51

NAME OF AUTHOR/*NOM DE L'AUTEUR* Marta Louise Holmberg

TITLE OF THESIS/*TITRE DE LA THÈSE* "The Effect of Specificity of Encoding on Human Memory"

UNIVERSITY/*UNIVERSITÉ* McMaster

DEGREE FOR WHICH THESIS WAS PRESENTED/ *GRADE POUR LEQUEL CETTE THÈSE FUT PRÉSENTÉE* Ph.D.

YEAR THIS DEGREE CONFERRED/*ANNÉE D'OBTENTION DE CE DEGRÉ* 1983

NAME OF SUPERVISOR/*NOM DU DIRECTEUR DE THÈSE* Dr. I.M. Begg

DATED/*DATÉ* January 31, 1983 SIGNED/*SIGNÉ* Marta L. Holmberg

PERMANENT ADDRESS/*RÉSIDENCE FIXÉ* 401 Burr Oak Rd. Unit 103B
Hinsdale, Illinois 60521
USA

THE EFFECT OF SPECIFICITY OF ENCODING ON HUMAN MEMORY

By

© MARTA LOUISE HOLMBERG

A Thesis

Submitted to the Faculty of Graduate Studies

in Partial Fulfillment of the Requirements

for the Degree

Doctor of Philosophy

McMaster University

(September) 1982

THE EFFECT OF SPECIFICITY OF ENCODING ON HUMAN MEMORY

DOCTOR OF PHILOSOPHY (1982)

McMASTER UNIVERSITY
Hamilton, Ontario

TITLE: The effect of specificity of encoding on human memory

AUTHOR: Marta Louise Holmberg, B.A. (Oberlin College)

SUPERVISOR: Professor Ian Begg

NUMBER OF PAGES: (x), 161

ABSTRACT

The thesis examines the effect that specificity of encoding has on human memory. Specificity of encoding is defined as the precision and detail of stimulus analysis required to fulfill the purpose of the encoding task. The memory trace is characterized as the description of the analysis used to interpret the stimulus event. Greater specificity of encoding influences the quality of the memory trace by requiring a more complete and specific description of the stimulus to differentiate the stimulus in the task environment. The resulting memory trace is more distinctive and generally is more accessible and more discriminable from alternative traces.

The experiments presented demonstrate that specificity of encoding in itself is not sufficient to ensure increased memory performance. Experiment 1 demonstrates that specificity of encoding enhances cued recall and recognition, and the cohesion of the memory trace, when the increased specificity stresses semantic, but not structural, aspects of the stimulus. The differential effect of specificity is attributable to the specificity of encoding and not to the cognitive effort expended during encoding. Experiment 2 demonstrates that specificity of encoding may occur without accompanying completeness of encoding, resulting in a relatively inaccessible (free recall), but cohesive, memory trace.

Experiments 3 and 4 demonstrate that the potential distinctiveness of specific encoding may be masked by incongruity of encoding. Specificity of encoding results in a cohesive, retrievable

memory trace. Despite the fact that incongruity interferes with specificity, such specificity may still result in higher recall than does a more general encoding for meaning. Experiment 5 demonstrates that the potential effects of specificity of encoding are severely undermined when incongruity leads to a superficial semantic encoding. Experiment 6 establishes that specificity of encoding is critical to the formation of a distinctive memory trace, but that the accessibility and discriminability of the trace may be hampered by a loss of specificity or completeness in the retrieval context. However, the cohesion of the memory trace is determined by the encoding and is not influenced by a change in context at the time of the memory test.

ACKNOWLEDGMENTS

I wish to express heartfelt thanks to my advisor and friend, Dr. Ian Begg, for the able counsel and constant encouragement he has offered me through all phases of this investigation and graduate career. I would also like to thank Dr. Lee Brooks and Dr. Larry Jacoby for suggestions and criticisms that have improved the quality of this work. Finally, I would like to thank Dr. Roy Pritchard for his kind assistance.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF TABLES

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION

The encoding of a stimulus is an efficient process. The purpose of the encoding (e.g., identification, or determining the significance of the stimulus) and the demands established by the encoding context primarily determine the nature (e.g., whether the analysis focuses on phonemic or semantic aspects of the stimulus) and the specificity of the encoding. If the purpose of the encoding "pineapple" is to determine how many letters the word contains, then the encoding of "pineapple" in that context does not require any consideration of meaning; a meaningful analysis is irrelevant to the task demands.

