "effectively demand any registration of that relation" (Slamecka & Graf, 1978, p. 603). But results reported by McFarland et al. (1980) seem to argue against this position. These authors instructed subjects to complete each of a series of sentences with a word that did not fit the rest of the sentence. It is unclear whether or how subjects attempted to relate these completion words to the sentences. Nevertheless, these unfitting completion words were remembered as well as, or even better than the appropriate completion words. Thus, McFarland et al. argued that distinctive "item-context integration must be rejected as an explanation for the generation effect" (1980, p. 222). However, the results reported by these authors are hopelessly confounded with material and idiosyncratic item selection effects, since the experimenters had no control over the words that subjects generated. Thus, the notion of a more distinctive encoding of the relationship between the words of a pair in the Generate than in the

Read condition still offers a possible, albeit incomplete, description of the generation effect. Its incompletenesss stems from focusing on a single locus of the effects of generating as compared to reading, while this thesis presents evidence that generating has effects at two loci.

A subtly different distinctiveness account was espoused by McFarland et al. (1980) and by Jacoby (1978). These authors emphasized the distinctiveness of the activity of generating. They suggested that while reading is a highly practiced skill, most subjects are novices at the task of generating. The well practiced execution of the routine activities involved in skilled reading leaves few opportunities for reinstating the same activities later. In contrast, the slow and unpracticed execution of the novel mental activities required in generating verbal materials includes a unique and highly distinctive set of activities, allowing more opportunities for recognition. This account also focuses on a single locus of the effects of generating as compared to reading.

In summary, distinctiveness interpretations of the generation phenomenon (a) receive some support from the work of Donaldson and Bass (1980), (b) are supported by numerous studies concerned with extensions of the levels of processing framework (e.g., Eysenck, 1979; Jacoby, Craik, & Begg, 1979), and (c) they successfully account for the memorial consequences of unskilled reading (e.g., Kolers, 1973, 1975; Kolers & Ostry, 1974). However, the profitability of a distinctiveness interpretation of the generation effect is limited since it focuses attention on a single locus of the effect. Evidence presented in this thesis shows that generating as compared to reading a sentence affects

both the memorability of the words of the sentence and the relationships among these words. The notion of distinctiveness does not afford an adequate description of these findings.

Generating as retrieval practice. A related but substantially different interpretation of the generation effect evolves when one compares generating to retrieval from semantic memory (e.g., Slamecka & Graf, 1978). Generation of a word may be viewed as the initial retrieval of that word from semantic memory which facilitates later recall of the same word. Since all materials are given to the subject in the Read condition, reading does not require retrieval of these materials from semantic memory. There is substantial evidence attesting to the memorial advantages of an initial recall test on a subsequent recall test (e.g., Bjork, 1975; Darely & Murdock, 1971; Brown, 1976), and on recognition tests (e.g., Rabinowitz, Mandler, & Patterson, 1977). These supporting results present this view of the generation effect as a possible albeit incomplete alternative. Its incompleteness stems also from limiting the effects of generating as compared to reading to a single locus.

In addition, as Slamecka and Graf (1978) have noted, "this type of explanation also comes close to being only a restatement of the very finding it seeks to explain, that is, that a generated word is better remembered than one that was read because it was generated (recalled)" (p. 603). McFarland et al. (1980) offer a way out of this circular argument suggesting that "retrieval modifies an item's representation in memory according to the way in which retrieval is conducted" (p. 222).

Furthermore, they observe that a difficult initial retrieval may require more decisions and more unsuccessful attempts than an easy retrieval (Gardiner et al., 1973; Jacoby et al., 1979), resulting in a more distinctive memory trace. Generating may involve a more difficult retrieval act than does reading.

An alternative possibility suggested by the observation that retrieval modifies an item's representation according to how retrieval was accomplished (McFarland et al., 1980) is indicated by the work of Morris, Bransford, and Franks (1977). The idea is that there is a greater similarity between the processes involved in generating and the usual tests used in the assessment of retention, than between reading and those tests. Morris et al. (1977) have demonstrated that the degree of overlap between the study and test conditions is an important determinant of performance.

In summary, the retrieval practice account of the generation effect remains a viable but incomplete alternative. The view does require a more explicit restatement which may, however, reveal it as indistinguishable from the distinctiveness account. This view and the distinctiveness view lack the generalizability that characterizes the comprehension notion. To illustrate, while completing a sentence with a word probably requires retrieval of that word from semantic memory, it is not readily apparent why, in these views, the constraining (but experimenter supplied) sentence should enjoy a memorial benefit comparable to that of the target word.

Other views. In addition to the interpretive notions of the generation effect just described, it is possible to entertain a number of more quantitative accounts. To illustrate, it may be argued that generating is a more effortful task than reading (e.g., Auble & Franks, 1978; Griffith, 1976), and that retention is a function of the effortfulness of the processing task. However while the amount of effort expended on a task may provide a useful correlate of cognitive activity under some conditions (see Auble and Franks, 1978; Kahneman, 1973), it does not advance understanding of the generation effect. It is the mental activities that require the expending of effort that determine retention performance (Walsh & Jenkins, 1973). Other quantitative accounts include the total time hypothesis, strength theory, and perhaps frequency theory. These three views postulate differential opportunities for studying the materials in the Generate and Read condition. However, since recent research on the generation effect carefully equated study opportunities in the two conditions, these alternatives can be rejected (Slamecka & Graf, 1978).

Summary of interpretive views. Various interpretive accounts have been offered. These accounts are intuitively appealing, and they have been successfully employed in the interpretation of other retention phenomena. These views can accommodate most of the evidence currently available for the generation effect. They vary in terms of how well they generalize to phenomena related to the generation effect. In addition, given their descriptive nature, the existing interpretive accounts do not permit the formulation of specific predictions, and the

design of experiments for testing these predictions. Thus, there is currently little direct empirical evidence that discriminates among the available alternatives. More important, however, this thesis presents new evidence that generating as compared to reading a sentence has an effect both on the memorability of the words of the sentence and on the relationships among these words. The existing views cannot readily account for the effects of generating as compared to reading at both of these loci. Consequently, the profitability of their application to the generation effect in future research is questionable, and the generation phenomenon remains without a well founded interpretive account (Slamecka & Graf, 1978).

2.3 Alternative interpretive framework

The theoretical arguments advanced in interpretation of the generation effect by previous investigators contributed significantly to the formulation of the present framework. In particular, the comprehension view articulated by Bobrow and Bower (1969) suggested an interpretation based on organizational processes (Mandler, 1967, 1979; Miller, 1956). The basic postulates of the comprehension view are first, that generating as compared to reading results in better comprehension of the study materials, and second, that increased comprehension mediates increased retention performance. Comprehension of a sentence implies, among other things, that the meaning or meanings of the words of the sentences are appreciated, and that these words are related to one another in a well-structured unit. Appreciating the meaning of a word may involve relating it to previous experiences, and

it may require translation of the word in terms of such experiences (Ausubel, 1963). The relationships among the words in a sentence can be used to identify distinctive previous experiences stored in long term memory, and they serve to modify such experiences. The meaning of a sentence is understood when a distinctive and unified representation of the sentence has been formulated (Ausubel, 1963; Wittrock, Marks, & Doctorow, 1975).

The comprehension view of Bobrow and Bower (1969) implies the involvement of two organizational processes in the generation effect: interword organization and intraword organization. Interword organization refers to the organization among the words of a sentence (Mandler, 1979). The notion of comprehending the meaning of a sentence implies that the words of the sentence form an integrated, well-organized, informational unit. The argument that generating, as compared to reading, leads to better comprehension suggests increased interword organization in the Generate as compared to the Read condition. Numerous authors (e.g., Mandler, 1967, 1979; Tulving, 1962, 1966) have documented a positive relationship between increased interword organization and retention performance. Thus, the generation effect may in part be attributable to increased interword organization in the Generate as compared to the Read condition.

One implication of this organizational process is that the magnitude of the generation effect should be influenced by variables that affect the comprehensibility of a sentence. The comprehensibility of a sentence is a function of its meaningfulness. The meaningfulness of a sentence depends critically on the relationships among the words in

a sentence. Meaningfulness may be defined by the potential for directing the reader to construct or reconstruct a particular mental representation for an event, emotional state, etc. (Ausubel, 1963). A list of randomly selected words arranged in a grammatical structure may be defined as an anomalous sentence. Such a sentence does not typically permit the construction of a unified mental representation; it remains incomprehensible. The absence of meaningful relationships among the words of an anomalous sentence prevents, or at least hinders substantially, the forming of a strong interword organization among the words of the sentence. Consequently, on the argument that generating, as compared to reading, increases interword organization, this increase in interword organization is expected to be larger for meaningful sentences than for anomalous sentences.

The second organizational process implicated by the comprehension view of Bobrow and Bower (1969) concerns the organization among the various elements of each word. This type of organization has been labelled intraword organization (Mandler, 1979). The elements of a word are, for example, its various meanings in different contexts, its diverse appearances, its letters, and the separate features of these letters (Mandler, 1979; 1980). An encounter with a word provides an opportunity for relating some of its elements into a unified memory trace. In an encounter with a highly familiar word, new relationships may be discovered, or old ones rediscovered and strengthened (Mandler, 1980). Most importantly, intraword organization involves only the word by itself, and not its relationship to other words.

The comprehension view (Bobrow & Bower, 1969) suggests the possibility that there is an increase in intraword organization in the Generate as compared to the Read condition. Since comprehension involves an appreciation of the meaning of each component word of a sentence, it will affect intraword organization. The argument that generating leads to better comprehension than reading implies more intraword organization in the Generate than in the Read condition. Intraword organization is concerned only with organization among the elements within a word, and thus it does not depend on whether the words of a sentence are meaningfully related or randomly put together. Consequently, one implication of this organizational process is that the size of the generation effect should not depend on the meaningfulness of the sentences read and generated.

In summary, this framework based on organizational processes suggests that generating, as compared to reading, results in increased interword organization, and that this increase in interword organization would be larger for meaningful sentences than for anomalous sentences. The second process suggests that generating as compared to reading results in increased intraword organization, whether or not the study sentences are meaningful or anomalous. The literature contains tests that are highly sensitive either to interword organization or to intraword organization (Mandler, 1979). Based on the proposed framework, on tests sensitive to interword organization, the size of the generation effect is expected to interact with the meaningfulness of the study sentences. On tests that are sensitive to intraword organization, both the meaningful and the anomalous sentences are expected to show

similar generation effects. These expectations are examined in chapters 3 and 4.

2.4 Summary

The generation effect is a robust phenomenon obtained with a variety of paradigms, materials, and test procedures. It has been extensively studied by psychologists as well as educators. Numerous theoretical arguments have previously been offered in its interpretation. However, these arguments concentrate on the effects of generating as compared to reading on a single locus. This thesis presents evidence that generating a sentence affects both the memorability of the words of the sentences and the relationships among these words. Existing interpretive notions cannot readily accommodate these findings, and consequently, the generation effect remains without a well-founded interpretive account. The present chapter detailed an interpetive framework that focuses on the effects of generating as compared to reading both on interword organization and on intraword organization. The implications of this framework for the generation effect have been outlined. The experiments included in the next two chapters will examine these implications.

Chapter 3

GENERAL METHOD AND EMPIRICAL WORK ON INTERWORD ORGANIZATION

The framework proposed in interpretation of the generation effect focuses on both interword organization and on intraword organization. This chapter consists of six experiments examining the effects of generating and reading on interword organization.[1] The framework suggests that there is an increase in interword organization in the Generate as compared to the Read condition. Furthermore, assuming that this increase in interword organization is based on a meaning dimension, it is expected to be larger when meaningful as compared to anomalous sentences are studied.

This general expectation was examined in the six experiments described in this chapter. In these experiments subjects read and self-generated both meaningful and anomalous sentences before they were tested. The main purpose of the first three experiments was to establish the generation effect with these sentences and to test it over a range of performance levels with cued recall tests based on two types of cues. The fourth experiment used a cued recall test to examine the generation effect as a function of study trials. The last two experiments investigated further implications of interword organization with different types of tests. Since the general method used in these experiments and in the experiments described in chapter 4, is identical it will be detailed once here with only deviations from this general method provided for each experiment.

3.1 General method

Materials. The same set of materials, or a subset thereof, was used in all experiments. Each experiment consisted of presenting meaningful and anomalous six-word sentences on the Cathode Ray Tube (CRT) of a laboratory computer. The grammar of these sentences was: Article (the), Adjective, Noun, Verb (-ed), Article (the), Noun. The nouns were common concrete words. The verb was always in the past tense. The basic set of materials was composed of 60 meaningful and 60 anomalous sentences that were obtained from this grammar. (These sentences are reproduced in Appendix A.) The anomalous sentences consisted of a random rearrangement of the content words of the meaningful sentences, with the constraint that no meaningful combinations of the words were formed. These two sets of sentences were used for study in the Read condition.

The same 60 meaningful and anomalous sentences were also used in the Generate condition. In this condition, the CRT displayed a list of the content words of one of these sentences. The words were arranged randomly within each list but the subject-noun of each sentence was marked by a dash indicating its role. In short, there were four sets of 60 sentences: a Meaningful-Read set, a Meaningful-Generate set, an Anomalous-Read set, and an Anomalous-Generate set. From each set, ten sentences were used for practice and two as fillers. The remaining 48 sentences formed the critical set. Table 3.1 lists an example of a Read and a Generate sentence for both the meaningful and the anomalous material. All material was displayed in capital letters.

TABLE 3.1

An example of the study material for each condition

Study material	Processing condition	
	Read	Generate
Meaningful	THE BLOND GIRL BAKED THE CAKE	CAKE BLOND BAKED -GIRL
Anomalous	THE BLOND LEAFLET FILLED THE PILOT	PILOT -LEAFLET FILLED BLOND

Some experiments required only a subset of the 48 critical sentences. To prevent confounding of specific words with the meaningfulness of the sentences in these experiments, the same words were used in constructing each subset of 16 meaningful and anomalous sentences.

Subjects and Design. The subjects were undergraduate volunteers of both sexes who received course credit for participating. Participation in an experiment depended on a subject's ability to generate the sentences according to instructions[2]. The failure to follow instructions would introduce a difference in the materials studied in the Generate and Read condition. Thus the restriction on participation was necessary to maintain counterbalancing of materials studied across the two processing conditions. The basic design included the presentation format or the processing condition (read or generate) as a within-subjects factor and the study material (meaningful or anomalous) as a between-subjects factor.

Procedure. The subjects were tested individually. First, they were shown the grammar that was used to generate all sentences (on an index card). Then, in the Read condition, a sentence appeared on the CRT for eight seconds and the subject was instructed to "simply read the sentence aloud, loudly and clearly, exactly once without errors". In the Generate condition the CRT showed a list of the content words of a sentence for eight seconds. Subjects were shown how to use the specified grammar to generate a sentence with these words. It was

explained that the marked word was to become the subject-noun of the sentence, and that the articles had to be THE's. The instructions stressed that each sentence was to be generated in the head ("Think before speaking!") before saying it "out loud, loudly and clearly, exactly once without errors." The importance of errorless performance was emphasized. Subjects were told to expect a memory test for the sentences, although the exact nature of the test was not specified.

During practice, a subject was shown five read and five generate sentences, with the material (meaningful or anomalous) depending on the experimental condition. The presentation format alternated from sentence to sentence with each format starting the sequence equally often across subjects. The practice set was repeated until a subject performed without errors on all sentences, and until he or she felt comfortable with the task.

To minimize interference from the practice list on the critical list, subjects were engaged in an unrelated activity for three minutes prior to the learning of the sentences from the critical list. The critical list was composed of blocks of sentences presented in the Read format or in the Generate format. Each block consisted of four sentences. The order of presentation of Read and Generate blocks was alternated with each format starting the sequence equally often across subjects. Two filler sentences were appended to the critical list, one of which was shown in each presentation format. The purpose of the fillers was to prevent subjects from maintaining and retrieving the last few study items from short term memory. The study material was

counterbalanced in such a manner that the identical nominal lexical items were shown in all conditions of each experiment.

In summary, the general procedure consisted of presenting either meaningful or anomalous sentences in the Read and in the Generate condition and having the subjects say them aloud. Requiring subjects to say the sentences out loud permitted the experimenter to monitor task performance. The goal of these experiments was to examine the expectation, based on the notion of interword organization that generating, as compared to reading, results in increased interword organization. On the additional assumption that this increased interword organization is based on a meaning dimension, the magnitude of the generation effect was expected to be larger with the meaningful sentences than with the anomalous sentences. Thus, on tests which are sensitive to interword organization the size of the generation effect was expected to interact with the meaningfulness of the sentences.

3.2 Experiment 1

In this experiment a cued-recall test was used to assess retention of the critical sentences. In this test, the verb from each sentence was provided as a cue for recalling the rest of the sentence. The selection of the verb as a retrieval cue was based on the central role played by the verb in defining the relationship among the other words in the sentence (Fillmore, 1968; Raeburn, 1979). The effectiveness of a cue for recall has been formally expressed in the encoding specificity principle (Tulving & Thomson, 1973). According to that principle, a recall cue can facilitate recall only to the extent

that it was integrated and stored in memory together with the rest of the sentence during study. A word that was well integrated, as compared to one that was poorly integrated, provides an effective cue for retrieval of the rest of the sentence (Mandler, 1979; Tulving & Thomson, 1973). A verb should be a powerful retrieval cue for a sentence with high interword organization. Performance on this cued recall test, then, was expected to show an interaction of the generation effect with the meaningfulness of the study sentences due to a larger generation effect for the meaningful sentences than for the anomalous sentences.

Method. The method was as described above. Twenty-four subjects participated in this experiment. They were randomly assigned to the Meaningful and the Anomalous condition until there were exactly 12 in each group. The experiment required only 16 of the 48 critical sentences, eight of which were shown in each presentation format.

Following study of the sentences, subjects received a six-minute cued recall test. It consisted of a random listing of the verbs of the critical sentences. Subjects were instructed to write, next to each verb, the sentence or any parts thereof that had contained each verb.

Results and Discussion. Throughout this thesis the following conventions will be observed. First, even though subjects were encouraged to recall complete sentences, the data that are presented are exclusively from the recall of the subject and object nouns of these sentences, these being scored independently. Focusing on the recall of the nouns permits a comparison of performance on identical materials

across a variety of tests of retention. The data were also scored for complete sentence recall and the same pattern of results emerged. Second, the data are presented as recall percentages to facilitate comparisons among experiments, although all analyses were conducted on the number of words correctly recalled. Finally, any statistical effect that reached the .05 level was considered significant.

In scoring the recall protocols a noun was considered correctly recalled if it was identical to one of the sentence words and if it was recalled in response to its correct verb-cue. When the protocols were scored on a more lenient criterion that did not consider where a word was written on the recall protocol, the pattern of results remained unchanged. The data, displayed in Table 3.2, show the expected generation effect on the meaningful sentences and no effect on the anomalous sentences.

The data were submitted to a two-factor mixed design analysis of variance, treating processing condition (generate and read) as a within-subjects factor and study material (meaningful and anomalous) as a between-subjects factor. The results of the analysis are presented in Appendix B. The analysis revealed a significant main effect for material, $F(1,22) = 36.18$, $MSe = 7.22$, and for processing condition, $F(1,22) = 4.89$, $MSe = 3.83$, as well as a significant interaction of these factors, $F(1,22) = 12.53$, $MSe = 3.83$. The effect due to material, which demonstrates that meaningful sentences are easier to learn than anomalous sentences, is as expected from previous research (e.g., Marks & Miller, 1964; Miller & Selfridge, 1950). The interaction was clarified by Simple Main Effects analyses (Kirk, 1968) that confirmed a

TABLE 3.2

Recall of nouns as a function of processing condition and material in Experiment 1 (using verbs as recall cues)

Study material	Processing condition			
	Generate		Read	
	M(%)	SD(%)	M(%)	SD(%)
Meaningful	52.08	14.86	31.77	22.19
Anomalous	10.42	6.19	15.10	10.50

generation effect for the meaningful sentences, $F(1,22) = 16.55$, $MSe = 1.92$, but not for the anomalous sentences, $F < 1$.

Although these findings demonstrate the expected interaction of the processing condition with the meaningfulness of the study sentences, they raise two questions. First, is the observed pattern of results peculiar to having provided the verbs of the sentences as cues for recall? The verb of a sentence is acknowledged to serve a central organizing function in specifying the relationships among its words (Fillmore, 1968; Raeburn, 1979). This unique property of the verbs may be crucial in obtaining the present pattern of results. This possibility was investigated in Experiment 2, where the adjective phrase (Article, Adjective, Noun) of each sentence was provided as a cue for recall. Second, is the obtained pattern of results a function of the different levels of recall obtained with the meaningful and the anomalous sentences? The low level of recall on the anomalous sentences may have prevented the appearance of a generation effect in that condition. This possibility was investigated in Experiment 3 where all critical sentences were studied three times before the cued recall test was administered.

3.3 Experiment 2

Method. Twenty-four subjects participated in this experiment. They were randomly assigned to the Meaningful and the Anomalous condition until there were exactly 12 in each group. Sixteen of the critical sentences were used, eight of which were shown in each presentation format.

Following study of the critical sentences subjects received a six-minute cued recall test. It consisted of a random listing of the adjective phrases (Article, Adjective, Noun) of the critical sentences. Subjects were instructed to write, next to each adjective phrase cue, the sentence or any parts thereof that had contained the cue.

Results and Discussion. The recall protocols were scored as in Experiment 1. Since the cue provided for recall included the subject noun of each sentence, the data included in Table 3.3 reflect the recall of the object nouns. The table shows the expected generation effect on the meaningful sentences; for the anomalous sentences, recall in the Read condition is better than recall in the Generate condition. These data were submitted to an analysis of variance identical to that employed in Experiment 1 (see Appendix B).

The analysis again revealed a significant main effect for material, $F(1,22) = 37.17$, $MSe = 1.76$, and for processing condition, $F(1,22) = 7.24$, $MSe = 1.15$, as well as a significant interaction of these factors, $F(1,22) = 35.03$, $MSe = 1.15$. The interaction was further investigated by Simple Main Effects analyses (Kirk, 1968) that showed a substantial generation effect for the meaningful sentences, $F(1,22) = 34.04$, $MSe = 1.15$. However, for the Anomalous condition, recall was better in the Read condition than in the Generate condition, $F(1,22) = 5.21$, $MSe = 1.15$.

The overall similarity in the results obtained in Experiments 1 and 2 establishes the generality of the observed pattern of results with different recall cues. With both types of cues, there was a generation

TABLE 3.3

Recall of object nouns as a function of processing condition and material in Experiment 2 (using adjective phrases as recall cues)

Study material	Processing condition			
	Generate		Read	
	M(%)	SD(%)	M(%)	SD(%)
Meaningful	68.75	13.59	35.41	14.91
Anomalous	16.67	14.44	29.16	17.14

effect when meaningful sentences were studied but there was no generation effect with the anomalous sentences. This pattern of results supports the notion that generating as compared to reading leads to increased interword organization along a meaning dimension.

The superior recall of the anomalous sentences in the Read condition as compared to the Generate condition of Experiment 2 is surprising. While the present framework anticipated the absence of a generation effect with these materials, it offers no theoretical arguments from which to expect the observed result. It is possible to offer a post hoc interpretation for this finding, attributing it either to the differential effectiveness of the cues used in Experiments 1 and 2, to the fact that performance levels in these experiments were noticeably different, or to the difference in functional study time in the Generate and Read condition (in the Generate condition some of the available time is taken up by arranging the displayed words into a grammatical frame). These possibilities are not pursued here since they were not central to the topic of this thesis. Moreover, speculation seemed unwarranted since this finding was not replicated in any of the other experiments included in this thesis. Overall, Experiment 2 replicates Experiment 1 and demonstrates that the observed interaction generalizes over different retrieval cues.

3.4 Experiment 3

This experiment examined the possibility that the overall pattern of results obtained in Experiments 1 and 2 was due to the different levels of performance obtained with the meaningful and

anomalous sentences. In particular, in Experiment 1, the low level of performance on the anomalous sentences may obscure a large difference in recall between the Generate and Read conditions. To explore this possibility, in this experiment all critical study sentences were shown three times before the cued recall test was administered.

Method. Twenty-four subjects participated in this experiment. They were randomly assigned to the Meaningful and the Anomalous condition until there were exactly 12 in each group. The study set consisted of 16 sentences, eight of which were shown in each presentation format. Each subject received three consecutive exposures to the critical study set (no pause between exposures) before cued recall was tested, using the verbs as recall cues, as in Experiment 1.

Results and Discussion. The recall protocols were scored as in Experiment 1. The data are summarized in Table 3.4. The table shows the expected generation effect on the meaningful sentences, and no difference between the Generate and Read condition on the anomalous sentences. These data were submitted to an analysis of variance identical to that used in Experiment 1 (see Appendix B).

The analysis showed a significant main effect of material, $F(1,22) = 33.76$, $MSe = 9.04$, as well as a main effect for processing condition, $F(1,22) = 17.9$, $MSe = 4.62$, and it supported the expected interaction between these two factors, $F(1,22) = 9.13$, $MSe = 4.62$. Simple Main Effects analyses (Kirk, 1968) revealed that the difference between the Generate and Read condition was significant for the

TABLE 3.4

Recall of nouns as a function of processing condition and material in Experiment 3 (using three study trials)

Study material	Processing condition			
	Generate		Read	
	M(%)	SD(%)	M(%)	SD(%)
Meaningful	81.80	11.75	53.60	19.30
Anomalous	38.54	17.40	33.85	15.87

meaningful sentences, $F(1,22) = 26.3$, $MSe = 4.62$, but not for the anomalous sentences, $F < 1$.

This pattern of results replicates and extends to a multiple exposure situation the findings of Experiment 1. The essentially unchanged pattern of results with substantially increased performance levels attests to the robustness of the generation effect. A comparison of these two sets of data allows an examination of the effects of study trials on the recall of the meaningful and anomalous sentences. Since with meaningful sentences, generating resulted in better retention performance than reading after one study trial, additional trials might be expected to increase the size of the generation effect. A t-test comparing the difference in recall between the Generate and Read condition of Experiments 1 (one study trial) and 3 (three study trials) showed a marginal increase in the size of the effect, $t(22) = 1.579$ ($p < .07$). While this marginal increase suggests that in the Meaningful condition generating continues to benefit the learner more than reading throughout the experiment, it does not provide a clearcut answer to what happens to the size of the effect over study trials. In addition, a comparison of the pattern of results from the Anomalous condition of Experiments 1 and 3 indicates that recall performance has benefitted more from additional study trials in the Generate than in the Read condition, suggesting a trend toward a generation effect. In short, neither the data from the meaningful sentences nor those from the anomalous sentences show clearly the effects of the processing conditions as a function of study trials. Experiment 4 was designed to

explore these effects further. An additional purpose of Experiment 4 was to examine the influence of multiple tests on these effects.

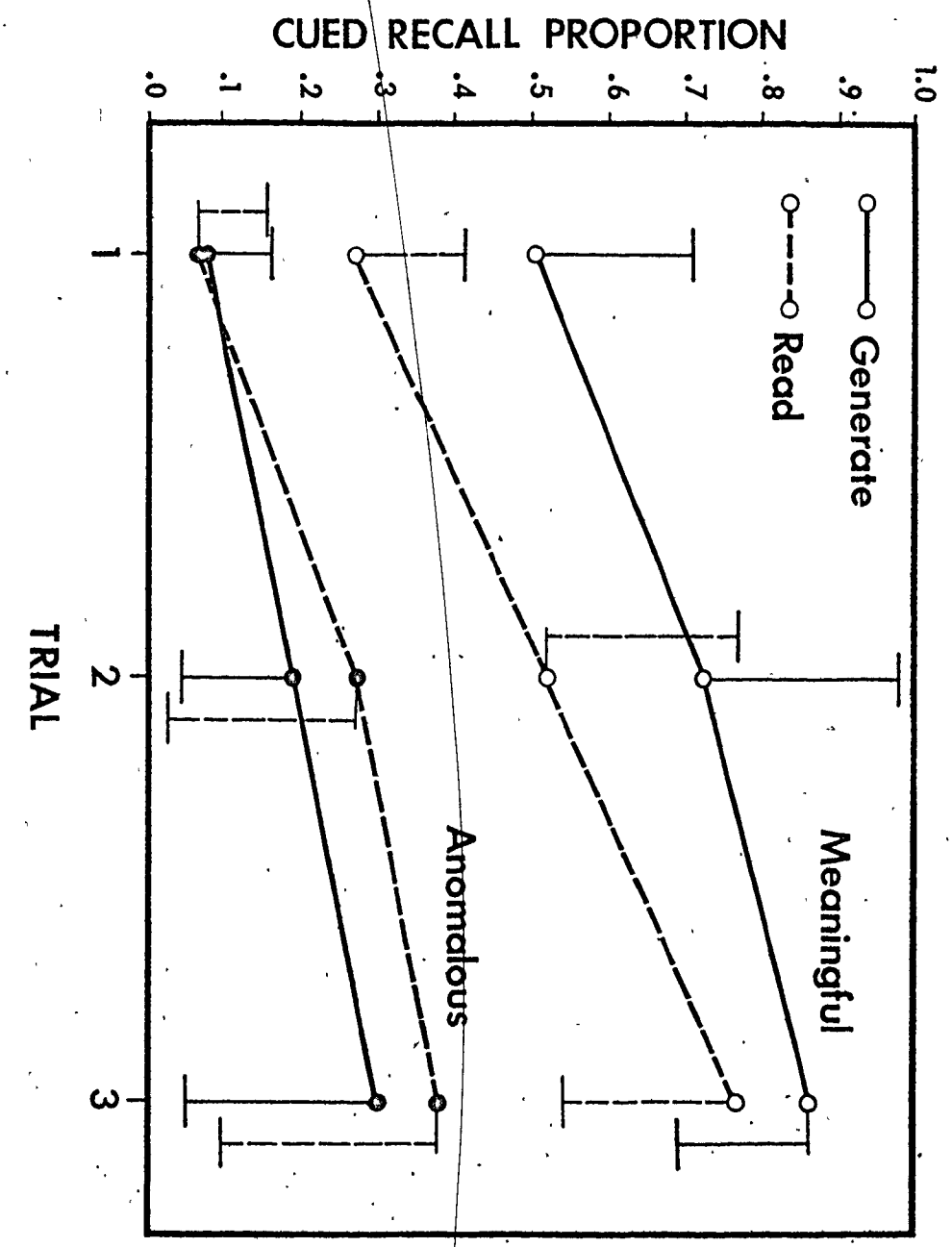
3.5 Experiment 4

Method. Twenty-four subjects participated in this experiment. They were randomly assigned to the Meaningful and the Anomalous condition until there were exactly 12 in each group. The experiment was identical to Experiment 1, except that there were three study-test trials in this experiment as compared to the single study-test trial in Experiment 1; it differed from Experiment 3 by having a test trial after each study trial.

Results and Discussion. The recall protocols were scored as in Experiment 1. The data are presented in Figure 3.1. The figure shows, (a) better recall in the Meaningful than in the Anomalous condition, (b) increased recall performance as a function of study trials, (c) a generation effect in the Meaningful condition but none in the Anomalous condition, and (d) the persistence of the generation effect over study trials in the Meaningful condition. These data were submitted to a three-factor mixed design analysis of variance, treating processing condition (generate and read) and trials as within-subjects factors and materials (meaningful and anomalous) as a between-subjects factor (see Appendix B).

The analysis revealed a significant main effect for material, $F(1,22) = 36.09$, $MSe = 39.31$, and for trials, $F(2,44) = 69.55$, $MSe = 5.34$, and interaction effects for material with trials, $F(2,44) = 3.51$,

Figure 3.1. Recall of nouns (means and standard deviations) as a function of processing condition, study-test trials, and study material in Experiment 4



MSe = 5.34, for material with processing condition, F(1,22) = 11.86, MSe = 9.89, and for trials with processing condition, F(2,44) = 4.21, MSe = 2.41, with no other significant effects.

The interaction of material with trials was due to the faster acquisition of meaningful sentences, as compared to anomalous sentences, over the three learning trials. The interaction of trials with processing condition occurred because recall performance, combined over both material types, showed a greater total increase over trials for the Read condition than for the Generate condition. This latter finding is partially due to the ceiling effect that is evident in the Meaningful-Generate condition on Trial 3. The interaction of material with processing condition was clarified by Simple Main Effects analyses (Kirk, 1968). These analyses showed a significant generation effect with meaningful material, F(1,22) = 14.32, MSe = 9.89, but no systematic differences with anomalous material, F(1,22) = 1.18.

The results of this experiment agree with those of Experiments 1, 2, and 3 in all essential respects. The data from the anomalous sentences showed no evidence for a generation effect over a considerable range of performance levels. In addition, the data from the meaningful sentences showed that the generation effect persisted at least over the number of trials used in this experiment, only to diminish when recall neared ceiling (in the Meaningful-Generate condition). This persistence of the generation effect replicates and extends the findings of Slamecka and Graf (1978) who worked with a different paradigm and different materials. At the same time the present findings seem to argue that generating benefits the learner only on the initial study test trial,

but not on subsequent trials, since the rate of learning for Generate and Read sentences appeared to be identical. If generating continued to benefit the learner more than reading, even on subsequent trials, one would expect a divergence of the learning curves for the Generate and the Read condition. The interesting question is why there is no such divergence in the present data particularly since a comparison of the data of Experiments 1 and 3 did show a marginal increase in the magnitude of the generation effect as a function of study trials. One interpretation suggested by these two sets of opposing data is that the test experience available in a multiple study-test trial experiment allows a subject to monitor recall performance in the Read and in the Generate conditions. This test experience may alert the subject to the ineffectiveness of reading as a learning strategy and it may then motivate him or her to try "harder" in the Read condition. This notion was pursued in an experiment. Since that experiment is not primarily concerned with the role of interword organization in the generation effect, it is not included in the body of this thesis, but in Appendix C instead.

The two major findings that have emerged from these four experiments are (a) the consistent generation effect in the Meaningful condition, and (b) the consistent absence of a generation effect in the Anomalous condition. This pattern of results demonstrates the expected interaction of the processing condition with the meaningfulness of the study material. These findings may be interpreted as evidence that generating as compared to reading results in an increase in the

interword organization of sentences that are meaningful but not of sentences that are anomalous.

A cued recall test offers one way of gaining access to the interword organization of memory traces. With a high degree of interword organization, the words of a sentence have become interrelated or meaningfully bonded to one another. Thus, another indicator of interword organization is the presence of relational bonds between the subject and object nouns of each study sentence. To test for the existence of such bonds, following the same study procedure as in the above experiments, subjects received a list of noun-pairs with instructions to indicate which pairs they had seen before. The words that appeared in these pairs had all been included in the study sentences, but for some pairs both words came from the same sentence (an intact pair), while for others the words were from two different sentences (a broken pair). It was expected that a subject's ability to recognize an intact pair and to reject a broken pair should depend on whether or not the words of the sentences had been well organized during study. Consequently, for the meaningful material, performance was expected to be better in the Generate than in the Read condition. The words of an anomalous sentence are difficult, if not impossible, to organize or interrelate into a well structured and stable memory trace, in both the Generate and the Read condition. Thus, overall recognition of intact pair was expected to be low with these materials. In addition, since it is difficult to interrelate the words of an anomalous sentence, the difference in performance between the Generate and the Read condition was expected to be small or totally absent. These

expectations were tested with a word-pair recognition test in Experiment 5. The experiment included two groups, with the first receiving one exposure to the study sentences and the second receiving two exposures. The inclusion of the second group was motivated by a desire to boost the level of performance particularly on the anomalous sentences; the two groups also allowed examination of the effects over a range of performance levels.

3.6 Experiment 5

Subjects and Design. Forty-eight subjects participated in this experiment. They were randomly assigned to two groups until there were exactly 24 in each, in such a way that 12 subjects from each group received meaningful sentences and 12 received anomalous sentences. Group 1 received one exposure to the study list; Group 2 received two consecutive exposures to the study list (no pause between presentations). Each exposure was as described in the general method section.

Materials and Procedure. The study list consisted of 32 critical sentences, half of which were shown in each presentation format. The word-pair recognition test included all the nouns of these sentences, arranged into 32 word pairs. Of these 32 pairs, 16 were intact, in that the words were from the same sentence. Eight of these intact pairs came from sentences in the Generate condition, and the remaining eight came from Read sentences. There were also 16 broken pairs. For eight of these the two words came from two Generate

sentences. For the other eight broken pairs, the two nouns came from two Read sentences. In total, the word-pair recognition test consisted of a random listing of 32 pairs. There were eight pairs from each of the following classes: intact Generate, intact Read, broken Generate, and broken Read. The broken pairs were required to assess subjects' ability to discriminate between pairs that contained words from the same sentence and pairs with words from different sentences. With the aid of a mask that exposed only one pair at a time, subjects proceeded through the test list inspecting each pair only once and indicating those that had appeared as pairs (in the same sentence) in the study list. The test was self-paced.

Results and Discussion. On this word-pair recognition test individual subjects could adopt different criteria for discriminating between intact and broken pairs. To illustrate, one subject may correctly identify all intact pairs and not falsely recognize any broken pairs as intact. Another subject may also correctly identify all intact pairs but he or she may falsely recognize every broken pair as intact. If the criterion used by a subject in discriminating between intact and broken pairs is ignored, these two subjects will receive the same recognition score. A more sensitive measure of retention would consider both the subject's ability to identify correctly any intact pairs as well as the ability to reject any broken pairs. A measure that is commonly used for this purpose in the literature is signal detection theory's d' (Green & Swets, 1966).

The procedures of signal-detection theory were employed in summarizing and analyzing the data. First, the number of intact pairs identified as intact was tallied as hits, and the number of broken pairs identified as intact was counted as false alarms. Second, the hits and false alarms obtained for each subject were converted into values of d'. In cases where the proportion of hits was 1.0 and the proportion of false alarms was 0.0, these values were adjusted to .99 and .01 in order to obtain d' scores. The mean number of hits and false alarms, expressed in percentage points, and the mean d' values are shown in Table 3.5. The d' data indicate (a) better performance on meaningful pairs than on anomalous pairs for both groups, (b) better overall performance in Group 2 (two study trials) than in Group 1 (one study trial), (c) a generation effect on the meaningful material for both groups, but no evidence for a generation effect on the anomalous material. In fact, the Group 2 data from the Anomalous condition show substantially better retention of the pairs in the Read than in the Generate condition, but this different did not achieve significance, $t(11) = 1.97$. The hit and false alarm data show essentially the same pattern of results. The d' data were submitted to a three-factor mixed design analysis of variance. This analysis treated processing condition (generate and read) as a within-subjects factor and both materials (meaningful and anomalous) and groups as between-subjects factors (see Appendix B.)

The analysis supported the above observations, revealing a main effect for material, $F(1,44) = 72.52$, $MSe = 1.11$, and for group, $F(1,44) = 14.79$, $MSe = 1.11$. There was also a significant interaction effect

TABLE 3.5

Recognition of word pairs as a function of group, processing condition, and material in Experiment 5 (using a word pair recognition test)

		Group I		Group II	
		Processing condition			
Study material		Generate	Read	Generate	Read
Meaningful	Hits (%)	66.7	38.5	90.6	69.8
	FA's (%)	5.2	6.3	0.0	1.0
	d'	2.57	1.62	3.85	3.04
Anomalous	Hits (%)	33.3	37.5	51.0	67.7
	FA's (%)	18.7	16.7	33.3	27.1
	d'	.72	.86	.65	1.54

for material with processing condition, $F(1,44) = 14.35$, $MSe = .81$, and for material with groups, $F(1,44) = 5.93$, $MSe = 1.11$.