A key aspect of the encoding is specificity. (The account of specificity of encoding presented here is essentially an adaptation of the account of distinctiveness of encoding presented by Jacoby, Craik, and Begg, 1979.) Specificity of encoding is defined as the precision and detail of analysis required to fulfill the purpose of the encoding. A primary goal of stimulus analysis is that of distinguishing the stimulus event from plausible alternatives in the environment. As stimuli are identified in the physical environment by finding distinctive features that differentiate them from similar alternatives, so are stimuli identified in the cognitive environment. The more similar an alternative is to the intended stimulus, the more specific and complete is the analysis needed to distinguish the two. It is relatively easy to distinguish "wine" from "rabbit" by making use of general contrasts such as "liquid" or "inanimate." Such general distinctions will not suffice to distinguish "wine" from "vodka." In the latter case, more specific distinctions such as "made from grapes" and "10-14% alcohol level" are required to discriminate between the two alternatives; the context requires greater specificity of encoding.

Specificity of encoding is considered to have a major impact on the memory trace. The memory trace is characterized as a description of the analysis used to interpret a stimulus event. The demands of an encoding task influence the quality of the resulting memory trace by focusing on particular aspects of the stimulus, with those aspects possibly ranging from physical features to subtle connotations of meaning. Greater specificity of encoding influences the nature of the memory trace by requiring that a more specific and complete description of the stimulus be utilized to differentiate the stimulus in the task environment. The greater specificity of the memory trace makes it distinctive and readily discriminable by severely limiting the set of plausible alternatives. In general, greater specificity of encoding results in memory traces that are increasingly memorable.

The goal of the thesis is to delineate the extent to which specificity of encoding affects memory. The thesis determines those aspects of memory that are influenced by specificity of encoding. In addition, critical limits are established for the effect that specificity required by the encoding task has on the distinctiveness and retrievability of the memory trace.

Specificity of encoding is manipulated in two different ways in the course of the thesis. All of the experiments in the thesis use pairs of words as stimuli. The first four experiments use pairs of words varied as to similarity along a particular dimension (e.g., referent size or pleasantness). In the first two experiments, subjects choose the greater (e.g., in terms of referent size) member of a pair (e.g., SNAIL, PEANUT) and rate the difficulty of the decision. The more similar pairs require a more difficult decision; it will be shown that the greater the similarity, the more specific is the encoding of a pair. A different measure of specificity of encoding is introduced in the third and fourth experiments; subjects directly rate the similarity of the word pair along the specified dimension. The last two experiments manipulate specificity more directly by varying the specificity of the encoding question for a stimulus pair.

The thesis begins to delineate the effect of specificity of encoding by determining that specificity of encoding is a more adequate explanation of the memory process than is an alternate account that stresses processing capacity rather than the nature of the encoding. When specificity of encoding is referentially based (e.g., referent size) rather than nominally based (e.g., word length), the resulting memory trace is recognizable, recallable, and cohesive to a degree corresponding to the relative specificity called for in the encoding task. Specificity of encoding is demonstrated to be more predictive of the level and pattern of memory performance than is processing capacity, or cognitive effort. A detailed comparison of the two accounts is presented in the context of the appropriate experiment, Experiment 1.

The thesis provides an increasingly refined picture of specificity of encoding. Experiment 2 extends the generality of the effect that specificity of encoding has when the encoding is meaningful rather than structural in nature, with a consideration of synonymy. Even when the level of similarity in meaning is quite high, a relative increase in the specificity of encoding called for by the comparison task results in a higher level of recognition and recall. Experiment 3 determines that specificity of encoding does not stem from requiring subjects to choose one member of a pair, but from the underlying similarity of the pair on the dimension of comparison.

Experiments 3 and 4 examine specificity in conjunction with another critical factor, congruity of encoding. Congruity will be discussed in the context of those experiments. The experiments demonstrate that specificity of encoding results in relatively distinctive, hence retrievable, memory traces only when the dimension of comparison is congruous with the nature of the stimuli. When the encoding task is incongruous with the nature of the stimli, other salient aspects of the stimuli predominate in affecting retrieval. The experiments demonstrate that congruity of encoding does not necessarily result in a better memory for the stimulus events. On a different level, the overall specificity (e.g., referent size vs. general understanding) of encoding called for by the encoding task does affect memory independently of the congruity of encoding. The results considered together support the idea that the level of specificity necessary to complete the encoding task does influence the accessibility and the cohesion of the memory trace.