These findings support the expectations based on the process of interword organization. The main effect for material demonstrates the difficulty in grouping a string of unrelated words. The group effect shows that there was an overall increase in performance when study consisted of two exposures as compared to only one, but the interaction of material with group implies that this increase was greater when the sentences were meaningful than when they were anomalous. The interaction of material with processing condition indicates that generated sentences are more likely to be integrated than sentences that were only read, provided the study sentences were meaningful. It appears that when the study material is anomalous (of low meaningfulness) the processing conditions do not differentially affect the likelihood that a sentence will be integrated.

This fifth experiment provided convergent evidence in support of the expected interaction of processing condition with the meaningfulness of the study sentences. The overall pattern of findings from the word-pair recognition test does not differ from that of the cued recall tests. Both of these tests are primarily sensitive to the interword organization of memory traces, that is, to whether or not a sentence was well integrated during study. The consistent finding, in these five experiments, of a generation effect in the Meaningful condition combined with its absence in the Anomalous condition, provides evidence for the critical nature of meaningfulness in the generation effect with these tests.

This chapter on the effects of generating and reading on interword organization includes a final experiment examining yet another implication of this organizational process. Only the meaningful sentences were used in this experiment where subjects were given a free recall test followed by a cued recall test. The implication stemming from interword organization is that cued recall performance should exceed free recall performance, particularly in the Generate condition (Begg, 1972). The notion underlying this implication is that recall cues are necessary for retrieval only for those items which would not be recalled without cues (Slamecka, 1972; Tulving & Pearlstone, 1966). The proposed framework argues that there is an increase in interword organization in the Generate as compared to the Read condition. On this argument, the retrieval cues are expected to be more beneficial to performance in the Generate than in the Read condition. Therefore, retention performance should show an interaction reflecting a greater difference in performance between the cued recall and the free recall test in the Generate condition than in the Read condition. This expectation was examined in Experiment 6.

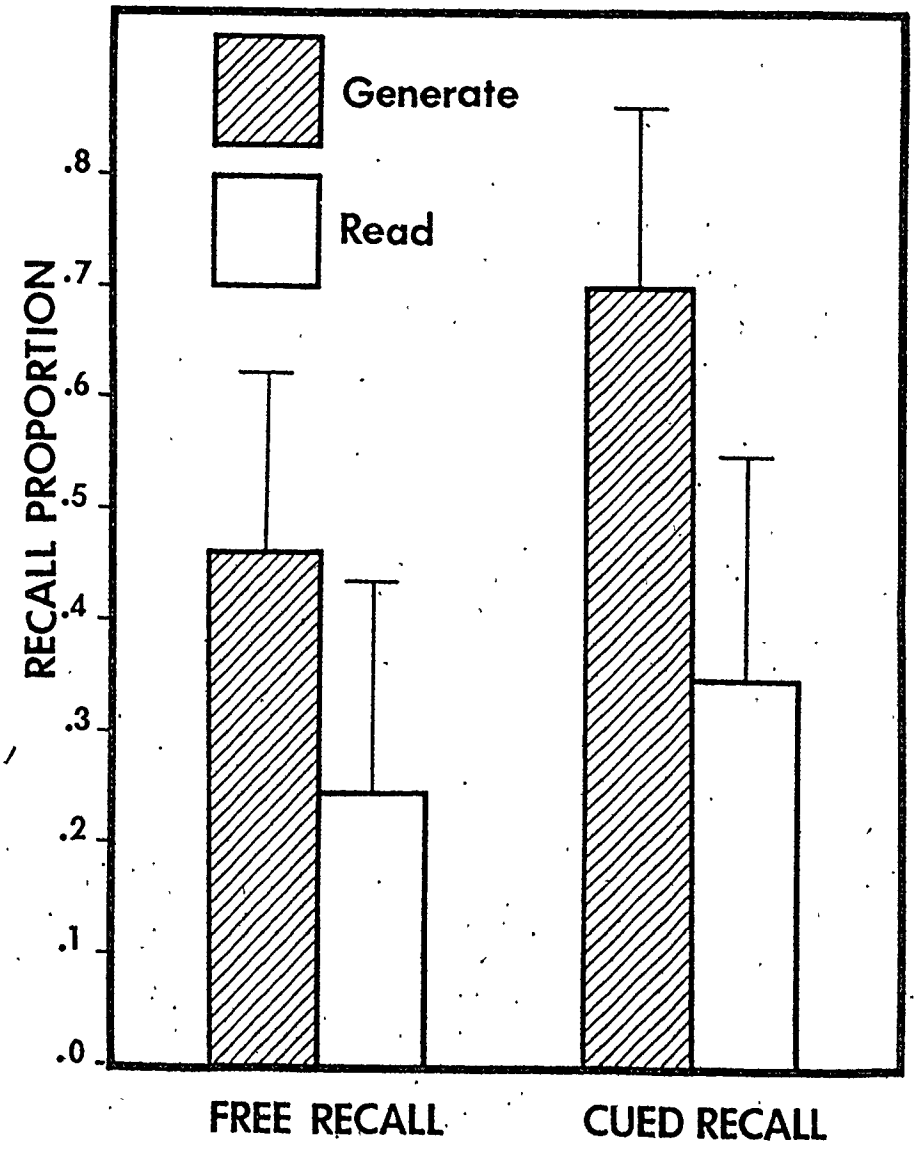
3.7 Experiment 6

Subjects and Design. Sixteen subjects were used in this experiment. The design included the processing condition (generate and read) and the test type (free recall and cued recall) as within-subjects factors.

Material and Procedure. The study list consisted of 16 sentences, eight of which were shown in each presentation format. Each subject received two consecutive exposures to the critical sentences (no pause between exposures) before being tested for recall. The testing phase included a free recall test followed after a 2 minute filled interval by a cued recall test. Six minutes were allowed for each test. In the free recall test subjects were presented a lined sheet of paper with instructions to recall all the sentences, or any parts thereof, in any order. The cued recall test was identical to that used in Experiment 1. The purpose of the filled interval, during which subjects were engaged in conversation, was to prevent direct transfer of materials from the free recall to the cued recall test. The experiment followed the general method in all other respects.

Results and Discussion. In scoring the free recall protocols a noun was considered correctly recalled if it was identical to one of the sentence words. The words could appear anywhere on the recall sheet. The cued recall protocols were scored as in Experiment 1. The mean number of nouns recalled on these two tests and in the Generate and Read condition is shown in Figure 3.2. The figure shows (a) better performance on the cued recall test than on the free recall test, (b) a generation effect on both tests, and (c) a greater difference in performance between the cued recall and the free recall test in the Generate than in the Read condition. These data were submitted to a two-factor repeated measures analysis of variance (see Appendix B). The analysis revealed a significant main effect for processing conditions,

Figure 3.2. Recall of nouns (means and standard deviations) as a function of test type and processing condition in Experiment 6



$F(1,15) = 27.19$, $MSe = 12.58$, for test type, $F(1,15) = 62.37$, $MSe = 2.03$, and an interaction effect between these factors, $F(1,15) = 6.90$, $MSe = 4.38$. Simple Main Effects analyses (Kirk, 1968) supported a generation effect on the free recall test, $F(1,15) = 9.96$, $MSe = 8.48$, as well as on the cued recall test, $F(1,15) = 33.95$, $MSe = 8.48$.

These findings support the expectations based on the process of interword organization. The generation effect observed on the cued recall test replicates the results from the earlier experiments, while the generation effect observed on the free recall test extends these findings to a different test situation. The process of interword organization anticipated the results on the free recall test. The essential difference between the free recall and the cued recall test is defined by the explicitness with which the retrieval cues are stated (Tulving & Watkins, 1973). In the typical cued recall test, these cues are physically presented on the test sheet. On the free recall test, the cues are provided by the test context which includes instructions to recall a specific set of materials, the subject's memory for having studied these materials, and perhaps the presence of the same experimenter in the same physical environment.

The results of the present experiment showing generation effects on both tests reflect the increased organization of the memory traces in the Generate as compared to the Read condition. The higher level of performance on the cued recall than on the free recall test suggests that the specific cues provided on the cued recall test were more effective for retrieval of the sentences than the general cues available for the free recall test. The observed interaction, due to a greater

difference in performance between the cued recall and the free recall test in the Generate than in the Read condition, indicates that the specific cues were substantially more effective in the Generate than in the Read condition. This interaction suggests that generating as compared to reading results in increased interword organization. This interaction also vindicates the focusing on cued recall tests in these experiments by revealing it to be highly sensitive to interword organization.

3.8 Summary of experiments on interword organization

The six experiments included in this chapter shared a pattern of results consistent with the interpretive framework proposed for the generation effect. One central component of the framework is the process of interword organization. According to the framework generating as compared to reading results in an increase in interword organization along a meaning dimension. On this argument, tests that are sensitive to interword organization were expected to show a generation effect with meaningful sentences but not with anomalous sentences. In support of this expectation, this pattern of results was obtained over a wide range of performance levels, with different types of cues provided for recall, over repeated study-test trials, and with different types of tests. The consistent finding of a generation effect on the meaningful sentences, combined with its absence in the anomalous sentences, provides evidence for the critical nature of meaningfulness in the generation effect with the tests used in these experiments. This finding argues strongly that the increased interword organization is

based on a meaning dimension. In short, the evidence warrants the conclusion that generating, as compared to reading, results in increased interword organization along a meaning dimension.

Chapter 4

EMPIRICAL WORK ON INTRAWORD ORGANIZATION

This chapter consists of two experiments on the effects of generating and reading on intraword organization. To repeat intraword organization refers to the organization among the perceptual and conceptual elements within a word (Mandler, 1979). This type of organization is affected by the exposure to a word which provides an opportunity for interrelating its elements into a unified memory trace; it is independent of the relationship among the words of a sentence. The framework proposed in interpretation of the generation effect argues that there is an increase in intraword organization in the Generate as compared to the Read condition.

There are at least two general reasons for expecting increased intraword organization in the Generate as compared to the Read condition. First, Bobrow and Bower (1969) have suggested that comprehension is better for sentences that were generated rather than read. Comprehending a sentence involves an appreciation of the meaning of its component words (Ausubel, 1963), which provides an opportunity for forming a unified memory trace for each word. Thus, better comprehension implies increased intraword organization in the Generate than in the Read condition.

Second, in the present generation paradigm, the requirement to generate a sentence demands close inspection of the presented words in order to decide how to combine them into a specified grammatical frame.

Generating a meaningful sentence seems rather effortless since the semantic constraints among its components almost "make the words fall into place". Griffith (1976) attempted to measure the mental effort required for generating and reading meaningful sentences. He had subjects perform a secondary choice reaction time task while generating and reading sentences. Based on slower choice reactions in the Generate than in the Read condition, Griffith speculated that subjects expend more processing capacity on the generate task than on the read task. The anomalous sentences are not semantically constrained. Thus, perhaps generating an anomalous sentence requires even closer inspection of the component words, particularly since the paradigm allows for only one correct outcome. In short, it seems inherent in generating (see instructions) that individual words are examined more closely and more often, than in reading. Therefore, generating is expected to result in more intraword organization than reading.

Intraword organization is independent of the relationship among the words of a sentence (Mandler, 1979, 1980). In the present experiments the identical words were used to make up the meaningful and the anomalous sentences, although they were combined into the sentences according to different criteria. The same grammatical structure was used to generate both types of sentences, but the words within each sentence were concatenated either meaningfully or randomly. Consequently, on the assumption that generating results in more intraword organization than does reading, it should have the same effect on both the meaningful and the anomalous sentences. Mandler (1980) indicated that intraword organization underlies the experience of word

familiarity. Word familiarity can be measured with a word recognition test (Mandler, 1979, 1980). In a word recognition test the relationships among the words in a sentence are largely irrelevant. Subjects base recognition decisions primarily on information available about each individual word. Therefore, on a word recognition test, a generation effect is expected with both the meaningful and the anomalous sentences.

This expectation was tested in two experiments, the first used a batch recognition test to assess word recognition while the second employed a Yes-No recognition test. These experiments followed the general method outlined in chapter 3.

4.1 Experiment 7

Method. Thirty-two subjects participated in this experiment. They were randomly assigned to the Meaningful and the Anomalous condition until there were exactly 16 in each. The study list consisted of 32 sentences, half of which were shown in each presentation format.

Following study of the sentences a batch recognition test (Brown, 1976) was given. For this test, 128 concrete nouns (similar to those used in the sentences) were selected as distractors. The distractors, and the two nouns from each sentence, were randomly arranged on two test sheets with the constraint that the 32 nouns from the sentences that were generated appeared on one sheet while the 32 nouns from the sentences that were read appeared on the other. The test required subjects to circle as OLD (seen in a study sentence) exactly 32 words on each sheet. About 12 minutes were required for this test.

Results and Discussion. The number of nouns correctly identified as OLD in the various conditions was tallied. The means, expressed in percentage points, are presented in Table 4.1. The table shows the expected higher recognition performance in the Generate than in the Read condition with both the meaningful and the anomalous sentences. These data were submitted to a two-factor mixed design analysis of variance, treating processing condition (generate and read) as a within-subjects factor and material (meaningful and anomalous) as a between-subjecs factor (see Appendix B). The analysis confirmed expectations, revealing a main effect for processing condition, $F(1,30) = 25.77$, $MSe = 9.32$, with no other effects approaching significance.

This pattern of results is as expected from the proposed framework based on two organizational processes. The superior recognition of words in the Generate as compared to the Read conditions implies a greater degree of intraword organization. The findings support the notion that generating requires more detailed inspection of individual words than reading. This detailed inspection of individual words appears to be required whether a to-be-generated sentence is meaningful or anomalous. As a consequence, there is an increase in intraword organization in the Generate as compared to the Read condition for both the meaningful and the anomalous sentences.

To examine the robustness of this finding, and to extend it to a different word recognition test, a second experiment was conducted. This experiment was identical to Experiment 7, except in that a Yes-No word recognition test was used to assess retention.

TABLE 4.1

Recognition of nouns as a function of processing condition and material in Experiment 7 (using a batch recognition test)

Study material	Processing condition			
	Generate		Read	
	M(%)	SD(%)	M(%)	SD(%)
Meaningful	68.75	8.23	55.10	12.75
Anomalous	64.45	9.81	54.10	12.69

4.2 Experiment 8

Method. Thirty-two subjects participated in this experiment. They were randomly assigned to the Meaningful and the Anomalous condition until there were exactly 16 in each. The study list consisted of 32 sentences, half of which were shown in each presentation format.

Following study of the sentences, a Yes-No word recognition test was given. For this test, all the words used on the batch recognition test of Experiment 7 were randomized and presented on the CRT, one at a time. For each word the subject pressed a key to indicate that he or she recognized the word as OLD(YES) or NEW(NO). About ten minutes were required for this self-paced test.

Results and Discussion. Performance on a Yes-No recognition test is influenced by subjects' ability to discriminate between OLD and NEW words. Since individual subjects adopt different discrimination criteria, recognition performance should be assessed with a metric that takes these criteria into account. Such a metric, which is commonly used in psychology, is offered by signal detection theory's d' (Green & Swets, 1966). Thus, the methods of signal detection theory were used in the summary and analysis of the data. These methods were not required for analyzing the data from the batch recognition test used in Experiment 7, since the batch recognition test forces subjects to identify as OLD a fixed number of words. It is assumed that subjects initially identify words that they clearly recognize as OLD, and that they subsequently guess in order to identify the required number of

words. Thus, the criterion adopted by each subject for discriminating between OLD and NEW words is assumed to be equal and determined by the test.

In summarizing the data from the Yes-No recognition test, for each subject the hits (old nouns correctly identified as old) in the Generate and Read condition, and the false alarms (distractor words identified as old) were tallied. (Since the nouns from the sentences that were generated and read were randomly distributed among the test items, the false alarm rate was the same for the Generate and Read condition.) The hits and false alarms for each subject were converted into d' values. Table 4.2 shows the mean number of hits and false alarms, expressed in percentage points, as well as the mean d' values. The table reflects the expected generation effects with both the meaningful and the anomalous sentences on the d' values as well as on the hits. The d' data were submitted to an analysis of variance identical to that used in Experiment 7 (see Appendix B). The analysis revealed a main effect for processing condition, $F(1,30) = 18.72$, $MSe = .18$, with no other effects approaching significance.

The present finding of a generation effect with both the meaningful and the anomalous sentences replicates the results of Experiment 7 and it extends them to a Yes-No recognition test.

4.3 · Summary of experiments on intraword organization

The two experiments included in this chapter show a pattern of results that is consistent with the interpretive framework proposed for

TABLE 4.2

Recognition of nouns as a function of processing condition and material in Experiment 8

(using a Yes-No recognition test)

Study material		Processing condition	
		Generate	Read
Meaningful	Hits (%)	74.60	56.30
	FA's (%)	16.60	16.60
	d'	1.77	1.22
Anomalous	Hits (%)	70.34	59.80
	FA's (%)	19.64	19.64
	d'	1.57	1.21

the generation effect. One central component of the framework is the process of intraword organization. According to the framework, generating as compared to reading results in an increase in intraword organization. Intraword organization is assumed to be independent of the relationship among the words of a sentence. Therefore, on these arguments, a test that is sensitive to intraword organization was expected to reveal a generation effect with both the meaningful and the anomalous sentences. The results obtained in the present experiments support this expectation. Thus, these findings warrant the assertion that generating as compared to reading results in increased intraword organization.

In conclusion, the results from the experiments described in chapters 3 and 4 provide evidence in support of the whole interpetive framework proposed for the generation effect. These experiments show two distinct patterns of results. First, the experiments on interword organization consistently showed an interaction of the generation effect with the meaningfulness of the study material. Second, the experiments on intraword organization consistently showed a generation effect with both the meaningful and the anomalous sentences. In combination, this evidence suggests that generating as compared to reading has two memorial consequences: 1) increased interword organization along a meaning dimension, and 2) increased intraword organization. The next chapter includes experiments that attempted to assess the effects of generating as compared to reading on both of these organizational dimensions simultaneously.

Chapter 5

PREDICTION FROM FRAMEWORK AND ITS EMPIRICAL TEST

The strength of a scientific or theoretical framework resides partially in its ability to make testable predictions. This chapter consists of three experiments that examined a prediction derived from the framework proposed in interpretation of the generation effect. The experiments included in chapters 3 and 4 focused either on an interword or an intraword organization, but they never considered the effects of generating and reading on both of these organizational dimensions simultaneously. The prediction derived from the framework concerns the combined effects of the act of generating on interword and intraword organization. The examination of this prediction requires a test that is highly sensitive to both of these organizational dimensions simultaneously. Such a test was inspired by the work of Humphreys (1976, 1978) and by Rabinowitz, Mandler, and Barsalou (1977). It will be referred to as a varied-context word recognition test.

The varied-context word recognition test takes advantage of the finding that recognition of a word in its study context is substantially better than recognition of a word in a new context (e.g., Humphreys, 1978, Light & Carter-Sobell, 1970; Marcel & Steel, 1973). The authors explain that while primarily information about a word must be relied upon for recognition of the word in a new context (as in a typical word recognition test), when the study context is reinstated during testing, both information about the word and about its relationship to other

words are available for making a recognition decision. More concretely, imagine studying a number of the sentences such as those used in the experiments of chapters 3 and 4, and then being given a word recognition test on the nouns of these sentences. On the test, the nouns are presented in pairs for the purpose of manipulating the test context. In order to reinstate the study context during testing, a noun may be paired with the other noun from the same sentence, constituting an old-context test pair (intact pair). A new-context test environment is created by pairing two nouns from two different sentences (broken pair), or by pairing a sentence noun with a word that was not in a sentence (old-new pair). The evidence on context effects in recognition tests suggests that recognition of a word in an old-context test pair (intact pair) should exceed recognition of the word in a new-context test pair (broken pair or old-new pair). In all types of pairs, words may be directly recognized via information about specific words, but in the old-context test pairs a word enjoys a recognition advantage because it can also be indirectly recognized via the context word. Therefore, when information about the relationships among the words is available, recognition of words tested in old context should exceed recognition of words tested in new context. When no such information is available, recognition performance in these two test contexts should be the same, determined solely by the availability of information about specific words.