Experiment 5 demonstrates a different effect of congruity in conjunction with specificity of encoding. If an incongruous encoding task discourages all but a superficial semantic encoding, then those items encoded in an incongruous context are recalled with far less frequency than are stimuli considered in a congruous context. Again, the experiment confirms that a specific encoding is more readily retrievable than a more general encoding only if the encoding task is congruous with the nature of the stimuli being considered.

Experiment 6 compares the impact of specificity at the time of encoding and at the time of test. When the encoding context is specific and the specificity is recreated at the time of the memory test, the memory search is quite successful, with the trace being accessible, discriminable, and cohesive. A loss of completeness or specificity in the context from time of encoding to time of test reduces the accessibility, but not the cohesion, of the memory trace. When the encoding context is general, the accessibility and cohesion of the trace are lessened relative to a more specific encoding, but the specificity of the context at retrieval is relatively unimportant as long as it recreates at least the general encoding context.

The findings of the thesis demonstrate that specificity of encoding is not sufficient to guarantee the formation of a distinctive memory trace. In some cases, encoding may be quite specific without being complete, so that traces are relatively inaccessible, but quite cohesive (e.g., synonymy in Exp. 2). If the context at the time of retrieval does not provide strong clues to one member of a stimulus pair, the trace may be located only infrequently. In addition to completeness, congruity of encoding is crucial if specificity of encoding is to result in a distinctive trace. If encoding is incongruous, the stimuli may remain meaningless or be processed for relatively superficial aspects of meaning. Either effect of incongruity will undermine the potential effect of specificity. Incongruity between the task and the nature of the stimuli may allow other salient aspects of the stimuli to overwhelm the effect of specificity. Finally, if the context at time of retrieval stresses an incomplete (general or specific) aspect of the encoding context, depressing the accessibility of the memory trace, the potential distinctiveness of the memory trace may be masked.

This investigation of specificity of encoding relies on previous work establishing that the quality of encoding is of critical importance to the nature and retrievability of the memory trace. The principle of specificity of encoding may be seen as stemming from two well-developed explanations of stimulus encoding, "levels-of-processing" and distinctiveness of encoding. The development of the underlying principles and significance of specificity of encoding may be clarified by a consideration of its historical context. The remainder of the introduction presents that historical context, with additional explanatory principles being developed in the context of individual experiments.

HISTORICAL CONTEXT

Levels-of-processing investigations have established that stimulus encoding is not an automatic process, carried out to the same degree regardless of the purpose of the encoding. Instead, stimulus encoding is an efficient process, with the nature and depth of analysis depending primarily on the demands of the encoding context. The quality of the encoding influences the nature and the retrievability of the memory trace.

Craik and Lockhart (1972) introduced the levels-of-processing account of human memory to focus attention on the encoding process. They view the memory trace as resulting from perceptual analysis at various levels. A cursory analysis entails examination of physical features. Deeper levels of analysis require the use of past knowledge to identify the stimulus. The deepest levels require a semantic analysis of the stimulus to determine its significance. The nature and depth of stimulus analysis are reflected in the nature, strength, and durability of the memory trace. The authors suggest a "continuum of analysis" (p. 676) rather than a set of discrete stages of encoding. The further along the continuum in the direction of rich, semantic analysis the encoding of a stimulus, the more elaborate, durable, and strong will be the resultant memory trace.

Craik and Lockhart (1972) view memory capacity as being a function of encoding depth. Deeper encoding makes increased use of pre-existing rules and knowledge, resulting in greater efficiency of encoding and an increase in memory capacity. Encoding a stimulus more deeply should increase retention.

Various investigators have substantiated the general finding that the level of initial encoding affects the subsequent level of retrieval (e.g., Coltheart, 1977; Craik and Tulving, 1975; Davies and Cubbage, 1976; Elias and Perfetti, 1973; Epstein, Phillips, and Johnson, 1975; Eysenck, 1974; Hyde and Jenkins, 1969, 1973; Johnston and Jenkins, 1971; Lane and Robertson, 1979; Nelson, Wheeler, Borden, and Brooks, 1974; Parkin, 1979; Postman and Kruesi, 1977; Schulman, 1971; Seamon and Murray, 1976). The findings of such investigators support the tenets of the levels-of-processing framework.

Levels of processing introduced the procedure of examining the particular encoding operation for its effect on memory for the encoded items. Differences in retention can be attributed to precise differences in encoding tasks. The shift in focus from the stimulus itself to the encoding operation emphasizes the complexity and efficiency of the memory process. The tenets established in the levels-of-processing work underlie the manner in which principles such as specificity of encoding or congruity are considered in the experiments presented here.

Although the levels-of-processing framework provides a general account of human memory, "levels" has been criticized for not presenting a complete account that is consistent with empirical findings. Criticisms of the framework suggest that the "levels" account requires revision and extension of its principles to be a plausible explanation of the workings of human memory. Criticisms exposing the theoretical limitations of the levels-of-processing account are discussed here because they point out those aspects of memory that must be accounted for by an alternate theory of human memory. (See Bransford, Morris, and Stein, 1979, and Morris, Bransford, and Franks, 1977, for arguments that standard memory tests are biased in favor of encoding stimuli as semantic units.)

Baddeley (1978) and Eysenck (1978) suggest that a major weakness in the levels framework is the lack of an independent index of depth. To avoid circularity of argument, the account must provide an index of processing depth that can be measured independently of subsequent retention levels. The lack of an independent measure of depth is not crucial when encoding tasks are grossly different in nature, as when the tasks require an orthographic as opposed to a semantic analysis of stimuli. However, the comparison of two semantic tasks cannot use subsequent differences in retention level as the basis for attributing differences in encoding depth. Without an independent measure of depth, specific differences in encoding tasks cannot be explained or predicted, severely limiting the power and scope of an account of memory.

In addition to an independent index of depth, the levels framework requires some additional principles, or modifications, to explain fully data generally consistent with the levels account (Baddeley, 1978). Two such principles are specificity of encoding and congruity (principles examined directly in the thesis). Specificity of encoding here refers to the richer and more detailed encoding a stimulus may receive in the context of a specific semantic decision, as opposed to a general encoding for meaning. Congruity refers to the compatibility between the encoding context and the stimulus being encoded. For example, "Is this a type of footwear?" is a congruous question for the stimulus "sandals", but it is incongruous for the stimulus "heroin". Stimuli that are congruous with the encoding context are better recalled than are stimuli incongruous with the encoding context (Craik and Tulving, 1975). Congruity of stimulus and encoding context may provide a coherent, integrated context for the encoding, resulting in a more fully developed trace than is formed when the meaning of the stimulus is incongruous with the encoding context. The concepts of specificity and congruity of encoding are not inconsistent with levels; however, neither do they necessarily follow from its principles. An account stressing distinctiveness of encoding encompasses the principles of levels-of-processing as well as additional principles necessary to the formation of a plausible theory of human memory.

The account of distinctiveness of encoding developed by Jacoby et al. (1979) is essentially that of specificity of encoding developed in the first section of the introduction. I have chosen to use the phrase "specificity of encoding" rather than "distinctiveness of encoding" as a matter of convenience, not to suggest any dissatisfaction with the concept of distinctiveness. Some of the manipulations employed in the experiments presented in the thesis may be explained directly in terms of specificity, making the relationship between the task and the encoding more natural and compelling. For the remainder of this section I will use the term distinctiveness, but arguments apply to specificity of encoding as well as to distinctiveness. Rather than repeat the distinctiveness argument, I will discuss the advantages that the distinctiveness hypothesis has over levels-of-processing.

Various investigators, in addition to Jacoby et al. (1979), have presented findings consistent with the idea that distinctiveness of encoding results in better memory performance (e.g., Auble, Franks, and Soraci, 1979; Begg, 1978; Bransford, Franks, Morris, and Stein, 1979; Craik and Tulving, 1975; Epstein, Phillips, and Johnson, 1975; Eysenck, 1979; Eysenck and Eysenck, 1980; Frase and Kamman, 1974; Gardiner, Craik, & Bleasdale, 1973; Goldman and Pellegrino, 1977; Johnson-Laird, Gibbs, and de Mowbray, 1978; Klein and Saltz, 1976; Lesgold and Goldman, 1973; Mathews, 1977; Moscovitch and Craik, 1976; Nelson et al., 1974; Norman and Bobrow, 1979; Packman and Battig, 1978; Smith, 1977; Stein, 1977, 1978; Stein, Morris, and Bransford, 1978; Winograd, 1981).

Distinctiveness of encoding, unlike levels-of-processing, can be specified independently of memory performance. The typical finding of superior memory for semantic as opposed to nonsemantic tasks is accounted for by the fact that meaning is usually specific to a relevant word in the task, while nonsemantic features are shared to a greater extent by stimuli in the task environment. The result is that nonsemantic features lack contrastive value in distinguishing stimuli. Hunt and Mitchell (1978) established that orthographically or phonemically distinct words were recalled more frequently than their more typical counterparts following encoding tasks of a semantic or nonsemantic nature. They suggest that distinctivess of encoding is a function of specific inherent features of the stimuli as well as an encoding task which encourages a distinctive encoding of the stimulus.