Humphreys (1976, 1978) has offered the term item or word information to describe the information available about a word. He has suggested the term relational information to refer to information

available about the relationship among words. Mandler (1980) argued that the psychological processes that underlie item and relational information are intraword and interword organization. The framework proposed for the generation effect asserts that generating as compared to reading has two memorial consequences: 1) increased interword organization along a meaning dimension, and 2) increased intraword organization. Translated in terms of the empirical manifestations of these psychological processes, the framework suggests that generating as compared to reading has two effects: 1) increasing the availability of meaning based relational information, and 2) increasing the availability of item information. Therefore the framework allows a specific prediction about performance on the varied-context word recognition test. Since the effects of generating on intraword organization is independent of the meaningfulness of a sentence, the framework predicts better overall recognition performance in the Generate than in the Read condition for both meaningful and anomalous sentences. This would reflect an effect due to the availability of item information. The overall effect of generating and reading on interword organization should be evidenced by better recognition performance on old-context as compared to new-context test items in the Generate condition. On the additional argument that this interword organization effect is based on a meaning dimension, the difference in performance between old-context and new-context test items should appear only with the meaningful but not with the anomalous sentences (since the words of the anomalous sentences are not meaningfully related nor relatable.)

This prediction was examined in three experiments. The first and third experiments involved meaningful sentences, and the second involved anomalous sentences. In the first two experiments, the processing condition (generate and read) was varied between subjects, while it was a within-subjects variable in the third experiment. The general method employed in these experiments was identical to that outlined in chapter 3. However, in these experiments the presentation of the study sentences was self-paced. The advantage of a self-paced procedure is that it eliminates the need to reject subjects from the experiment for making errors during learning, since they can take as much time as required to generate each sentence. The effects of a self-paced as compared to a paced procedure on the generation effect appears to be minimal, as demonstrated in the experiment presented in Appendix C.

5.1 Experiment 9

Subjects and Design. Thirty-two subjects participated in this experiment. The design included the processing condition as a between-subjects factor. Sixteen subjects were randomly assigned to the Generate and to the Read condition.

Material and Procedure. The study list consisted of 32 meaningful sentences all of which were shown either in the Generate format or in the Read format, depending on the experimental condition. The practice and the filler sentences were presented in the same format as the critical sentences. The varied-context word recognition test

included all the nouns from these sentences plus an equal number of similar new nouns. The 64 new nouns were randomly selected from those used as distractors in Experiments 7 and 8. All words were grouped into pairs in order to create diverse test contexts for the words. Twelve pairs included two nouns that had been in the same sentence, defining an old-context test pair (intact pair). New-context test pairs were created in two ways. First, for 12 pairs two nouns from two different sentences were grouped together (broken pairs). Second, 16 pairs contained one noun from a sentence plus one new noun (old-new pairs). These two types of new-context test items provide alternative measures of recognition performance in new context. On the assumption that recognition of a word in a new context is independent of relational information, the same level of performance was expected on the broken and on the old-new pairs. Finally, the test also included 24 pairs with two new nouns (new-new pairs). All of these word pairs were randomly arranged on the test form. Subjects were informed of the composition of the test. They were instructed to proceed through the test, word by word, circling any word that they recognized as OLD and striking out any that they did not recognize. This procedure insured that each test word was attended to. The test was self-paced.

Results and Discussion. In order to compensate for guessing on the test, the methods of signal detection theory were employed in summarizing the data. First, the number of words correctly identified as OLD in each type of test pair was tallied, giving the hits for intact pairs, broken pairs, and old-new pairs. Second, the number of new words

incorrectly recognized as OLD was established, giving the false alarms for old-new pairs and new-new pairs. The mean number of hits and false alarms, expressed in percentage points, for the different test pairs and processing conditions are shown in Table 5.1a. The table reveals that for each processing condition, recognition of old words was similar on the two new-context test pairs (broken pairs and old-new pairs). This suggests that in these pairs recognition decisions were not differentially affected by the different test contexts, but that they were based on item information alone. Thus, the results from the broken pairs and from the old-new pairs were combined to obtain a single hit score for words tested in new-context test pairs. Similarly, the old-new pairs and the new-new pairs provide alternative sources of false alarms, and they were also combined to yield a single false alarm score for each processing condition. The combined scores are presented in Table 5.1b. The data shown in that table were converted to obtain d' values for each subject. The mean d' values show better overall performance in the Generate than in the Read condition. In addition, performance on the old-context test items was higher than performance on the new-context test items for the Generate condition; in the Read condition performance was similar in both test contexts.

The d' data were submitted to a two-factor mixed design analysis of variance, treating processing condition (generate and read) as a between-subjects factor and test context (old and new) as a within-subjects factor (see Appendix B). The results from the analysis support the above observations showing a significant main effect for processing condition, $\underline{F}(1,30) = 18.0$, $\underline{MSe} = .56$, and for test context, $\underline{F}(1,30) =$

TABLE 5.1

Recognition of nouns as a function of processing condition and test pair in Experiment 9

A) Test results

Processing condition		Test pair			
		Intact	Broken	Old-New	New-New
Generate	Hits (%)	80.99	63.80	66.01	
	FA's (%)			10.55	9.11
Read	Hits (%)	60.4	58.3	57.42	
	FA's (%)			18.36	15.49

b) Combined results

Processing condition		Test context		
		Old		New
Generate	Hits (%)	80.99		64.68
	FA's (%)		9.75	
	d'	2.39		1.84
Read	Hits (%)	60.42		57.97
	FA's (%)		16.19	
	d'	1.35		1.30

20.29, MSe = .07, as well as a significant interaction effect of these two factors, $F(1,30) = 14.35$, MSe = .07. The interaction effect was clarified by Simple Main Effects analyses that revealed a significant difference in performance between the old- and new-context test items in the Generate condition, $F(1,30) = 34.46$, MSe = .07, but not in the Read condition, $F < 1$.

These results are exactly as predicted from the interpretive framework proposed for the generation effect. The framework predicts that generating as compared to reading increases both the availability of meaning based relational information and, the availability of item information. The observed overall superior performance in the Generate as compared to the Read condition reflects an increase in the availability of item information. The superior performance on old-context as compared to new-context test items in the Generate condition indicates the greater availability of relational information. Whether or not this increase in relational information is based on a meaning dimension, as predicted by the framework was examined Experiment 10. Experiment 10 was identical to Experiment 9, except that the anomalous sentences were studied. According to the framework, the pattern of results of Experiment 10 should be identical to that observed in the present study, except that the context within which a word is tested should not affect performance in either processing condition.

5.2 Experiment 10

Method. Thirty-two subjects participated in the experiment. Sixteen were randomly assigned to the Generate and to the Read

condition. The study list consisted of 32 anomalous sentences. All sentences were presented as in Experiment 9. The method of that experiment was also followed in constructing and administering the varied-context word recognition test.

Results and Discussion. The recognition protocols were scored and summarized as in Experiment 9. The mean hits and false alarms for the different test pairs, expressed in percentage points, are shown in Table 5.2a. These data were combined as in Experiment 9 in order to obtain a single hit score for performance on new-context test items as well as a single false alarm score for each processing condition. The combined scores are shown in Table 5.2b. The data presented in that table were converted into values of d'. The d' values show better overall performance in the Generate than in the Read condition, but no difference in performance between old- and new-context test items in either processing condition.

The d' data were submitted to an analysis of variance identical to that used in Experiment 9 (see Appendix B). In support of the above observations, the analysis showed a main effect for processing condition, $F(1,30) = 13.26$, $MSe = .41$, with no other effects approaching significance.

These results support the prediction that generating as compared to reading increases the availability of item information as well as the availability of meaning based relational information. The higher performance in the Generate as compared to the Read condition is attributable to an increase in the availability of item information.

TABLE 5.2

Recognition of nouns as a function of processing condition and test pair in Experiment 10

a) Test results

Processing condition		Test pair			
		Intact	Broken	Old-New	New-New
Generate	Hits (%)	48.94	45.38	44.86	
	FA's (%)			17.56	9.25
Read	Hits (%)	50.81	47.69	47.38	
	FA's (%)			24.22	25.91

b) Combined results

Processing condition		Test context		
		Old		New
Generate	Hits (%)	48.94		45.16
	FA's (%)		11.04	
	d'	1.37		1.27
Read	Hits (%)	50.81		47.50
	FA's (%)		25.49	
	d'	.78		.70

The finding that recognition performance on words tested in old- and new-context test items was similar with the anomalous sentences reflects the lack of relational information in both processing conditions. In Experiment 9, but not in the present experiment, recognition performance on old-context test items exceeded performance on new-context test items in the Generate condition. This difference in the pattern of results must be attributed to the materials studied in these experiments. The difference in results provides evidence that the relational information effect observed in the Generate condition of Experiment 9 was based on a meaning dimension.

In combination, these two experiments support the interpretive framework proposed for the generation effect. These experiments examined a prediction derived from the framework. The prediction was tested in a design that treated processing condition as a between-subjects factor. In order to replicate the present findings and to extend them to a within-subjects manipulation, an additional experiment was carried out, using the meaningful sentences only. In addition, the present results came from experiments where a self-paced study procedure was used. By using a paced study procedure, Experiment 11 further extends the present results. A secondary question addressed by this experiment was whether word recognition in new-context test pairs (broken and old-new) was independent of relational information and based exclusively on item information. It is possible that a context word directs a subject to search memory for inappropriate information. Such an inappropriate search would not be encouraged if a word were presented alone on the test, yet the single word would also present a new-context

test item. Therefore, the varied-context word recognition test used in Experiment 11 included some single old words as well as some single new words. These words add an additional source of recognition in new context, as well as an additional source of false alarms.

5.3 Experiment 11

Subjects and Design. Twenty-four subjects were used. The design included the processing condition (generate and read) as a within-subjects factor and the order of testing items that were generated and read as a between-subjects factor. Twelve subjects were randomly selected for each test order group.

Material and Procedure. The experiment followed the general procedure outlined in chapter 3. The study list consisted of all 48 meaningful sentences. Of these sentences, 24 were presented in each processing condition. Study of the sentences was paced at eight seconds per sentence as described in the general method.

The varied-context word recognition test consisted of two pages, one for testing generated sentences and the other for testing the sentences that were read. The test included all the nouns from the 48 study sentences, plus 96 new ones, randomly selected from those words used as distractors in Experiments 7 and 8. Each test page included eight intact pairs (old-context items), eight broken pairs, eight old-new pairs, 16 new-new pairs, eight single old words, and eight single new words. Subjects were informed of the exact composition of the test. Half the subjects were first tested on the generated sentences and then

on the read sentences; the others were tested in the reverse order. They were instructed to circle words that they recognized as OLD and to strike out the new ones. The test was self-paced.

Results and Discussion. The recognition protocols were scored and summarized as in Experiments 9 and 10, except that the single words on the test provided an additional source of hits and of false alarms. The mean hits and false alarms, expressed in percentage points, for the different test items are shown in Table 5.3a. The table shows that for the new-context test items (broken, old-new, and alone) correct recognition is similar on the different test items in both processing conditions. Thus, it appears that on new-context test items, recognition is independent of relational information and based exclusively on item information. The results from these items were combined to obtain a single hit score for new-context test items, for each processing condition. The test provided for three sources of false alarms (old-new, new-new, and alone items). These data were combined to obtain a single false alarm score for each subject in each processing condition. The combined data, shown in Table 5.3b, were converted into values of d'. The d' values show better overall performance in the Generate than in the Read condition. In addition, performance on old-context test items is substantially better than performance on new-context test items, particularly in the Generate condition.

The d' data were submitted to an analysis of variance. The analysis treated test order as a between-subjects factor and both processing condition and test context (old-context and new-context) as

TABLE 5.3

Recognition of nouns as a function of processing condition and test pair in Experiment 11

a) Test results

Processing condition		Test pair				
		Intact	Broken	Old-New	Alone	New-New
Generate	Hits (%)	94.66	79.08	73.33	81.38	
	FA's (%)			16.96	20.00	17.00
Read	Hits (%)	72.71	65.17	64.83	65.33	
	FA's (%)			16.88	26.75	23.67

b) Combined results

Processing condition		Test context		
		Old		New
Generate	Hits (%)	94.66		78.13
	FA's (%)		17.42	
	d'	2.95		1.97
Read	Hits (%)	72.71		65.08
	FA's (%)		22.96	
	d'	1.56		1.26

within-subjects factors (see Appendix B). The analysis revealed a significant main effect for processing condition, $F(1,22) = 59.67$, $MSe = .44$, and for test context, $F(1,22) = 60.58$, $MSe = .16$, as well as a significant interaction efffect of these two factors, $F(1,22) = 7.74$, $MSe = .35$. The interaction effect was subjected to Simple Main Effects analyses that revealed a significant difference between old- and new-context test items for the Generate condition, $F(1,22) = 44.71$, $MSe = .26$, as well as for the Read condition, $F(1,22) = 4.38$, $MSe = .26$.

These results replicate those of Experiment 9 in all essential respects. The main effect for processing condition reflects the greater availability of item information in the Generate than in the Read condition. The main effect of test context stems from the superior performance on old-context test items as compared to new-context test items. This superiority is attributable to the usefulness of relational information in recognizing words in old context. The interaction, which is due to a greater difference between old-context and new-context test items in the Generate as compared to the Read condition reflects the greater availability of relational information in the Generate condition. The finding of a significant difference between the old- and new-context test items in the Read condition is hardly surprising, since successful reading is an activity that demands an appreciation of the relationships among the words that are read. It is surprising that a similar effect was not also observed in Experiment 9. In summary, the present results replicate those of Experiment 9, and they extend them to a within-subject design and to a paced study procedure.

5.4 Summary of experiments

The three experiments included in this chapter examined a prediction derived from the interpretive framework proposed for the generation effect. The experiments employed the varied-context word recognition test. This test is highly sensitive to both item, and relational information. The results from this test showed a higher level of performance in the Generate condition than in the Read condition with both meaningful and anomalous sentences. This overall difference in performance reflects the greater availability of item information in the Generate than in the Read condition. The test also revealed that in the Generate condition words were better recognized when tested in old- as compared to new-context test items. This difference reflects the greater availability of relational information in the Generate condition. Since this difference between old- and new-context test items was only observed with the meaningful sentences, it appears that the increase in relational information is based on a meaning dimension. In summary, these test results demonstrate that generating as compared to reading has two effects: 1) increasing the availability of item information and 2) increasing the availability of meaning based relational information. These effects were predicted by the framework which asserts that generating as compared to reading has two memorial consequences: 1) increased intraword organization and 2) increased interword organization along a meaning dimension. Thus, the results offer strong support for the framework.

Chapter 6

EXTENSION OF FRAMEWORK TO GENERATION EFFECTS OBTAINED WITH WORD PAIRS

This chapter examines whether the interpretive framework proposed for the generation effect obtained with sentences provides a useful investigative tool that can be extended to generation effects observed with word pairs. To recapitulate, the framework asserts that generating as compared to reading has two memorial consequences: 1) increased interword organization along a meaning dimension and 2) increased intraword organization. The empirical manifestation of those organizational processes are: 1) increased availability of meaning based relational information and 2) increased availability of item information. This section includes an experiment examining whether the generation effect observed with rhyme pairs can also be accounted for in terms of item and relational information.

At least two reasons motivated the extension of the framework to generation effects observed with word pairs. First, considerable recent research on the generation effect has focused on words and on word pairs. An attempt to formulate a comprehensive account of the generation effect necessitates generalizability to this research. Second, numerous authors have considered variants of the notion of increased item and relational information in interpretation of the generation effect. The available data remain ambiguous with respect to these informational sources. The present framework may advance our

understanding of the role of these informational sources in the generation effect with word pairs.

The importance of relational information in the generation effect with word pairs has been suggested by several authors (e.g., Donaldson & Bass, 1980; Jacoby, 1978; Slamecka & Graf, 1978). Donaldson and Bass offered the most intensive investigation of the role of this informational source in the generation effect. In a series of experiments, they had subjects study pairs of related words (e.g., WARM-SUMMER; TIGER-STRIPES). In the Read condition subjects simply read these word pairs. In the Generate condition, each pair was presented with an incomplete right-hand member (e.g., WARM-S _ _ MER; TIGER-STRI _ _ S). Subjects had to generate the incomplete word. The critical findings stem from three groups of subjects. The first group was instructed to generate the incomplete words. The second group simply read the word pairs. The third group read the word pairs first, and then evaluated on a three-point scale, how closely the words of each pair were related to one another. This Read-and-evaluate group was explicitly directed to attend to relational information. The results showed that the Generate group remembered significantly more than the Read group. However, the Read-and-evaluate group performed as well as the Generate group. The requirement to evaluate the relationship between the words of a pair had boosted the recallability of these words to the same level as in the Generate condition. Based on this similarity in results, Donaldson and Bass concluded that the generation effect seems to stem from the increased availability of relational information.

McFarland et al. (1980) argued that relational information cannot account for the generation effect. They instructed subjects to complete sentences with words. Some subjects were asked to generate completion words that followed from the rest of the sentence; other subjects were instructed to produce words that were clearly inappropriate completion words. Compared to a condition where the completion words were experimenter supplied, both the appropriate and the inappropriate completion words were substantially better recalled. Since subjects were unlikely to relate these inappropriate completion words to the sentences, relational information was ruled out as an explanation for the observed superior retention of all generated words. However, the procedure used by McFarland et al. (1980) to obtain inappropriate completion words allowed subjects to generate idiosyncratically memorable words, making interpretation of the results impossible.

Slamecka and Graf (1978) offer data which appear more compatible with an interpretation based on item information. These data come from experiments on rhyming word pairs. In these experiments subjects were provided with pairs of rhyming words (e.g., HALL-BALL; SILK-MILK), and they were told of the rule (rhyme) that related the two words of each pair. In the Read condition, subjects simply read these pairs aloud. In the Generate condition, the right-hand member of each pair was left incomplete (e.g., HALL-B _ _ _; SILK-M _ _ _), and subjects had to generate a rhyming word starting with the initial letter given for the right-hand word. Slamecka and Graf (1978) suggested that in order to generate the right-hand word of a pair, subjects had to examine the

left-hand word and then generate a word related to it via the rhyme rule. On this suggestion, the authors expected better retention of both members of each pair on a later test of memory. However, the results from a recognition test (Experiment 3) revealed a generation effect only on the right-hand member of the pairs and not on the left-hand members. Since only the words actually generated benefitted on the later retention test, it might be argued that with such word pairs relational information contributes little to the generation effect. Slamecka and Graf (1978) abstained from attributing the generation effect to item information, in part, because the results from a later experiment (Experiment 5) were inconsistent with those of Experiment 3. Instead, these authors concluded that the generation effect with word pairs remains without a well-founded interpretation.

The framework proposed for the generation effect with sentences includes both sources of information which have been considered in explanation of the generation effect observed with word pairs. Thus, the present framework was used to illuminate the relative contribution of item and relational information in the generation effect with word pairs. One experiment was carried out. This experiment required subjects to generate and to read rhyming word pairs. The varied-context word recognition test was adapted for testing recognition of these materials. Based on the evidence available on the generation effect obtained with word pairs, an increase in the availability of item information was expected in the Generate as compared to the Read conditions. No specific prediction was made about the availability of relational information.

6.1 Experiment 12

Subjects and Design. Twenty-four subjects participated in this experiment. The design included the processing condition (generate and read) as a within-subjects factor and the order of testing generate and read items as a between-subjects factor. Twelve subjects were randomly assigned to each test order group.

Materials. Forty quadruples of rhyming words were required for this experiment (e.g., CAGE-RAGE-PAGE-WAGE). (These are presented in Appendix A.) Twenty-four quadruples were randomly selected as the study items; the remaining 16 served as distractors in the varied-context word recognition test. The 24 critical quadruples were randomly divided and 12 were assigned to each processing condition. These items were counterbalanced across subjects in such a manner that each quadruple was studied in each processing condition equally often. The 12 quadruples from the Generate and the Read condition were arranged into 24 pairs of words. Each pair was typed on an index card in capital letters. For the Read condition, each pair was typed in its complete form (e.g., CAGE-RAGE). For the Generate condition, the right-hand word of each pair was left incomplete (e.g., CAGE-R ___). An additional 10 rhyming pairs were selected for presentation in the generate format during practice.

Varied-context word recognition test. The test consisted of two pages, including a total of 80 rhyming word pairs. Sixteen of these pairs were old-context test items; they contained two words that had

also appeared together during study (intact pair). Of the intact pairs, eight came from each processing condition. The test included two types of new-context test items (broken and old-new). Sixteen of these items were broken pairs, with eight coming from each processing condition. A broken test item was obtained by re-pairing the two pairs forming a quadruple. Thus, if the two pairs SCORE-CORE and BORE-MORE had appeared in the study list, the broken pairs were SCORE-MORE and BORE-CORE. The remaining 16 new-context test items were old-new pairs. These pairs contained one word from a study pair and one rhyming new word. Of these 16 old-new pairs, eight were contributed by each processing condition. The remaining 32 pairs were obtained by pairing the 16 quadruples set aside as distractors. These 32 pairs, as well as the old-new pairs, served as alternative sources of false alarms. In summary, the test contained the following items: 16 old-context test pairs (intact pairs), 32 new-context test pairs (16 broken pairs and 16 old-new pairs), and 32 new-new pairs. These 80 rhyming pairs were randomly arranged on two test pages. One page contained all the words studied in the Generate condition, plus 16 new-new pairs, while the other page contained all the pairs with words from the Read condition as well as the remaining 16 new-new pairs.

Procedure. Subjects were tested individually. They were instructed to read out loud the two rhyming words appearing on each card, or to generate the incomplete rhyming word and say the pair out loud. The pairs were read in step with a timer which presented a click every four seconds. Subjects were asked to say the words of each pair

aloud during the four seconds and to turn to the next card when they heard the click. Four seconds provided ample time for studying each pair of words. However, when a subject was unable to generate a word in the alloted time, he or she was allowed the next four seconds before continuing on. On average, only 1% of the pairs given in the Generate condition could not be completed in the initial four seconds.

The procedure included a practice phase, a study phase, and after a 2 minute filled interval the test phase. During practice, subjects generated 10 pairs of words. They were then instructed to study the critical pairs for a later test of memory. The exact nature of the test was not specified. The two sets of 24 pairs used in each processing condition were shuffled and the cards containing them presented to the subjects, who proceeded through them twice, turning one card over every four seconds. Following study of the sentences, subjects were engaged in conversation for two minutes to prevent maintenance of the last few items in short term memory. Finally, subjects were given the recognition test with instructions to proceed through it, word by word, circling those words that they recognized as OLD and striking out the new words. About 10 minutes were required for this test.

Results. The methods of signal detection theory were employed in summarizing the data. First, the number of words correctly identified as OLD in each type of test pair was tallied, giving the hits for intact pairs, broken pairs, and old-new pairs. Second, the number of new words incorrectly recognized as OLD was counted, giving the false

alarms for old-new pairs and new-new pairs. The mean number of hits and false alarms, expressed in percentage points, for the different test pairs and processing conditions are shown in Table 6.1a. The table shows that in both processing conditions recognition was similar on the new-context test items (broken pairs and old-new pairs). The results from broken pairs and old-new pairs were combined to yield a single hit score for words tested in new context. Similarly, the old-new pairs and the new-new pairs provided alternative sources of false alarms, and they were also combined to obtain a single false alarm score for each processing condition. The combined data are presented in Table 6.1b. Those data were converted to obtain d' values for each subject. The mean d' values show better overall performance in the Generate than in the Read condition. In addition, in both processing conditions, words were better recognized in old-context test items than in new-context test items.

The d' data were submitted to a three-factor mixed design analysis of variance, treating test order (generate tested first or read tested first) as a between-subjects factor, and processing condition and test context (old and new) as within-subjects factors (see Appendix B). The analysis revealed a significant main effect for processing condition, $F(1,22) = 5.49$, $MSe = .80$, and for test context, $F(1,22) = 53.86$, $MSe = .21$, with no other effects approaching significance.

Discussion. These findings can be discussed in terms of the framework proposed in interpretation of the generation effect obtained with sentences. The overall better performance in the Generate than in

TABLE 6.1

Recognition of nouns as a function of processing condition and test pair in Experiment 12

a) Test results

Processing condition		Test pair			
		Intact	Broken	Old-New	New-New
Generate	Hits (%)	85.37	70.42	65.38	
	FA's (%)			11.15	13.83
Read	Hits (%)	73.38	59.46	59.04	
	FA's (%)			16.96	15.54

b) Combined results

Processing condition		Test context		
		Old		New
Generate	Hits (%)	85.37		68.79
	FA's (%)		13.63	
	d'	2.56		1.79
Read	Hits (%)	73.78		59.25
	FA's (%)		16.17	
	d'	2.05		1.45

the Read condition indicates the greater availability of item information in the Generate condition. The higher level of performance observed on words tested in old-context items as compared to new-context items demonstrates that subject rely on relational information in recognizing individual words. The lack of an interaction between test context and processing condition reveals that similar amounts of relational information were available in both processing conditions. In combination, the present findings suggest that generating as compared to reading a rhyming word pair increases the availability of item information but not the availability of relational information.

The absence of a differential effect of generating and reading on the availability of relational information permits at least two interpretations. First, generating as compared to reading a rhyming pair may not require more careful inspection of the relationship between the words of the pair. As a consequence it may not differentially increase the availability of relational information. Second, generating as compared to reading may differentially increase the availability of relational information, but the varied-context word recognition test did not measure it. The recognition test used in this experiment included new-context test items (broken and old-new pairs) which preserved the rhyming relationship between the words of each pair. The difference between the old-context test items and the new-context test items was that the former items preserved the original pairing of the words. The results show that the observed benefits due to this specific pairing of words was not differentially affected by generating and reading. These

results do not eliminate the possibility that some other type of relational information was more available in the Generate than in the Read condition. In particular, generating rhyming pairs may focus attention on a phonemic dimension. Since the new-context test items used in the present experiment preserved this dimension intact, the test was insensitive to differential amounts of rhyme based relational information in the two processing conditions. Thus the available results remain inconclusive.

6.2 Implications of the framework for generation effects with word pairs

The attempt to use the framework proposed in interpretation of the generation effect obtained with sentences as an investigative tool for examining generation effects observed with word pairs was successful in suggesting directions for future research. The framework postulates an increase in the availability of meaning based relational information in the Generate condition. Perhaps most important, the lack of a differential effect of generating and reading on relational information in Experiment 12 raises questions about other types of relational information. The framework attributes the increase in relational information to an increase in the underlying interword organization of memory units. It is possible that any organizational dimension (rhyme, alphabetical, etc.) may provide a basis for increasing the availability of relational information. Meaning is a very powerful organizational dimension (e.g., Bower, 1972; Postman, 1972) and it is probably the preferred dimension for interrelating the words of a sentence. Yet

other organizational dimensions may prove equally effective with different materials. The varied-context word recognition test can be redesigned to make it sensitive to any of these dimensions. The test results may then show differential effects on relational information in the Read and Generate condition.

Chapter 7

EXTENSION OF THE FRAMEWORK TO THE TRANSFORMATION EFFECT

The transformation effect refers to the superior retention of sentences read in geometrically transformed as compared to normal typography. The framework proposed in interpretation of the generation effect postulates two consequences of generating: 1) increased availability of meaning based relational information and 2) increased availability of item information. The two experiments included in this chapter examined whether the transformation effect can be accounted for in terms of item and meaning based relational information. Thus, this chapter offers an attempt to extend the framework proposed for the generation effect to the transformation effect.

The attempt to extend the framework to the transformation effect was inspired by the similarity in the cognitive activities that are assumed to be involved in generating a sentence and in reading one in a geometrically transformed typography. A sentence displayed in a transformed typography presents a novel and unfamiliar stimulus that is difficult to read. The reader may be forced to analyze such a sentence letter by letter and laboriously attempt to arrange those letters into words. Alternatively, he or she may read only some of its letters and words and attempt to generate the rest of the sentence with the help of these words and with the help of knowledge of the language accumulated in previous encounters with text (Kolers, 1975). The reader may generate hypotheses about the words to follow in a line of text, and

rely on the text only in order to confirm or reject these hypotheses (e.g., Rumelhart, 1977; Smith, 1973). In short, reading a sentence in transformed typography may depend on considerable generative activity. Thus the cognitive activites involved in reading a transformed sentence would bear a close resemblance to the activites involved in generating a sentence. This similarity in the underlying cognitive activities suggests that the transformation effect, like the generation effect, may be decomposable into an item and a relational information component.

The two experiments included in this chapter examined this possibility. One experiment involved meaningful sentences and the other anomalous sentences. The sentences used in these experiments were those from the experiments on the generation effect. These sentences were presented to subjects on index cards, either in normal typography or rotated 180^0 along the plane of the page. The varied-context word recognition test was used to measure retention of the nouns of these sentences. Based on the similarity in the cognitive activities that are assumed to be involved in reading a sentence in a transformed typography and in generating one, the pattern of results was expected to be similar to that observed in the generation effect with sentences.

7.1 Experiment 13

Subjects and Design. Sixteen subjects were used in this experiment. The design included the typography of the study sentences (normal and rotated) as a within-subjects factor and the order of testing the normal and the rotated sentences as a between-subjects factor. Eight subjects were randomly assigned to each test order group.

Material. All 60 meaningful sentences prepared for the generation experiments were used. Each sentence was typed in capital letters on a white index card. The beginning of each sentence was marked with a star; it was provided to aid subjects in locating the beginning of the sentences presented in rotated typography. Table 7.1 displays a sentence in normal and in rotated typography. The varied-context word recognition test was exactly as described in Experiment 11. The test consisted of two pages, one for testing sentences read in normal typography and the other for testing sentences read in rotated typography. Each test page included eight old-context test items (intact pairs), 24 new-context test items (eight broken pairs, eight old-new pairs, and eight single old words), as well as 16 new-new pairs and eight single new words.

Procedure. The procedure was exactly as described in the general method section except that instead of generating sentences subjects were required to read them in transformed typography in the present experiment. They were first given a deck containing the 10 practice sentences, with five in each typography. The sentences were arranged in the deck in such a manner that the typographies alternated. Of the 48 critical sentences, 24 were presented in each typography. They were grouped into blocks of four, and these blocks were presented as described in the general method section. Subjects read each sentence out loud, loudly and clearly, once without errors. They were informed of the subsequent test of retention, although the exact nature of that test was not revealed. The last two sentences of the study deck served

TABLE 7.1

An example of the study material in each typography condition

Sentence typography	
Normal	*THE BLOND GIRL BAKED THE CAKE
Rotated	*THE BLOND GIRL BAKED THE CAKE

as fillers. Following study of the sentences the varied-context word recognition test was administered as in Experiment 11.

Results and Discussion. The recognition test forms were scored and summarized as in Experiment 11. The mean number of hits and false alarms for the different test items, expressed in percentage points, are presented in Table 7.2a. The table shows that for the new-context test items (broken, old-new, and old alone) correct recgonition is similar on the different test items for both sentence typographies. The results from these items were combined to obtain a single hit score for new-context test items for each typography condition. The test provided for three sources of false alarms (old-new, new-new, and new alone). These sources were combined to obtain a single false alarm score for each typography condition. The combined data were converted into values of d'. Table 7.2b shows the combined hits, false alarms and d' scores for the two sentence typographies. The d' values show better overall performance on the rotated sentences than on the sentences that were read in normal typography. In addition, performance on words tested in old-context as compared to new-context test items was substantially better, particularly for sentences read in transformed typography.

The d' data were submitted to an analysis of variance. The analysis treated test order as a between-subjects factor and both sentence typography (normal and rotated) and test context (old and new) as within-subjects factors (see Appendix B). The analysis revealed a significant main effect for typography, $F(1,14) = 31.45$, $MSe = .32$, and for test context, $F(1,14) = 36.54$, $MSe = .13$, as well as a significant

TABLE 7.2

Recognition of nouns as a function of typography and test pair in Experiment 13

a) Test results

Typography		Test pair				
		Intact	Broken	Old-New	Alone	New-New
Rotated	Hits (%)	86.63	70.44	67.50	65.13	
	FA's (%)			14.44	22.88	23.44
Normal	Hits (%)	60.63	55.19	47.13	48.81	
	FA's (%)			18.19	24.50	21.44

b) Combined results

Typography		Test context		
		Old		New
Rotated	Hits (%)	86.63		68.38
	FA's (%)		20.25	
	d'	2.17		1.35
Normal	Hits (%)	60.63		51.44
	FA's (%)		21.38	
	d'	1.09		.83

interaction effect of these two factors, $F(1,14) = 6.54$, $MSe = .20$. No other effects approached significance. The interaction effect was clarified by Simple Main Effects analyses that revealed a significant difference between old- and new-context test items for the sentences read in rotated typography, $F(1,14) = 33.61$, $MSe = .16$, but not for the sentences read in normal typography, $F = 3.34$.

This pattern of results is identical to that observed in Experiment 11 where the same sentences were read and generated. The better overall performance on the sentences that were read in rotated as compared to normal typography indicates an increase in the availability of item information in that condition. The superior performance on words tested in old as compared to new context, for the sentences presented in rotated typography, reflects an increase in relational information for these sentences. The lack of a similar difference in performance between old- and new-context test items for the normal sentences reveals that such sentences can be read without carefully inspecting the relationships among the words of each sentence (Anderson, 1970). Whether or not the relational information effect observed on the rotated sentences is based on a meaning dimension was examined in Experiment 14. In that experiment the anomalous sentences were studied. The overall pattern of results obtained with the anomalous sentences was expected to be the same as that observed with the meaningful sentences, except that neither the rotated nor the normal sentences were expected to benefit from the old-context test items. This last expectation was based on the notion that the relational information effect observed in Experiment 13 was based on a meaning dimension.

7.2 Experiment 14

Method. This experiment was identical to Experiment 13 except that the anomalous sentences were studied. These sentences were typed on index cards and presented as in Experiment 13. The procedure of that experiment was also followed in constructing and administering the varied-context word recogniton test. Sixteen subjects were required for this experiment, with eight randomly assigned to each test order group.

Results and Discussion. The recognition test protocols were scored and summarized as in Experiment 13. The mean number of hits and false alarms, expressed in percentage points, for the different test items are shown in Table 7.3a. The table shows comparable performance levels on the new-context test items (broken, old-new, and old alone) for each typography condition. These data were combined to obtain a single hit score for new-context test items. There were three sources of false alarms for each sentence typography (old-new, new-new, and new alone). These sources were combined to obtain a single false alarm score for each typography condition. The combined data were converted into values of d'. Table 7.3b displays the combined hits, false alarms, and d' scores for each sentence typography. The d' data show better overall performance on the nouns from sentences read in rotated as compared to normal typography, with similar levels of performance on words tested in old- and new-context test items.

The d' data were submitted to an analysis of variance identical to that used in Experiment 13 (see Appendix B). The analysis supported the above observations, confirming a main effect for typography, $F(1,14)$

TABLE 7.3

Recognition of nouns as a function of typography and test pair in Experiment 14

a) Test results

Typography		Test pair				
		Intact	Broken	Old-New	Alone	New-New
Rotated	Hits (%)	63.88	61.81	62.75	70.44	
	FA's (%)			23.00	29.94	17.88
Normal	Hits (%)	41.13	37.69	35.38	33.81	
	FA's (%)			20.56	15.81	16.56

b) Combined results

Typography		Test context		
		Old		New
Rotated	Hits (%)	63.88		64.06
	FA's (%)		21.25	
	d'	1.27		1.21
Normal	Hits (%)	41.13		35.94
	FA's (%)		17.06	
	d'	.81		.67

= 31.24, MSe = .13. There was also a significant interaction effect of test order with test context, F(1,14) = 5.77, MSe = .05, with no other effects approaching significance. The interaction effect reflects a greater difference in performance between old- and new-context test items when the normal sentences were tested first than when the rotated sentences were tested first. Similar interactions involving test order were not obtained in any other experiments.

This pattern of results is identical to that observed in Experiment 10 where anomalous sentences were read and generated. The higher level of performance obtained on the sentences read in rotated as compared to normal typography reflects an increase in the availability of item information with the rotated sentences. The finding that recognition performance on words tested in old-context and new-context test items was similar for both typographies indicates a lack of relational information with these materials. In Experiment 13, but not in the present experiment, recognition performance on old-context test items exceeded performance on new-context test items with the rotated sentences. This difference in the pattern of results must be attributed to the materials studied in these two experiments and it provides evidence that the relational information effect observed with the rotated sentences in Experiment 13 was based on a meaning dimension.

7.3 Summary of experiments

The two experiments included in this chapter offer an attempt to use the framework proposed for the generation effect as an investigative tool for examining the transformation effect. The results from these

experiments indicate that reading a sentence in rotated as compared to normal typography has two effects: 1) it increases the availability of meaning based relational information and 2) it increases the availability of item information. This pattern of results is identical to that observed when sentences were read and generated. The framework proposed in interpretation of the generation effect attributes this pattern of results to two memorial consequences: 1) increased interword organization along a meaning dimension and 2) increased intraword organization. The present results suggest that the same memorial consequences underlie the transformation effect.

The interpretation of the transformation effect in terms of inter- and intraword organization differs substantially from the view presented by Kolers (1973, 1975). Kolers has offered an extensive investigation of the transformation effect. He has suggested that the requirement to read a geometrically transformed sentence changes the average college student into an unskilled reader (Kolers, 1973, 1975). The skilled reader is portrayed as an efficient pattern analyzer, highly selective in the sampling of critical stimulus features, ignoring redundant and irrelevant information, with the ability to focus on the distinguishing characteristics of a stimulus (Kolers, 1975). Unskilled reading is not just slower than skilled reading, it is less selective and consequently involves more encoding of relevant and irrelevant information. Thus, Kolers argues that the skilled reader is able to pay more attention to the message embodied in the text, while the unskilled reader spends more time analyzing its superficial aspects. He maintains that the initial performance of an activity facilitates the subsequent

performance of the same activity. The more unique, individualized, and varied the cognitive activities involved in the initial task, the greater the opportunity for reinstating some of these activities at a later time. In short, Kolers' explanation of the transformation effect is based on the argument that reading a geometrically transformed text requires more extensive analysis of the study sentences than reading normal text, and thus there is a greater chance for reinstantiation of the cognitive activities involved in the initial reading of the sentence during testing.

The present findings are partially inconsistent with Kolers' view. Consistent with his view, the overall increased memorability of the individual words from both types of sentences read in rotated as compared to normal typography may be described in terms of the memorial consequences of skilled and unskilled reading. However, the observed increase in meaning based relational information for meaningful sentences read in rotated as compared to normal typography suggests that the "unskilled" reader, rather than the skilled reader, attends more to the message embodied in a text. The findings imply that the skilled reader, who is turned into an unskilled reader by a geometrical transformation of the text, begins to rely more on knowledge accumulated in previous encounters with text, perhaps in order to compensate for the laborious processing of the transformed text (Graf, 1981).

The framework proposed in interpretation of the generation effect attributes the transformation effect to the organizational processes involved in reading normal and transformed sentences. The laborious, time consuming, and detailed analysis of a geometrically

transformed sentence encourages the reader to draw upon all sources of information about the presented sentence. Some of that information has been accumulated in memory from previous encounters with text. By relying on stored information about specific words and about the relationships among the words of the sentence, the reader formulates hypotheses about words that may follow next in a line of text. Partial information from the text may be used to confirm or reject these hypotheses. The net result of this word guessing game (Goodman, 1967) is an increase in the interword organization along a meaning dimension and an increase in intraword organization.

Thus, the advantages of the organizational account for the transformation effect are, first that it identifies that phenomenon as one in a class of similar memory phenomena that seem to be characterized by forcing the learner to attend carefully to the components of the learning materials as well as to the relationships among those components. Second, Kolers' account attributes the different results observed with the meaningful and the anomalous sentences to differences in the cognitive activities that are assumed to be required for reading these sentences in both typographies. However, his account offers no explicit basis for expecting separable effects on the availability of item and of relational information. In short, Kolers' view requires elaboration before it can accommodate the different pattern of results observed with meaningful and anomalous sentences. Third, Kolers' account suggests a positive relationship between retention performance and the complexity of an initial processing task. The superior retention of rotated as compared to normal sentences provides some

evidence in support of such a relationship. While the present organizational account also anticipates an increase in retention performance with an initial change away from normal reading conditions, it does not postulate a positive relationship between retention performance and the complexity of the processing task, since the knowledge sources available to the reader are assumed to be limited. Evidence from Kolers (1973) and from Graf (1981) indicates that although different geometrical transformations of sentences substantially increase the difficulty of reading, there is no corresponding increase in retention.