The principle of distinctiveness of encoding requires no modification to explain specificity or congruity. Specificity is an integral factor in determining the distinctiveness of an encoding by requiring that more specific and complete contrasts be made to distinguish the stimulus from alternatives. A congruous encoding is generally more distinctive than an incongruous encoding, because the congruous encoding task requires more specific contrasts to be made to distinguish the stimulus. An incongruous encoding context does not require any specific meaningful analysis. It seems that distinctiveness of encoding encompasses the important principles of levels-of-processing and provides a more adequate account of memory by more precisely specifying the nature of the encoding process and the resultant memory trace.

CHAPTER 2

EXPERIMENT 1

Some areas of cognitive research have been concerned with specificity of encoding, but have not examined the memory for the encoding. Studies of mental comparisons have established a phenomenon known as the "symbolic distance effect" (Moyer and Bayer, 1976). The basic nature of the phenomenon is that the more similar two words or symbols on the dimension of comparison (e.g., size of referent), the greater the time required to make a decision relevant to that dimension (g.v., Griggs and Shea, 1977; Holyoak, Dumais, and Moyer, 1979; Moyer and Dumais, 1978; Potts, 1975). It takes considerably longer to decide which is the larger of the pair snail-peanut than to decide the larger of the pair man-elephant. The differences in time required to make comparisons would seem to reflect directly the relative specificity of encoding within a single dimension.

The symbolic distance effect has been demonstrated for perceptual dimensions such as size (e.g., Moyer, 1973; Paivio, 1975) as well as for inferential dimensions such as pleasantness (Paivio, 1978; Paivio and Marschark, 1980), ferocity (Kerst and Howard, 1977), and intelligence (Banks and Flora, 1977). A similar effect has been demonstrated for semantic orderings such as number (Potts, 1972; Potts, Banks, Kosslyn, Moyer, Riley, and Smith, 1978), and time, quality, and temperature (Holyoak and Walker, 1976).

If the words in a pair are highly similar on the relevant dimension, deciding between them is more difficult and a more specific encoding is necessary than if the words are quite disparate on that dimension. In the latter case, it may suffice to tag one member of the pair as being large, or whatever, without fully contrasting it with the other member. The more similar the pair on the dimension of comparison, the more specific is the encoding required to distinguish between the members of the pair. Decision difficulty seems to provide an estimate of the specificity of encoding afforded a stimulus pair.

The generality of the symbolic distance effect makes it an ideal candidate for use in investigating the effects of specificity of encoding on human memory. Specificity of encoding as measured by decision difficulty is related directly to the similarity of the paired words on the dimension of comparison. It is possible to investigate a wide range of specificity by systematically varying the dimensional difference of word pairs. An independent measure of specificity avoids the problem of circularity, since specificity of encoding is established separately from level of recall.

Jacoby et al. (1979, Exp. 2) demonstrated the utility of the symbolic distance effect for investigating the effect of specificity of encoding on memory. Subjects rated the difference in referent size between pair members for a list of pairs that varied systematically as to difference in referent size. Probability of correct recall varied directly with decision difficulty.

The argument presented suggests that decision difficulty enhances memory performance because the specific encoding required to distinguish stimuli from potential alternatives results in a more distinctive and retrievable memory trace. An alternative explanation of the effects of difficulty, that of cognitive effort, does not regard specificity of encoding as a significant factor in memory performance.

The argument for cognitive effort is presented next along with evidence requiring modification of the effort account. The two accounts, specificity of encoding and cognitive effort, will be examined to see how they differ in emphasis.

Tyler, Hertel, McCallum, and Ellis (1979) define cognitive effort as "...the amount of the available processing capacity of the limited-capacity central processor utilized in performing an information-processing task" (p. 608). Thus, effort represents a quantitative account of the manner in which task difficulty affects memory performance. The greater the amount of processing capacity required by a cognitive task, the stronger is the memory for the items in the task. Two tasks that differ significantly in nature (e.g., semantic vs. phonemic) but which require the same amount of processing capacity should result in equivalent memory performance.

Tyler et al. (1979) used a divided-attention paradigm to measure the cognitive effort expended in carrying out a task. The paradigm consists of a primary task and a secondary reaction-time task. The greater the effort used in performing the primary task, the longer will be the reaction time for the secondary task.

It is stressed that amount of cognitive effort can vary independently of processing level. Tyler et al. (1979) used anagrams as a nonsemantic task and sentence completion as a semantic task, varying the effort required by the task at each level. The anagrams required little or extensive rearrangement; the semantic decisions were easy or difficult. The investigators found a significant effect of effort, but none for processing level, in free recall (Exp. 1). However, the effect of effort on the anagram task alone was not significant. In a similar experiment (Exp. 4) with the target absent during the decision period, both effort and processing level had significant effects on recall, with effort having an effect at each level of processing.

Tyler et al. propose two possible mechanisms for the memorial advantage bestowed by greater cognitive effort. Effort may work to integrate the context and target, facilitating retrieval. Alternatively, the item processed with more effort may be stored as a trace of greater strength and durability. If effort enhances memory by affecting the integration of the encoding context with the stimulus, the effect of effort may be regarded as being similar to that of specificity of encoding. However, the integration is seen as being caused by expended processing capacity rather than the nature of the task, so it remains possible to distinguish effort from specificity of encoding.

Krinsky and Nelson (1981) examined cognitive effort with different semantic and structural encoding tasks. They felt the anagram task used by Tyler et al. (1979) may have led to semantic processing, especially when the anagrams were difficult to solve. Krinsky and Nelson used pupillary dilation as an autonomic measure of task effort. The easy structural task consisted of determining whether a target word had more than one syllable. The difficult structural task consisted of deciding whether a target matched a presented pattern of consonants and vowels. The easy semantic task consisted of deciding whether the target represented something that could be held in the hand. The difficult semantic task consisted of determining whether a target represented something man-made, too large to be carried in a pocket, and that could be used as a weapon. Krinsky and Nelson found effects of difficulty in number of decision errors, decision latency, and pupillary dilation for both structural and semantic tasks. Difficulty also affected recall for both types of task. The easy semantic task did not produce a higher level of recall than did the difficult structural task.

There are potential problems with the Krinsky and Nelson work. Although errors are at a comparable level for difficult structural and semantic tasks, the decision latency and pupillary dilation measures for the difficult semantic task more closely resemble corresponding measures for the easy structural task; recall is much higher for the semantic ($p. = .41$) than for the structural ($p. = .17$) task in that case. The difficult structural task seems to require much more effort than the corresponding semantic task, without a corresponding superiority in recall ($p. = .30$). One could argue that subjects adopt a strategy of regarding the targets as meaningful units when the structural task is difficult. Such a strategy might make it simpler to maintain the target while comparing it with the standard; regarding the target as a string of consonants and vowels would render the task more difficult. As a result of these problems, Krinsky and Nelson's demonstration seems inconclusive as a defense of the effort hypothesis.

Auble and Franks (1978) suggest that effort toward comprehension facilitates recall, but only if comprehension results. They presented subjects with "incomprehensible" sentences, and cues. For example, the sentence "The girl spilled her popcorn because the lock broke" is considered incomprehensible without the cue "lion cage," which provides a meaningful context. Effort produced some effect in that sentences with the cues embedded were recalled less frequently than sentences with the cues presented separately. However, when the cues were presented prior to the sentences, recall was as high as or higher than when the cues were presented following the sentences. The former situation should require less effort to comprehend the sentences than does the latter. All conditions with cues resulted in higher recall than did conditions with no cues.

Although the Auble and Franks (1978) paper is presented in terms of cognitive effort, it fits readily into a specificity of encoding framework. The more difficult conditions are those in which the sentences are best recalled. Those sentences which remain incomprehensible for a time will be analyzed more specifically in an attempt to understand them. The no-cue conditions can be explained by either specificity of encoding or cognitive effort. If the task cannot be performed, so that the sentence remains anomalous, the trace of the event will remain general and nondistinctive, because the context provides no contrastive information. An effort explanation would thus have to suggest that cognitive effort must be of an effective or appropriate nature to make an event memorable.

A comparison of the specificity of encoding and cognitive effort hypotheses may serve to emphasize key aspects of the two theories. Specificity of encoding is a qualitative, relative view of memory. The memory for a stimulus is affected by the quality as well as the number of contrasts used to distinguish the stimulus in the encoding context. The value of such contrastive information is tied to the encoding context; the distinctiveness of a memory trace is affected by changes in the context from time of encoding to time of retrieval. By contrast, the effort hypothesis is a quantitative, absolute view of memory. The amount of attention or capacity utilized in a processing task determines the strength of the memory trace. The nature or quality of information processed is not a critical factor in retention; rather, the effort expended is the critical factor. Such differences in emphasis between specificity and effort accounts underlie specific aspects of comparison.