Chapter 8

GENERAL DISCUSSION

This thesis presented and empirically tested an interpretive framework for the generation effect obtained with sentences. It includes a systematic collection of evidence in support of this framework. The framework has specific implications for the generation effect with sentences. These implications permit the discrimination of the framework from existing alternative views offered in interpretation of the generation effect. The empirical examination of these implications provided evidence in support of the framework. Alternative interpretive views, in their present form, cannot readily accommodate these findings. Thus, by eliminating alternative interpretive views offered for the generation effect and by accommodating it under a new framework, this thesis contributes to our understanding of the generation phenomenon. Moreover, the extension of the framework to the transformation effect reveals it as an effective investigative tool that promises to be useful in the examination not only of the generation effect but also of related memory phenomena.

8.1 Summary of research

The framework presented in interpretation of the generation effect obtained with sentences concentrates on the effects of generating as compared to reading on interword organization and on intraword organization. The relationship between these organizational processes

and retention performance has been extensively documented. The involvement of these processes in the generation effect was assessed in two series of experiments. The experiments on interword organization demonstrated a substantial retention advantage in the Generate as compared to the Read condition when meaningful sentences were studied but no generation effect when the anomalous sentences were studied. This pattern of results was obtained over a wide range of performance levels, multiple study-test trials, with different types of tests and test-cues, and when study was paced and unpaced. The consistent finding of a generation effect with meaningful sentences combined with its absence on the anomalous sentences establishes that generating as compared to reading results in increased interword organization along a meaning dimension.

A very different pattern of results was observed in experiments on intraword organization. These experiments revealed a generation effect on both the meaningful and the anomalous sentences. Thus, unlike the effects of generating and reading on interword organization, the effect of generating on intraword organization is independent of the meaningfulness of the study sentences. Generating as compared to reading increases intraword organization in both the meaningful and the anomalous sentences.

In combination, the evidence in support of each organizational process indicates that generating as compared to reading has two consequences: 1) increasing interword organization along a meaning dimension and 2) increasing intraword organization. The empirical manifestations of these organizational processes are: 1) an increase in

the availability of meaning based relational information and 2) an increase in the availability of item or word information. The combined effects of generating and reading on both of these sources of information were examined in a separate series of experiments. The varied-context word recognition test was used in these experiments. This test was designed for assessing the combined effects of generating as compared to reading on relational and on item information simultaneously. The findings from this test provide additional evidence in support of the framework as a whole.

The framework was used as an investigative tool in the examination of the generation effects obtained with pairs of related words. In particular, retention of rhyming word pairs was assessed with the varied-context word recognition test adapted for these materials. The findings suggest that while generating as compared to reading a pair of rhyming words increases the availability of item information, the two processing conditions do not appear to have a differential influence on the availability of relational information.

However, any strong conclusions about the effects of generating as compared to reading rhyme pairs on the availability of relational information must remain tentative. All test items included in the varied-context word recognition test used in the measurement of relational information contained two words that rhymed with one another. Some test items included two words that had appeared together in the same study pair while others included two words that had appeared in two different study pairs. Thus, the relational information available in the former test items, but not in the latter, was based on the specific

combination of two words in a pair. The test results showed no differential availability of this type of relational information when pairs of rhyming words were read and self-generated. However, the test was insensitive to other types of relational information, in particular to phonemically based relational information. An appreciation of the phonemic relationship between the words of a rhyming pair seems to be required in the Generate but not in the Read condition (Slamecka & Graf, 1978). Thus, the possibility remains to be investigated that generating as compared to reading rhyme pairs does increase the availability of phonemically based relational information. The varied-context word recognition test can readily be adapted for measuring differences in phonemically based relational information.

The framework was also applied in the examination of the detailed memorial consequences underlying the transformation effect. The pattern of results from two experiments on the transformation effect was identical to that observed in the generation effect. Therefore, in revealing the basic similarities underlying these two memory phenomena, the framework also contributes to our understanding of the transformation effect.

Overall, the framework offered in interpretation of the generation effect is supported by a systematic collection of evidence. This evidence comes from experiments on the retention of sentences that were read and self-generated. The additional evidence from the generation effect with rhyming word pairs and from the transformation effect reveals the framework as a powerful investigative tool that

promises to be useful in the interpretation of phenomena related to the generation effect obtained with sentences.

8.2 Theoretical advances

The major objective of this thesis was to present and to test an interpretive framework for the generation effect obtained with sentences. The empirical investigation of predictions derived from the framework provides supporting evidence. Thus, the objective of this thesis has been met by accommodating the generation effect obtained with sentences under a framework based on well established psychological processes.

This framework questions extant alternative views as serious interpretive candidates for the generation effect obtained with sentences. The evidence accumulated in support of the framework cannot readily be accommodated by the alternative interpretive views. The most prominent of these views are: (a) the comprehension view, (b) the levels of processing view, (c) the distinctiveness view, and (d) the retrieval practice view. The comprehension view is based on the notion that generating as compared to reading leads to better comprehension which in turn mediates increased retention performance (Bobrow & Bower, 1969). The levels of processing view implies that generating involves deeper semantic processing than does reading. The distinctiveness view asserts that a more distinctive memory episode is created when verbal materials are self-generated rather than just read. The distinctiveness of a memory episode is positively related to its retention (Jacoby et al., 1979; Eysenck, 1979). Finally, the retrieval practice view argues

that generating, but not reading, requires retrieval of materials from semantic memory. An initial retrieval is assumed to facilitate subsequent retrieval of the same materials (Bjork, 1975; Rabinowitz, Mandler, & Patterson, 1977). The common denominator of these views is that they attribute the generation phenomenon to one critical component or process (i.e., comprehension, depth of processing, distinctiveness, or retrieval from semantic memory) that is assumed to be more involved in generating than in reading.

A single component view can account for the effects of generating as compared to reading at a single locus. Such a view can accommodate main effects in the results but it does not predict and cannot readily accommodate effects at two loci. Thus, the views previously offered in interpretation for the generation effect are able to interpret the evidence accumulated in the demonstration of the effect of generating as compared to reading on intraword organization. In the experiments focusing on intraword organization, generating resulted in a higher level of performance than reading whether the study sentences were meaningful or anomalous. This consistent difference in performance between generating and reading may be attributed to a difference in comprehension, a difference in the level of processing or in distinctiveness, or a difference in retrieval practice, just as it was interpreted as reflecting a difference in intraword organization. The author's theoretical preference for the concept of intraword organization was based on a number of considerations. First, the notion of comprehension is vague and it is difficult to assess it unambiguously with the materials that are commonly used in generation experiments

(Anderson, 1972). Moreover, the present framework emphasizing inter- and intraword organization subsumes the memorial consequences that are clearly implied by the notion of comprehension. Second, the notions of depth of processing and distinctiveness are also vague and they have been criticized for lacking independent indexes (Baddeley, 1978; Postman, 1975). These notions were intended and successfully employed as heuristics for describing and relating diverse memory phenomena. Their focus on a single locus of the effects of generating as compared to reading, however, limits the profitability of their application to the generation phenomenon. Third, the retrieval practice account has been criticized as bordering on the circular (Slamecka & Graf, 1978). In addition, this notion requires elaboration before it can even accommodate the present finding of a generation effect on words that were supplied to, rather than generated by the subjects. In contrast, the process of intraword organization is well established in the literature and it has been subjected to extensive investigation.

The previous interpretive views lack a component for dealing with the effects of generating as compared to reading at two loci, that is, with the observed interaction effects. Interaction effects were obtained in the experiments focusing on interword organization. In these experiments, there was a consistent generation effect when meaningful sentences were studied but no generation effect on the anomalous sentences. While the single component views include a component for explaining the generation effect obtained with the meaningful sentences, the same component cannot also explain the absence of an effect on the anomalous sentences. These views could be modified

to enable them to accommodate the effects of generating on interword organization. However, such modifications would render them indistinguishable from the framework proposed in this thesis. After such modifications, the present framework would still enjoy numerous advantages stemming primarily from an extensive literature on the organizational processes involved in the framework.

The inability to accommodate the observed interaction effects also eliminates the more quantitative views offered in interpretation of the generation phenomenon. These views include the effort view, strength theory, frequency theory, and the total time hypothesis. Since these views attribute the generation effect found with meaningful sentences to a quantitative difference in effort, in frequency of exposure, in strength of memory traces, or in study time, they also predict a generation effect with anomalous sentences. Modification of these views to enable them to accommodate the present results would be difficult.

Since the organizational processes involved in the proposed framework have been extensively investigated and discussed in the psychological literature, this rich body of literature can serve as a source of ideas for guiding future investigations on the generation effect. For example, the literature contains evidence on the differential effectiveness of various organizational dimensions, such as meaning based, structurally based or phonemically based organization (Bower, 1972; Bruce & Crowley, 1970). The ability to impose such organizations upon to-be-remembered materials develops with age (Liberty & Ornstein, 1973), while subjects can benefit from such organizations

long before they organize materials in an effective manner spontaneously. Thus, the literature on these organizational processes offers concrete suggestions for questions that remain to be asked about the generation phenomenon. Moreover, in many cases, it also offers experimental procedures that can be applied in the investigation of these questions.

Finally, the framework offers an interpretive tool for investigating and relating the generation effect to the transformation effect. By relating these two previously unrelated memory phenomena, the insights gained into one phenomenon can advance investigation into the other phenomenon. Progress in science depends on such cross-fertilization of ideas and on the establishment of relationships between empirical phenomena.

8.3 Methodological advances

The work presented in this dissertation also offers a novel paradigm for gaining control over materials that are generated by subjects, and thus for limiting the source of the generation effect to the activities invovled in generating and in reading. The lack of such a paradigm prevented previous investigators from working with sentences, forcing them to focus on generation effects obtained with words and word pairs. This limitation made it difficult to examine many of the theoretical arguments offered in interpretation of the generation effect (such as the notion of comprehension). The present paradigm is highly flexible. It can not only be used to generate meaningful and anomalous sentences, but also a series of sentences composing a paragraph or

nonsense words (a string of letters ordered according to a rule). The examination of the generation effect with paragraphs may permit an assessment of the role of comprehension in the generation effect. The examination of the generation effect obtained with various types of nonsense words could provide additional insights into the nature of intraword organization. In short, the flexibility of this paradigm offers a convenient methodology for examining new aspects of the generation phenomenon.

This work also pioneers the application of the varied-context word recognition test in the examination of the generation and the transformation effect. This test has been used by previous investigators to examine recall and recognition processes (Humphreys, 1976, 1978). This test is highly adaptible. It can be made sensitive to meaning based relational information, phonemically based relational information, as well as many other types of relational information. However, its main advantage stems from its simultaneous sensitivity to both item and relational information.

8.4 Directions for future research

Many important questions about the generation phenomenon remain to be answered. Perhaps most directly related to the present work, questions remain about alternative organizational dimensions that might contribute to an increase in relational information in the Generate as compared to the Read condition. Phonemically based relational information is an obvious candidate when rhyming word pairs are studied. But differences in interword organization could also be based on

grammatical function, temporal contiguity, and various perceptual attributes of verbal materials (Bower, 1972). Whether or not these diverse organizational dimensions are as effective, and as long lasting, as meaning based organization remains to be investigated.

The process of intraword organization also merits further investigation; it refers to the organization among the perceptual and conceptual elements within a word (Mandler, 1979). A change in interword organization depends on exposure to and inspection of a word. In the present generation paradigm, words may be examined in a slightly different manner when a meaningful rather than an anomalous sentence is generated. In order to generate a meaningful sentence, a subject may examine primarily the meaning of a word to place it into the specified sentence frame; generating an anomalous sentence might require determining the grammatical function of a word. The subtle differences in generating these two types of sentences may have different effects on intraword organization.

It also remains to be determined whether the observed effects of generating as compared to reading on interword and intraword organization depend critically on subjects being supplied with the words composing each sentence. In previous experiments on the generation effect subjects always actually generated words. In the present experiments they generated sentences with the aid of words that were supplied by the experimenter. Whether or not this difference in the generate task is critical could be determined by supplying subjects with words that are incomplete (i.e., C_R, STR_CK, BIC_CLE). Subjects would

have to generate these words (complete them) before they could order them into a sentence frame.

Future research should also examine the role of these two organizational processes in generation effects observed with word pairs. While this thesis offers an attempt in this direction, it is far from being complete. Since considerable recent research on the generation effect has focused on word pairs, a comprehensive interpretive view of the effect should be able to accommodate these findings.

Finally, in focusing on the memorial benefits associated with generating as compared to reading, the present work has revealed the ineffectiveness of reading as a learning strategy. More specifically, the evidence demonstrated that subjects retain too little meaning based relational information when they read a sentence. This finding is surprising and alarming since reading is generally thought to be a task that involves semantic analysis of the presented materials. It is possible that the reading strategy used by subjects in the present experiments is different from the strategy they employ in normal reading. This difference in reading strategy may stem from being presented with unrelated sentences rather than with connected text. On this argument, the observed pattern of results may not hold when subjects generate and read paragraphs rather than single sentences. This possibility must be investigated before the present findings are used to reach conclusions about the ineffectiveness of reading as a learning strategy.

While recognizing these important avenues for future research, the primary objective of this thesis has been met. An interpretive

framework for the generation effect obtained with sentences has been offered and empirically tested. The present work adds important insights into the generation and the transformation effect, and it has potentially important implications for future research on the generation effect and on related memory phenomena.

FOOTNOTES

[1] Some of the experiments described in this dissertation have previously been published in P. Graf, Two consequences of generating: Increased inter- and intraword organization of sentences, Journal of Verbal Learning and Verbal Behavior, 1980, 19, 316-327. This paper includes experiments 1, 3, 4, 5, and 8.

[2] In a few experiments, some subjects did not meet this requirement. In some cases, a subject was unable to generate a sentence in the available time. In other cases, an incorrect sentence was generated (i.e., by incorrect grammar, adding unacceptable words, misreading a word, or reversing subject and object nouns). In the latter case, a subject was allowed to continue in the experiment provided that only one such error occurred and provided that all available words had been used in a grammatical sentence. The number of subjects who failed to meet the participation requirement was as follows: five in Experiment 1, three in each of Experiments 2 and 3, one in Experiment 4, seven in Experiment 5, five in Experiment 7, and six in Experiment 8. Of these 30 subjects, nine attempted to generate meaningful sentences and 21 anomalous sentences.

REFERENCES

Adams, J.A., & Montague, W.E. Retroactive inhibition and natural language mediation. Journal of Verbal Learning and Verbal Behavior, 1967, 6, 528-535.

Anderson, R.C. Control of student mediating processes during verbal learning and instruction. Review of Educational Research, 1970, 40, 349-369.

Anderson, R.C. How to construct achievement tests to assess comprehension. Review of Educational Research, 1972, 42, 145-170.

Anderson, R.C., Faust, G.W., & Roderick, M.C. "Overprompting" in programmed instruction. Journal of Educational Psychology, 1968, 59, 88-93.

Anderson, R.C., Goldberg, S.R., & Hidde, J.L. Meaningful processing of sentences. Journal of Educational Psychology, 1971, 62, 395-399.

Anderson, R.C., & Kulhavy, R.W. Learning concepts from definitions. American Educational Research Journal, 1972, 9, 385-390.

Anderson, R.C., Royer, J.M., Kulhavy, R.W., Thornburg, S.D., & Klemt, L.L. Thematic prompting in paired-associate learning. Journal of Educational Psychology, 1971, 62, 315-321.

Auble, P.M., & Franks, J.J. The effects of effort toward comprehension on recall. Memory & Cognition, 1978, 6, 20-25.

Ausubel, D.P. The psychology of meaningful verbal learning. New York, N.Y.: Grune & Stratton, Inc., 1963.

Baddeley, A.D. The trouble with levels: A reexamination of Craik and Lockhart's framework for memory research. Psychological Review, 1978, 85, 139-152.

Begg, I. Recall of meaningful phrases. Journal of Verbal Learning and Verbal Behavior, 1972, 11, 431-439.

Bjork, R.A. Retrieval as a memory modifier: An interpretation of negative recency and related phenomena. In R.L. Solso (Ed.), Information processing and cognition: The Loyola Symposium. Hillsdale, N.J.: Erlbaum, 1975.

Bobrow, S.A., & Bower, G.H. Comprehension and recall of sentences. Journal of Experimental Psychology, 1969, 80, 455-461.

Borkowski, J.G., & Eisner, H.C. Meaningfulness and abstractions in short-term memory. Journal of Experimental Psychology, 1968, 76, 57-61.

Bower, G.H. A multicomponent view of the memory trace. In K.W. Spence and J.T. Spence (Eds.), The psychology of learning and motivation: Advances in research and theory, Vol. 1. New York: Academic Press, 1967.

Bower, G.H. A selective review of organizational factors in memory. In E. Tulving and W. Donaldson (Eds.), Organization of memory. New York: Academic Press, 1972.

Bower, G.H., & Winzenz, D. Comparison of associative learning strategies. Psychonomic Science, 1970, 20, 119-120.

Brown, J. (Ed.), Recall and recognition, London: Wiley, 1976.

Bruce, D., & Crowley, J.J. Acoustic similarity effects on retrieval from secondary memory. Journal of Verbal Learning and Verbal Behavior, 1970, 9, 190-196.

Bruner, J.S. Some elements of discovery. In L.S. Shulman and E.R. Keislar (Eds.), Learning by discovery. Chicago: Rand McNally, 1966, 101-114.

Bruner, J.S. The act of discovery. Harvard Educational Review, 1961, 31, 21-32.

Bruner, J.S. The process of education. Cambridge, Mass.: Harvard University Press, 1960.

Cermack, L.S., & Craik, F.I.M. (Eds.), Levels of processing in human memory. Hillsdale, N.J.: Erlbaum, 1979.

Craik, F.I.M. A "levels of analysis" view of memory. In P. Pliner, L. Krames, and T.M. Alloway (Eds.), Communication and affect: Language and thought. New York: Academic Press, 1973, 45-65.

Craik, F.I.M. Depth of processing in recall and recognition. In S. Dornic and P.M.A. Rabbitt (Eds.), Attention and performance, Vol. VI, 1977.

Craik, F.I.M. Human memory. Annual Review of Psychology, 1979, 30, 63-102.

Craik, F.I.M. & Jacoby, L. Elaboration and distinctiveness in episodic memory. In L.G. Nilsson (Ed.), Perspectives in memory research. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1979.

Craik, F.I.M., & Lockhart, R.S. Levels of processing: A framework for memory research. Journal of Verbal Learning and Verbal Behavior, 1972, 11, 671-684.

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

Cronbach, L.J. The logic of experiments on discovery. In L.S. Shulman and E.R. Keislar (Eds.), Learning by discovery. Chicago: Rand McNally, 1966, 77-92.

Darley, C.F., & Murdock, B.B. Effects of prior free recall testing on final recall and recognition. Journal of Experimental Psychology, 1971, 91, 66-73.

Della-Piana, G. An experimental evaluation of programmed learning. Journal of Educational Research, 1962, 55, 495-501.

Dember, W.N. Motivation and the cognitive revolution. American Psychologist, 1974, 29, 161-168.

Dewey, J. How we think. Boston: Heath, 1910.

Doctorow, M., Wittrock, M.C., & Marks, C. Generative processes in reading comprehension. Journal of Educational Psychology, 1978, 70, 109-118.

Donaldson, W., & Bass, M. Relational information and memory for problem solutions. Journal of Verbal Learning and Verbal Behavior, 1980, 19, 26-35.

Erdelyi, M., Buschke, H., & Finkelstein, S. Hypermnesia for Socratic stimuli: The growth of recall for an internally generated memory list abstracted from a series of riddles. Memory & Cognition, 1977, 5, 283-286.