The specificity of encoding hypothesis stresses analysis conducted on stimuli in terms of the specificity required by the task. Specificity is a process account of memory, with the nature and specificity of a task influencing the distinctiveness of the resulting memory trace. A distinctive memory trace is highly discriminable and hence readily retrievable. A task may be difficult in ways that provide no distinctive information to the encoding, no distinctiveness to the memory trace; in such a case, the increased difficulty should result in no memorial advantage.

An effort account is not concerned with the concepts of specificity of encoding or quality of the memory trace, not in the sense of "quality" reflecting the nature of the trace. Instead, the quality of a memory trace is determined by the processing capacity utilized in carrying out the task. The effort hypothesis does not distinguish between effortful tasks that result in highly distinctive as opposed to nondistinctive information about the stimulus. The effort hypothesis might include trace distinctiveness as a byproduct of effort, but would not consider distinctiveness to be effective in enhancing memory.

The first experiment investigated the effect of specificity by varying the similarity of items being compared within a processing task, along with the level of processing required by a task. Referent size and word length were used as the dimensions of comparison. The two different encoding dimensions allow for a comparison of the effect of similarity in a task requiring referential or nominal processing of word pairs. Both tasks require judgments that have an objective, perceptual basis, but the referent size judgment requires that decisions be made from memory rather than from the physical appearance of the words.

It is expected that the referent size decisions would demonstrate an effect of specificity of encoding, a replication in essence of the Jacoby et al. (1979) findings. Pairs of referents similar in size require more specific encoding to make size comparisons than do pairs of referents dissimilar in size. Specific encoding should result in more distinctive traces for the encoding events, so that pairs of words similar in referent size should be remembered better than are pairs of words less similar in referent size.

The word length conditions provide a strong contrast. Decisions regarding word length discourage meaningful or distinctive encoding. Although decisions involving word length comparisons differ in the specificity required, the relatively specific encoding of a similar pair will center on physical features of the stimulus pair, adding little or no distinctiveness to the encoding. Hence, memory performance should be low and should reflect no benefit from greater specificity of encoding. The lack of a specificity effect would contradict the expectations of an effort account.

Three different memory tests were used with each stimulus list. The tests are as follows:

1) free recall,

2) combined cue recognition ~~and~~ associate recall, with semantic similarity distractors in the recognition phase, and

3) combined cue recognition and associate recaal, with size distractors in the recognition phase -- size list

or

with <u>physical similarity</u> distractors in the recognition phase - length list.

Two different types of distractors were chosen for the recognition targets from each stimulus list. For each of the targets from the size list, distractors were chosen to differ minimally from the target in size of referent (e.g., "thumb-cigar") or meaning (matched for specific category and other semantic features, e.g., "thumb-finger"). The two types of distractors may provide an indication of the specificity of encoding of pairs aside from their difference in size. Distractors that are similar in size to targets may be readily confusable with targets from decisions which require relatively little contrastive information. Distractors that also match targets on semantic features might cause more confusion than the size distractors for targets from initially difficult decisions, as the distinctive information in the memory trace may fit the distractor as well as the target.

For each of the recognition items from the length list, distractors were chosen to differ minimally from a target in terms of semantic (e.g., "leopard-panther") or physical (e.g., "leopard-leotard") features. The structural bias created by the processing task might cause the physically similar distractor to be more confusable with the target than the semantically similar distractor.

The effect of type of distractor on recognition for the targets from the length list may suggest the extent to which structural analysis dominates the word length conditions. The effect of type of distractor may indicate whether specificity of encoding in the length task affects the degree of semantic analysis carried out in the comparison task.

The different memory tests (free recall, and cue recognition plus associate recall with different types of distractors) provide information about the nature of the effect that specificity of encoding has on memory, whether all or only selective aspects of memory are affected by specificity of encoding. Free recall provides an estimate of the relative availability of individual words from the processing task. Free recall also provides a measure of the cohesion of the pair trace, from the frequency with which words from a pair are recalled together. The other retention tests can demonstrate the effect of specificity of encoding on recognition and on the retrieval of items paired during presentation.

Method

Subjects. Twenty subjects participated in each of six conditions, three with each of the two presentation lists. The 120 subjects were introductory psychology students who participated in the experiment as a course requirement. Subjects were tested in groups, with an average of ten subjects in each group.

Materials. Two presentation lists were constructed, with two different recognition test lists for each. (See Appendix 1 for a complete listing of stimulus pairs and recognition distractors used in the experiment.)