Eysenck, M.W. Depth, elaboration, and distinctiveness. In L.S. Cermack and F.I.M. Craik (Eds.), Levels of processing in human memory. Hillsdale, N.J.: Erlbaum, 1979.

Eysenck, M.W. Levels of processing: A critique. British Journal of Psychology, 1978, 69, 157-169.

Fillmore, C.J. The case for case. In E. Bach and R.T. Harms (Eds.), Universals in linguistic theory. New York: Holt, Rinehart, & Winston, 1968.

Gardiner, J.M., Craik, F.I.M., & Bleasdale, F.A. Retrieval Difficulty and Subsequent Recall. Memory & Cognition, 1973, 1, 213-216.

Garner, W.R. The processing of information and structure. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1974.

Goodman, K.S. Reading: A psycholinguistic guessing game. Journal of the Reading Specialist, 1967, 6, 126-135.

Graf, P. Reading and generating normal and transformed sentences. Canadian Journal of Psychology, 1981, 35, in press.

Green, D.M., & Swets, J.A. Signal detection theory and psychophysics. New York: John Wiley and Sons, Inc., 1966.

Griffith, D. The attentional demands of mnemonic control processes. Memory & Cognition, 1976, 4, 103-108.

Harris, G., Begg, I., & Upfold, D. On the role of the speaker's expectations in interpersonal communication. Journal of Verbal Learning and Verbal Behavior, 1980, 19, 597-607.

Humphreys, M.S. Item and relational information: A case for context independent retrieval. Journal of Verbal Learning and Verbal Behavior, 1978, 17, 175-188.

Humphreys, M.S. Relational information and the context effect in recognition memory. Memory & Cognition, 1976, 4, 221-232.

Jacoby, L. On interpreting the effects of repetition: Solving a problem versus remembering a solution. Journal of Verbal Learning and Verbal Behavior, 1978, 17, 649-667.

Jacoby, L. The role of mental contiguity in memory: Registration and retrieval effects. Journal of Verbal Learning and Verbal Behavior, 1974, 13, 483-496.

Jacoby, L., Craik, F.I.M., & Begg, I. Effects of decision difficulty on recognition and recall. Journal of Verbal Learning and Verbal Behavior, 1979, 18, 585-600.

Jensen, A.R., & Rohwer, W.D., Jr. Verbal mediation in paired-associate and serial learning. Journal of Verbal Learning and Verbal Behavior, 1963, 1, 346-352.

Kahneman, D. Attention and effort. Englewood Cliffs, N.J.: Prentice-Hall, 1973.

Kane, J.H., & Anderson, R.C. Depth of processing and interference effects in the learning and remembering of sentences. Journal of Educational Psychology, 1978, 70, 626-635.

Keenan, J.M., MacWhinney, B., & Mayhew, D. Pragmatics in memory: A study of natural conversation. Journal of Verbal Learning and Verbal Behavior, 1977, 16, 549-560.

Kemp, F.D., & Holland, J.G. Blackout ratio and overt responses in programmed instruction: Resolution of disparate results. Journal of Educational Psychology, 1966, 57, 109-114.

Kirk, R.E. Experimental design: Procedures for the behavioral sciences. Belmont, California: Wadsworth Publishing Company Inc., 1968.

Kolers, P.A. Memorial consequences of automatized encoding. Journal of Experimental Psychology: Human Learning and Memory, 1975, 1, 689-701.

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

Kolers, P.A. Remembering operations. Memory & Cognition, 1973, 1, 347-355.

Kolers, P.A., & Ostry, D.J. Time course of loss of information regarding pattern analyzing operations. Journal of Verbal Learning and Verbal Behavior, 1974, 13, 599-612.

Liberty, Ch., & Ornstein, P.A. Age differences in organization and recall: The effects of training in categorization. Journal of Experimental Child Psychology, 1973, 15, 169-186.

Light, L.L., & Carter-Sobell, L. Effects of changed semantic context on recognition memory. Journal of Verbal Learning and Verbal Behavior, 1970, 9, 1-11.

Lockhart, R.S., & Craik, F.I.M. Levels of processing: A reply to Eysenck. British Journal of Psychology, 1978, 69, 171-175.

Lockhart, R.S., Craik, F.I.M., & Jacoby, L. Depth of processing in recognition and recall: Some aspects of a general memory system. In J. Brown (Ed.), Recognition and Recall. London: Wiley, 1975.

Mandler, G. Organization and memory. In K.W. Spence and J.T. Spence (Eds.), The psychology of learning and motivation, Vol. 1. New York: Academic Press, 1967.

Mandler, G. Organization and repetition: Organizational principles with special reference to rote learning. In L.G. Nilsson (Ed.), Perspectives in memory research. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1979.

Mandler, G. Recognizing: The judgment of previous occurrence. Psychological Review, 1980, 87, 252-271.

Marcel, A.J., & Steel, R.G. Semantic cueing in recognition and recall. Quarterly Journal of Experimental Psychology, 1973, 25, 368-377.

Marks, L.E., & Miller, G.A. The role of semantic and syntactic constraints in the memorization of English sentences. Journal of Verbal Learning and Verbal Behavior, 1964, 3, 1-5.

Masson, M.E.J., & Sala, L.S. Interactive processes in sentence comprehension and recognition. Cognitive Psychology, 1978, 10, 244-270.

McFarland, C.E., Jr., Frey, T.J., & Rhodes, D.D. Retrieval of internally versus externally generated words in episodic memory. Journal of Verbal Learning and Verbal Behavior, 1980, 1, 210-225.

Miller, G.A. The magical number seven, plus or minus two: Some limits on our capacity for processing information. Psychological Review, 1956, 63, 81-97.

Miller, G.A., & Selfridge, J.A. Verbal context and the recall of meaningful material. American Journal of Psychology, 1950, 63, 176-185.

Mistler-Lachman, J.L. Depth of comprehension and sentence-memory. Journal of Verbal Learning and Verbal Behavior, 1974, 13, 98-106.

Montague, W.E., Adams, J.E., & Kiess, H.D. Forgetting and natural language mediation. Journal of Experimental Psychology, 1966, 72, 829-833.

Morris, C.D., Bransford, J.D., & Franks, J.J. Levels of processing versus transfer appropriate processing. Journal of Verbal Learning and Verbal Behavior, 1977, 16, 519-533.

Moscovitch, M., & Craik, F.I.M. Depth of processing, retrieval cues, and uniqueness of encoding as factors in recall. Journal of Verbal Learning and Verbal Behavior, 1976, 15, 447-458.

Nelson, T.O. Repetition and depth of processing. Journal of Verbal Learning and Verbal Behavior, 1977, 16, 151-171.

Noble, C.E. The role of stimulus meaning (m) in serial verbal learning. Journal of Experimental Psychology, 1952, 43, 437-446.

Olson, D.R. Language and thought: Aspects of a cognitive theory of semantics. Psychological Review, 1970, 77, 257-273.

Paivio, A., Yuille, J.C., & Madigan, S.A. Concreteness, imagery, and meaningfulness values for 925 nouns. Journal of Experimental Psychology Monograph Supplement, 1968, 76, No. 1, Part 2.

Pavlov, I. Conditioned reflexes. Oxford, England: Oxford University Press, 1927.

Pelton, L.H. Mediational construction vs. mediational perception in paired-associate learning. Psychonomic Science, 1969, 17, 220-221.

Postman, L. A pragmatic view of organization theory. In E. Tulving and W. Donaldson (Eds.), Organization of memory. New York: Academic Press, 1972.

Postman, L. Verbal learning and memory. Annual Review of Psychology, 1975, 37, 291-335.

Rabinowitz, J.C., Mandler, G., & Barsalou, L.W. Recognition failure: Another case of retrieval failure. Journal of Verbal Learning and Verbal Behavior, 1977, 16, 639-663.

Rabinowitz, J.C., Mandler, G., & Patterson, K.E. Determinants of recognition and recall: Accessibility and generation. Journal of Experimental Psychology: General, 1977, 106, 302-329.

Raeburn, V.P. The role of the verb in sentence memory. Memory & Cognition, 1979, 7, 133-140.

Raye, C.L., Johnson, M.K., & Taylor, T.H. Is there something special about memory for internally generated information? Memory & Cognition, 1980, 8, 141-148.

Restle, F. Critique of pure memory. In R.L. Solso (Ed.), Theories in cognitive psychology: The Loyola Symposium. Potomac, Md.: Lawrence Erlbaum Associates, 1974.

Rickards, J.P., & August, G.J. Generative underlining strategies in prose recall. Journal of Educational Psychology, 1975, 67, 860-865.

Rogers, T.B., Kuiper, N.A., & Kirker, W.C. Self-reference and the encoding of personal information. Journal of Personality and Social Psychology, 1977, 35, 677-688.

Rumelhart, D. Toward an interactive model of reading. In S. Dornic (Ed.), Attention and Performance, Vol. 6. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1977.

Schulman, A.L. Memory for words recently classified. Memory & Cognition, 1974, 2, 47-52.

Schwartz, M. Subject-generated versus experimenter-supplied mediators in paired-associate learning. Journal of Experimental Psychology, 1971, 87, 389-395.

Schwartz, M., & Walsh, M.F. Identical subject-generated and experimenter-supplied mediators in paired-associate learning. Journal of Experimental psychology, 1974, 103, 878-884.

Skinner, B.F. The behavior of organisms. New York: Appleton-Century-Crofts, 1938.

Slamecka, N.J. The question of associative growth in the learning of categorized material. Journal of Verbal Learning and Verbal Behavior, 1972, 11, 324-332.

Slamecka, N.J., & Graf, P. The generation effect: Delineation of a phenomenon. Journal of Experimental Psychology: Human Learning and Memory, 1978, 4, 592-604.

Smith, F. Psycholinguistics and reading. New York: Holt, Rinehart, and Winston, 1973.

Stolurow, L.M., & Walker, C.C. A comparison of overt and covert responses in programmed learning. Journal of Educational Research, 1962, 55, 421-429.

Tulving, E. Subjective organization and effects of repetition in multitrial free-recall learning. Journal of Verbal Learning and Verbal Behavior, 1966, 5, 193-197.

Tulving, E. Subjective organization in free recall of "unrelated" words. Psychological Review, 1962, 69, 344-354.

Tulving, E., & Pearlstone, Z. Availability versus accessibility of information in memory for word. Journal of Verbal Learning and Verbal Behavior, 1966, 5, 381-391.

Tulving, E., & Thomson, D.M. Encoding specificity and retrieval processes in episodic memory. Psychological Review, 1973, 80, 352-373.

Tulving, E., & Watkins, M.J. Continuity between recall and recognition. American Journal of Psychology, 1973, 86, 739-748.

Walsh, D.A., & Jenkins, J.J. Effects of orienting task on free recall in incidental learning: "Difficulty," "Effort," and "Process" explanations. Journal of Verbal Learning and Verbal Behavior, 1973, 12, 481-488.

Wittrock, M.C. Learning as a generative process. Educational Psychologist, 1974, 11, 87-95.

Wittrock, M.C. The learning by discovery hypothesis. In L.S. Shulman and E.R. Keislar (Eds.), Learning by discovery. Chicago: Rand McNally, 1966, 33-76.

Wittrock, M.C. (Ed.), Learning and instruction. Berkeley, Calif.: McCutchan, 1977.

Wittrock, M.C., & Carter, J.F. Generative processing of hierarchically organized words. American Journal of Psychology, 1975, 88, 489-501.

Wittrock, M.C., Marks, C., & Doctorow, M. Reading as a generative process. Journal of Educational Psychology, 1975, 67, 484-489.

Appendix A

Sentences used in all experiments involving sentences

Meaningful sentences:

Practice Sentences	1	THE WEALTHY LANDLORD SMOKED THE PIPE
	2	THE ANTIQUE RING PLEASED THE BRIDE
	3	THE LARGE WAREHOUSE SOLD THE MATERIAL
	4	THE DUSTY BOOK CONTAINED THE RIDDLE
	5	THE FLYING SAUCER CARRIED THE CAPTIVES
	6	THE JOKING FOOL AMUSED THE KING
	7	THE CLUMSY ACTOR MISSED THE CUE
	8	THE LONG DISCUSSION TIRED THE AUDIENCE
	9	THE ALERT GUARD NOTICED THE TRESPASSER
	10	THE HATEFUL BANKER CALLED THE LAWYER
Filler Sentences	11	THE HEAVY CARPET PROTECTED THE FLOOR
	12	THE SQUIRMING WORM ATTRACTED THE FISH
Critical Sentences	13	THE ANGRY MOTHER RETURNED THE GIFT
	14	THE GOLDEN KEY LOCKED THE DOOR
	15	THE KIND PRIEST DRANK THE TEA
	16	THE COLD WATER FILLED THE BOTTLE
	17	THE HUNGRY CHILD PLAYED THE PIANO
	18	THE REMOTE COTTAGE SHELTERED THE TRAVELLER
	19	THE BLOND GIRL BAKED THE CAKE
	20	THE POPULAR MAGAZINE REPORTED THE MURDER
	21	THE WISE JUDGE READ THE LEAFLET
	22	THE FALLING TREE CRUSHED THE TENT
	23	THE OLD MAN GRASPED THE CANE
	24	THE CLEVER FOX FOLLOWED THE TRAIL
	25	THE CHARMING MAID POLISHED THE GLASS
	26	THE SOFT SNOW COVERED THE ROAD
	27	THE HANDSOME PILOT ENTERED THE HOSPITAL
	28	THE PIERCING SCREAM WOKE THE NEIGHBOUR
	29	THE PASSING CAR STRUCK THE BICYCLE
	30	THE TIMID HAMSTER ATE THE CHEESE
	31	THE BLAZING FIRE HEATED THE ROOM
	32	THE RUDE SALESMAN INSULTED THE CUSTOMER
	33	THE SMALL ROCK BROKE THE WINDOW
	34	THE ANCIENT CASTLE OVERLOOKED THE BAY

35 THE SMART STUDENT QUESTIONED THE TEACHER
36 THE COMFORTABLE CHAIR RELAXED THE PATIENT
37 THE CUNNING DETECTIVE SOLVED THE CRIME
38 THE EMPTY VASE DECORATED THE TABLE
39 THE DIM LIGHT BRIGHTENED THE HALLWAY
40 THE IMPATIENT BOSS DISMISSED THE EMPLOYEE
41 THE YELLOW BIRD SMELLED THE FLOWER
42 THE POOR WRITER SIGNED THE CONTRACT
43 THE CAUTIOUS BEAR WATCHED THE TRAP
44 THE FEEBLE GRANDMOTHER LEFT THE HOUSE
45 THE FAMOUS MUSICIAN BOUGHT THE PAINTING
46 THE BRAVE POLICEMAN STOPPED THE THIEF
47 THE STORMY WIND SLAMMED THE GATE
48 THE CUTE KITTEN TANGLED THE KNITTING
49 THE YOUNG DOCTOR FOLDED THE NEWSPAPER
50 THE SLOW TURTLE ESCAPED THE FISHERMAN
51 THE BUSY SCIENTIST DISCOVERED THE DRUG
52 THE WITTY CLOWN TRICKED THE SOLDIER
53 THE DIRTY TRUCK OBSTRUCTED THE TRAFFIC
54 THE FRIENDLY CARPENTER BUILT THE SHACK
55 THE TINY MOUSE FRIGHTENED THE COOK
56 THE CHEERFUL TUNE CONSOLED THE TEAM
57 THE HEAVY CARPET PROTECTED THE FLOOR
58 THE SQUIRMING WORM ATTRACTED THE FISH
59 THE STRONG COFFEE DELIGHTED THE GUEST
60 THE CLINGING VINE FRAMED THE POSTER

Anomalous sentences:

Practice Sentences	1	THE WEALTHY RIDDLE SMOKED THE SAUCER
	2	THE FLUID WAREHOUSE CARRIED THE RING
	3	THE DUSTY PIPE PRODUCED THE CAPTIVES
	4	THE LARGE MATERIAL CONTAINED THE LANDLORD
	5	THE ANTIQUE ISLAND CLEARED THE BRIDE
	6	THE LONG AUDIENCE ERASED THE FOOL
	7	THE THICK DISCUSSION PACED THE PRISON
	8	THE HATEFUL FIELD FLUSHED THE BANKER
	9	THE FLOWING GUARD RULED THE TIRE
	10	THE THIRSTY PAGE NOTICED THE CUE
Filler Sentences	11	THE HANDY SCRATCH MOTIONED THE HEADLINE
	12	THE DREAMING RANSOM STAMPED THE LOGIC
Critical Sentences	13	THE ANGRY ROAD RETURNED THE MAID
	14	THE GOLDEN NEIGHBOUR CRUSHED THE WATER
	15	THE KIND CAKE LOCKED THE CHILD
	16	THE COLD PIANO READ THE TRAVELLER
	17	THE HUNGRY COTTAGE DRANK THE PRIEST
	18	THE REMOTE MURDER GRASPED THE TREE
	19	THE BLOND LEAFLET FILLED THE PILOT
	20	THE POPULAR TENT FOLLOWED THE MOTHER
	21	THE WISE CANE PLAYED THE GIRL
	22	THE FALLING GIFT POLISHED THE MAGAZINE
	23	THE CLEVER DOOR SHELTERED THE SNOW
	24	THE OLD GLASS BAKED THE JUDGE
	25	THE CHARMING TRAIL ENTERED THE KEY
	26	THE SOFT TEA COVERED THE FOX
	27	THE HANDSOME HOSPITAL REPORTED THE SCREAM
	28	THE PIERCING MAN WOKE THE BOTTLE
	29	THE PASSING CHEESE QUESTIONED THE SALESMAN
	30	THE CAUTIOUS FIRE OVERLOOKED THE CAR
	31	THE BLAZING CUSTOMER SOLVED THE ROCK
	32	THE RUDE BIRD BROKE THE HALLWAY
	33	THE SMALL BAY SIGNED THE DETECTIVE
	34	THE FEEBLE WINDOW SMELLED THE STUDENT
	35	THE ANCIENT CHAIR HEATED THE GRANDMOTHER
	36	THE SMART HOUSE BRIGHTENED THE EMPLOYEE
	37	THE COMFORTABLE FLOWER STRUCK THE TABLE
	38	THE CUNNING CONTRACT LEFT THE CASTLE
	39	THE EMPTY TRAP INSULTED THE WRITER

40 THE DIM HAMSTER DECORATED THE TEACHER
41 THE IMPATIENT CRIME ATE THE LIGHT
42 THE YELLOW PATIENT DISMISSED THE ROOM
43 THE POOR VASE RELAXED THE BOSS
44 THE TIMID BICYCLE WATCHED THE BEAR
45 THE CUTE NEWSPAPER SLAMMED THE FISHERMAN
46 THE STORMY PAINTING BUILT THE THIEF
47 THE BRAVE KNITTING STOPPED THE SOLDIER
48 THE FAMOUS FLOOR BOUGHT THE SCIENTIST
49 THE CLINGING POSTER PROTECTED THE FISH
50 THE STRONG WORM ATTRACTED THE VINE
51 THE SQUIRMING POLICEMAN FRIGHTENED THE TRUCK
52 THE HEAVY TEAM FOLDED THE GUEST
53 THE CHEERFUL CARPET ESCAPED THE MOUSE
54 THE TINY TUNE OBSTRUCTED THE CARPENTER
55 THE FRIENDLY COFFEE TRICKED THE DOCTOR
56 THE DIRTY WIND CONSOLED THE TURTLE
57 THE WITTY TRAFFIC DISCOVERED THE COOK
58 THE BUSY DRUG FRAMED THE KITTEN
59 THE SLOW SHACK TANGLED THE CLOWN
60 THE YOUNG GATE DELIGHTED THE MUSICIAN

Rhyming quadruples used in Experiment 12

Study Items	1	CAGE-PAGE-RAGE-WAGE
	2	CARE-DARE-FARE-RARE
	3	FARM-HARM-CHARM-ARM
	4	RING-KING-SING-WING
	5	MAT-SAT-CAT-RAT
	6	BRIGHT-NIGHT-LIGHT-TIGHT
	7	SCORE-CORE-BORE-MORE
	8	LAST-FAST-CAST-MAST
	9	PACE-LACE-FACE-RACE
	10	SLOW-GLOW-SNOW-ROW
	11	FATE-MATE-HATE-LATE
	12	FAME-SAME-GAME-NAME
	13	CHEAT-MEAT-SEAT-HEAT
	14	NEAR-DEAR-CLEAR-FEAR
	15	FIT-BIT-SIT-HIT
	16	BACK-LACK-PACK-SACK
	17	MAN-CAN-PAN-FAN
	18	MOON-NOON-SPOON-SOON
	19	CAMP-DAMP-RAMP-LAMP
	20	ROCK-SOCK-LOCK-DOCK
	21	HILL-MILL-BILL-FILL
	22	FAIL-SAIL-PAIL-TAIL
	23	HAND-BAND-LAND-SAND
	24	BELL-CELL-FELL-TELL
Distractor items	25	READ-DEAD-LEAD-HEAD
	26	REST-BEST-WEST-TEST
	27	WRONG-LONG-ALONG-SONG
	28	GLOOM-DOOM-BROOM-ROOM
	29	BEND-END-MEND-LEND
	30	SLEEP-DEEP-CHEAP-HEAP
	31	BALL-WALL-FALL-HALL
	32	CATCH-MATCH-PATCH-HATCH
	33	GRAIN-TRAIN-PLAIN-PAIN
	34	BAD-MAD-PAD-SAD
	35	FOLD-MOLD-COLD-HOLD
	36	BET-WET-LET-PET
	37	DAY-BAY-MAY-SAY
	38	FROWN-DROWN-CROWN-BROWN
	39	FAR-BAR-CAR-TAR
	40	FIND-MIND-KIND-BIND

Practice items	1	MILD-WILD
	2	TIME-DIME
	3	CREAM-DREAM
	4	ROUGH-TOUGH
	5	SNEEZE-FREEZE
	6	DITCH-WITCH
	7	BEAST-EAST
	8	LOVE-DOVE
	9	HOOK-BOOK
	10	SILK-MILK

Appendix B

Analysis of variance: Cued recall of nouns by study material and processing condition in Experiment 1

Source of Variance	Sum of Squares	Degrees of Freedom	Mean Square	F Ratio	Significance Level
Between subjects					
A) Study material	261.333	1	261.333	36.18	<.01
Subjects w. groups	158.917	22	7.223		
Within subjects					
B) Processing condition	18.750	1	18.750	4.89	<.05
Interaction of A x B	48.000	1	48.000	12.53	<.01
B x Subjects w. groups	84.250	22	3.830		
Total	571.250	47	12.154		

Analysis of variance: Cued recall of object nouns by study material and processing condition in Experiment 2

Source of Variance	Sum of Squares	Degrees of Freedom	Mean Square	F Ratio	Significance Level
Between subjects					
A) Study material	65.333	1	65.333	37.17	<.01
Subjects w. groups	38.667	22	1.758		
Within subjects					
B) Processing condition	8.333	1	8.333	7.24	<.01
Interaction of A x B	40.333	1	40.333	35.03	<.01
B x Subjects w. groups	25.333	22	1.152		
Total	178.000	47	3.787		

Analysis of variance: Cued recall of nouns by study material and processing condition in Experiment 3

Source of Variance	Sum of Squares	Degrees of Freedom	Mean Square	F Ratio	Significance Level
Between subjects					
A) Study material	305.021	1	305.021	33.76	<.01
Subjects w. groups	198.791	22	9.036		
Within subjects					
B) Processing condition	82.687	1	82.687	17.90	<.01
Interaction of A x B	42.187	1	42.187	9.13	<.01
B x Subjects w. groups	101.625	22	4.619		
Total	730.311	47	15.538		

Analysis of variance: Cued recall of nouns by study material, trials and processing condition in Experiment 4

Source of Variance	Sum of Squares	Degrees of Freedom	Mean Square	F Ratio	Significance Level
Between subjects					
A) Study material	1418.771	1	1418.771	36.09	<.01
Subjects w. groups	864.858	22	39.31		
Within subjects					
B) Trials	742.593	2	371.297	69.55	<.01
Interaction of A x B	37.514	2	18.757	3.51	<.05
B x Subjects w. groups	234.888	44	5.338		
C) Processing condition	35.999	1	35.999	3.64	<.08
Interaction of A x C	117.360	1	117.360	11.86	<.01
C x Subjects w. groups	217.638	22	9.894		
Interaction of B x C	20.292	2	10.146	4.21	<.05
Interaction of A x B x C	4.597	2	2.299	.95	--
B x C x Subjects w. groups	106.109	44	2.412		
Total	3800.619	143	26.578		

Analysis of variance: d' scores for recognition by study material, groups, and processing condition in Experiment 5

Source of Variance	Sum of Squares	Degrees of Freedom	Mean Square	F Ratio	Significance Level
Between subjects					
A) Groups	16.376	1	16.376	14.79	<.01
B) Study material	80.319	1	80.319	72.52	<.01
Interaction of A x B	6.568	1	6.568	5.93	<.05
Subjects w. groups	48.730	44	1.107		
Within subjects					
C) Processing condition	.812	1	.812	.99	--
Interaction of A x C	1.195	1	1.195	1.47	--
Interaction of B x C	11.683	1	11.683	14.35	<.01
Interaction of A x B x C	.569	1	.569	.70	--
C x Subjects w. groups	35.834	44	.814		
Total	202.086	95	2.127		

Analysis of variance: Free and cued recall of nouns by processing condition in Experiment 6

Source of Variance	Sum of Squares	Degrees of Freedom	Mean Square	F Ratio	Significance Level
Subjects	221.438	15	14.762		
A) Processing condition	342.250	1	342.250	27.19	<.01
A x Subjects w. groups	188.750	15	12.583		
B) Test type	126.563	1	126.563	62.37	<.01
B x Subjects w. groups	30.438	15	2.029		
Interaction of A x B	30.250	1	30.250	6.90	<.01
A x B x Subjects w. groups	65.750	15	4.383		
Total	1005.438	63	15.959		

Analysis of variance: Recognition of nouns by study material and processing condition in Experiment 7

Source of Variance	Sum of Squares	Degrees of Freedom	Mean Square	F Ratio	Significance Level
Between subjects					
A) Study material	12.250	1	12.250	.78	--
Subjects w. groups	470.188	30	15.673		
Within subjects					
B) Processing condition	240.250	1	240.250	25.77	<.01
Interaction of A x B	5.063	1	5.063	.54	--
B x Subjects w. groups	279.688	30	9.323		
Total	1007.438	63	15.991		

Analysis of variance: Recognition of nouns by study material and processing condition in Experiment 8

Source of Variance	Sum of Squares	Degrees of Freedom	Mean Square	F Ratio	Significance Level
Between subjects					
A) Study material	.183	1	.183	.27	--
Subjects w. groups	20.213	30	.674		
Within subjects					
B) Processing condition	3.294	1	3.294	18.72	<.01
Interaction of A x B	.131	1	.131	.75	--
B x Subjects w. groups	5.280	30	.176		
Total	29.102	63	.462		

Analysis of variance: Recognition of nouns by processing condition and test context in Experiment 9

Source of Variance	Sum of Squares	Degrees of Freedom	Mean Square	F Ratio	Significance Level
Between subjects					
A) Processing condition	10.033	1	10.033	18.00	<.01
Subjects w. groups	16.722	30	.557		
Within subjects					
B) Test context	1.464	1	1.464	20.29	<.01
Interaction of A x B	1.035	1	1.035	14.35	<.01
B x Subjects w. groups	2.165	30	.072		
Total	31.419	63	.499		

Analysis of variance: Recognition of nouns by processing condition and test context in Experiment 10

Source of Variance	Sum of Squares	Degrees of Freedom	Mean Square	F Ratio	Significance Level
Between subjects					
A) Processing condition	5.452	1	5.452	13.26	<.01
Subjects w. groups	12.332	30	.411		
Within subjects					
B) Test context	.137	1	.137	1.97	--
Interaction of A x B	.002	1	.002	.03	--
B x Subjects w. groups	2.090	30	.070		
Total	20.013	63	.318		

Analysis of variance: Recognition of nouns by test order, processing condition and test context in Experiment 11

Source of Variance	Sum of Squares	Degrees of Freedom	Mean Square	F Ratio	Significance Level
Between subjects					
A) Test order	.675	1	.675	.81	--
Subjects w. groups	18.413	22	.837		
Within subjects					
B) Processing condition	26.471	1	26.471	59.67	<.01
Interaction of A x B	.744	1	.744	1.68	--
B x Subjects w. groups	9.760	22	.444		
C) Test context	9.940	1	9.940	60.58	<.01
Interaction of A x C	.092	1	.092	.56	--
C x Subjects w. groups	3.609	22	.164		
Interaction of B x C	2.724	1	2.724	7.74	<.01
Interaction of A x B x C	.006	1	.006	.02	--
B x C x Subjects w. groups	7.739	22	.352		
Total	80.166	95	.844		

Analysis of variance: Recognition of nouns by test order, processing condition and test context in Experiment 12

Source of Variance	Sum of Squares	Degrees of Freedom	Mean Square	F Ratio	Significance Level
Between subjects					
A) Test order	.333	1	.333	.24	--
Subjects w. groups	31.002	22	1.409		
Within subjects					
B) Processing condition	4.386	1	4.386	5.49	<.05
Interaction of A x B	.076	1	.076	.10	--
B x Subjects w. groups	17.582	22	.799		
C) Test context	11.138	1	11.138	53.86	<.01
Interaction of A x C	.232	1	.232	1.13	--
C x Subjects w. groups	4.550	22	.207		
Interaction of B x C	.143	1	.143	.70	--
Interaction of A x B x C	.032	1	.032	.16	--
B x C x Subjects w. groups	4.504	22	.205		
Total	73.979	95	.779		

7

Analysis of variance: Recognition of nouns by test order, study typography and test context in Experiment 13

Source of Variance	Sum of Squares	Degrees of Freedom	Mean Square	F Ratio	Significance Level
Between subjects					
A) Test order	.015	1	.015	.02	--
Subjects w. groups	10.289	14	.735		
Within subjects					
B) Typography	10.168	1	10.168	31.45	<.01
Interaction of A x B	.173	1	.173	.54	--
B x Subjects w. groups	4.527	14	.323		
C) Test context	4.725	1	4.725	36.54	<.01
Interaction of A x C	.089	1	.089	.69	--
C x Subjects w. groups	1.811	14	.129		
Interaction of B x C	1.280	1	1.280	6.54	<.05
Interaction of A x B x C	.157	1	.157	.81	--
B x C x Subjects w. groups	2.740	14	.196		
Total	35.974	63	.571		

Analysis of variance: Recognition of nouns by test order, study typography and test context in Experiment 14

Source of Variance	Sum of Squares	Degrees of Freedom	Mean Square	F Ratio	Significance Level
Between subjects					
A) Test order	.044	1	.044	.07	--
Subjects w. groups	8.817	14	.630		
Within subjects					
B) Typography	4.091	1	4.091	31.24	<.01
Interaction of A x B	.176	1	.176	1.35	--
B x Subjects w. groups	1.833	14	.131		
C) Test context	.139	1	.139	2.85	--
Interaction of A x C	.281	1	.281	5.77	<.05
C x Subjects w. groups	.681	14	.049		
Interaction of B x C	.032	1	.032	.43	
Interaction of A x B x C	.038	1	.038	.52	
B x C x Subjects w. groups	1.029	14	.074		
Total	17.161	63	.272		

Appendix C

Experiment on the persistence of the generation effect over study-test trials

The central purpose of this experiment was to examine whether the generation effect would persist if subjects were given the opportunity to monitor performance in the Generate and Read condition. Experiment 4 indicated that while there is a generation effect after one study-test trial, the rate of learning in the two processing conditions was similar subsequently. This finding suggested that the test experience may alert subjects to the ineffectiveness of reading as a learning strategy and it may motivate him or her to try "harder" in the Read condition later. Based on this suggestion it is expected that if subjects studied and were tested on one set of materials, and then on another set of materials, there may no longer be a generation effect on the second test. This notion was examined in the present experiment. An additional purpose of the experiment was to test whether paced study (as in all the other experiments) as compared to self-paced study is critical for obtaining the pattern of results characterizing the

remaining experiments. Only the meaningful sentences were used in this experiment.

Method

Subjects and Design. Thirty-two subjects were used in this experiment, 16 of which were randomly assigned to the Paced group and the reamining 16 to the Self-paced group. The design included the processing condition (generate and read) as a within-subjects factor and the groups (paced and self-paced) as a between-subjects factor.

Material and Procedure. The study list consisted of 32 meaningful sentences arranged into two sets of 16 each. In each set eight were shown in each presentation format. The sentences were counterbalanced such that each occurred equally often in set 1 and in set 2. The procedure exposed each subject once to each sentence set following each with a cued recall test. The studying and testing of the two sets was separated by a five minute interval during which subjects were engaged in conversation. The verbs were used as cues on the recall test. The instructions given to the Self-paced group were identical to those received by the Paced group excepting that the former group was required to press a key to advance from one sentence to the next. Subjects were asked to press the key when they had finished reading a displayed sentence or generating one with the words provided on the CRT.

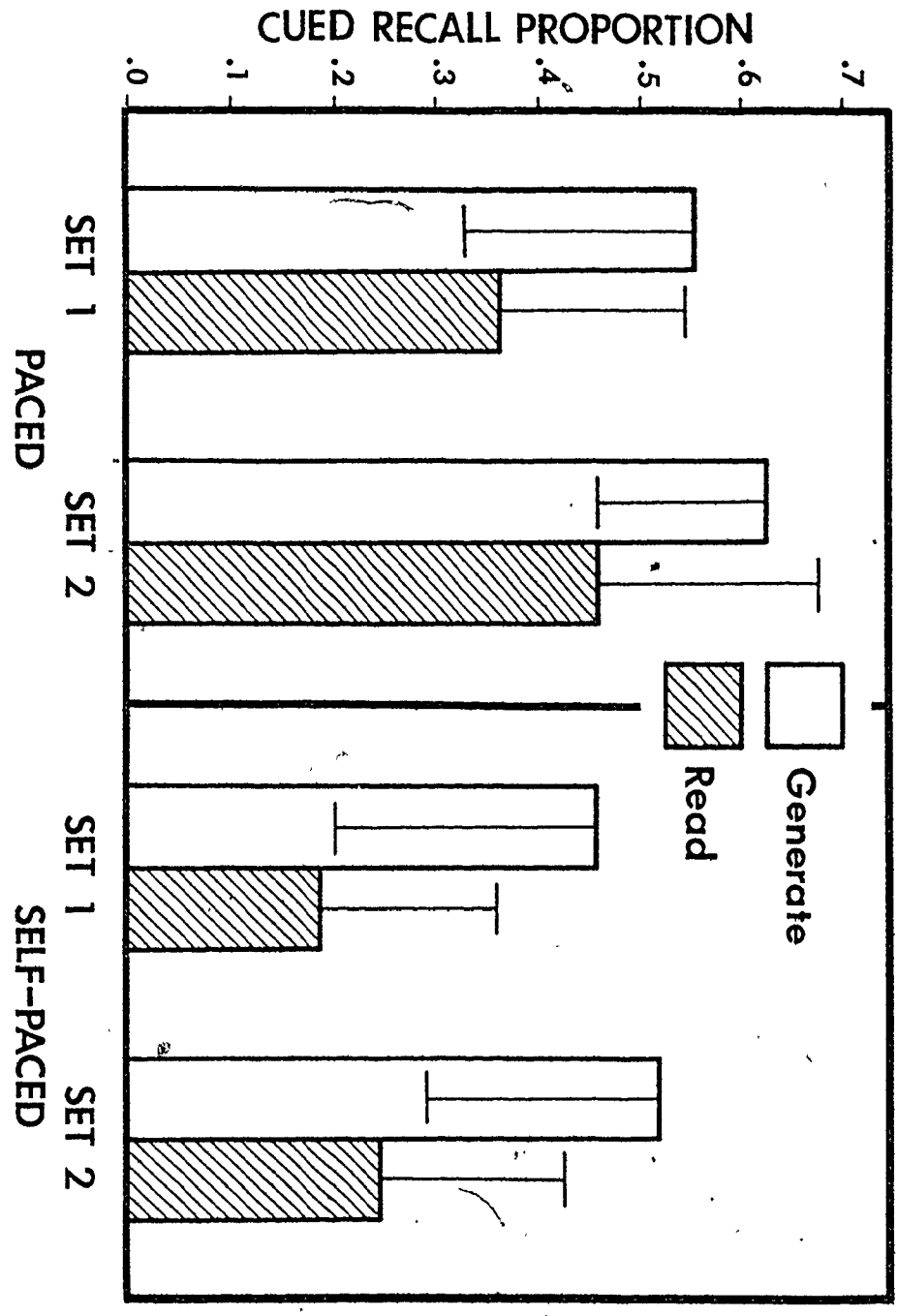
Results and Discussion

The recall protocols were scored and summarized as in Experiment 1. The mean number of nouns recalled in each condition is displayed in Figure C.1. The figure shows (a) a substantial generation effect on each set in each group, (b) better performance in the Paced group than in the Self-paced group, and (c) better performance on the second set than on the first set for both groups.

The data were submitted to an analysis of variance. The analysis treated groups (paced and self-paced) as a between-subjects factor and both set and processing condition (generate and read) as within-subjects factors. The results of the analysis are presented in Table C.1. The analysis confirmed observations showing a main effect for processing condition, $F(1,30) = 46.05$, $MSe = 8.97$, for group, $F(1,30) = 8.01$, $MSe = 21.94$, and for set, $F(1,30) = 9.68$, $MSe = 4.42$, with no other effects approaching significance.

The observed generation effects replicate the findings from other experiments on interword organization, and they extend these findings to a self-paced study situation. The lower performance in the Self-paced group, as compared to the Paced group, might be due to a difference in study time. While the paced subjects were given eight seconds to study each sentence, the self-paced subjects allowed themselves only an average 6.9 seconds to generate each sentence and 4.4 seconds to read each sentence. The increased performance on the second set, as compared to the first set, suggests that subjects try harder

Figure C.1. Recall of nouns (means and standard deviations) as a function of group, set, and processing condition



after having been tested or that the first test experience teaches them how to study the sentences more effectively for the second test. Most importantly, however, the presence of a generation effect on both sets indicates that the test experience does not differentially affect the learning strategies that subjects may be using in the Read and Generate condition.

In Experiment 4, where the same sentences were studied over three study-test trials, it was observed that the rate of learning was identical in the two processing conditions after one study-test trial. The present results indicate that the outcome of Experiment 4 cannot be attributed to the notion that subjects change the learning strategies used in the Generate and Read condition following the test experience. Instead, this experiment suggest that generating as compared to reading a sentence will aid performance only if the same sentence has not been encountered shortly before (see Jacoby, 1978). The sentences read and generated on the second and third study-test trial of Experiment 4 had all been encountered during the previous study-test trial. Thus the rate of learning was the same in the Generate and Read condition. The sentences used in set 2 in the present experiment had not been studied in set 1, and subjects learned more in the Generate condition than in the Read condition.

Table C.1

Analysis of variance: Cued recall of nouns by group, set, and processing condition

Source of Variance	Sum of Squares	Degrees of Freedom	Mean Square	F Ratio	Significance Level
Between subjects					
A) Group	175.781	1	175.781	8.01	<.01
Subjects w. groups	658.219	30	21.941		
Within subjects					
B) Set	42.781	1	42.781	9.68	<.01
Interaction of A x B	1.125	1	1.125	.25	--
B x Subjects w. groups	132.594	30	4.420		
C) Processing condition	413.281	1	413.281	46.05	<.01
Interaction of A x C	18.000	1	18.000	2.01	--
C x Subjects w. groups	269.219	30	8.974		
Interaction of B x C	.125	1	.125	.02	--
Interaction of A x B x C	.281	1	.281	.04	--
B x C x Subjects w. groups	228.094	30	7.603		
Total	1939.500	127	15.272		