Presentation Lists. For the size list, words were chosen from Paivio's (1975) list of size ratings. Thirty critical pairs of items were constructed so that the size ratio of the larger to the smaller item was approximately 1.5, 2.0, 2.5, 3.0, or 3.5, with six pairs of items representing each of the five ratios (e.g., "chair-quail", "raccoon-ribbon", "lobster-dot", "dog-thimble", "goat-raisin"). That is, a pair of items with a 2.5 size ratio consists of two items, one of which, according to the size ratings, is 2.5 times as large as the other. The range of chosen ratios provides an extensive range of specificity of encoding. In addition to the critical pairs, ten buffer pairs were constructed by random selection and pairing of 20 other items from the ratings list. Items of approximately the same size were not paired.

Construction of the word length list was accomplished by random selection of 80 concrete nouns from Paivio, Yuille, and Madigan's (1969) ratings and pairing them so as to vary the difference in number of letters between members of a pair from one to five letters (e.g., "flood-cord", "chasm-fox", "daybreak-elbow", "skillet-nun", "furniture-hoof"). Eight pairs represented each of the five differences in word length. Thirty of these pairs, six from each of the five differences in word length, were used as critical pairs for the presentation list, while the remaining ten were used as buffer pairs.

For each presentation list, the 40 pairs were recorded on videotape so that a single word pair appeared on the screen at a time. A five-second presentation rate was used. Five buffer pairs preceded the 30 critical pairs and five buffer pairs followed. The presentation order within the structure of critical and buffer pairs was predetermined in a random fashion.

<u>Test Lists</u>. Two recognition test lists were constructed for use with each of the presentation lists. Each test list consisted of a target and a distractor to represent each of the critical pairs in the

presentation list. A distractor was chosen individually for each target item, in a manner to be discussed later. Personal judgment determined the appropriateness of a distractor. The targets and distractors were presented singly on the test sheet in random order. The nature and selection of the distractors differ for the two presentation lists and will be explained separately.

The two recognition test lists for the size conditions were constructed so as to offer distractors which were highly similar to targets either in terms of size or meaning. A size distractor for each target was chosen by finding an item in the Paivio (1975) size ratings that varied minimally in rated size from the critical item. Matching a target and distractor for the semantic similarity test list was a less precise procedure. The semantic distractor for a target was a member of the same specific category (e.g., insect, mammal), and was as similar as possible in terms of sharing salient semantic features (e.g., "bee-wasp," "mule-donkey"). These selection criteria resulted in distractors being similar to targets in terms of size as well as specific aspects of meaning. Half of the targets were the larger member of a presented pair and half the smaller.

The two recognition test lists for the length conditions were constructed so that distractors varied minimally from their targets either in terms of physical appearance or meaning. As much as possible, distractors were of similar word length to the corresponding target items. Distractors for the physical similarity list were as similar as possible to their targets in terms of visual appearance (e.g., "friend-fiend"). Distractors for the semantic similarity list were as similar as possible to the corresponding targets in terms of meaning, being synonyms or members of the same category sharing some attributes (e.g., "friend-ally"). Half of the target items were the longer item in a presented pair and half the shorter.

Procedure.

Decision and Rating. Subjects in the size conditions were told they were participating in an experiment examining decision difficulty. They would be shown 40 pairs of words; when a pair appeared on the videoscreen they were to indicate which referent was larger by indicating left or right, corresponding to its position on the screen. It was stressed that the decision was to be based on the size of the objects represented by the words rather than the words themselves. After making the decision, subjects were to indicate the difficulty of the decision by rating it on a 7-point scale, with 1 representing extremely difficult decisions. Subjects were informed of the 5-second presentation rate, as the task required paying close attention to the screen.

Subjects in the length conditions were given instructions identical to those for the size conditions, with the exception of the basis of their comparisons. They were told to choose the member of a pair which had the greater number of letters.

Memory Tests. After subjects in the size conditions completed the decision and rating task for all pairs, they were given one of three retention tests. Subjects in the free recall condition were asked to write down as many words as they could recall from the presentation list. Subjects in the size or semantic similarity recognition plus recall conditions were given the corresponding recognition list and asked to circle the 30 words on the sheet which had appeared in the presentation list, and write in the word which had appeared with each on the list. Subjects in each test condition were given five minutes to complete the task.

Subjects in the length conditions were also given one of three retention tests. Instructions for the free recall condition and for the physical similarity and semantic similarity recognition plus recall conditions were identical to those used for the tests in the size conditions.

Results and Discussion

Results will be reported separately for size and length conditions. Specific comparisons between the two decision dimensions will be made following the individual discussions.

The .05 $\underline{\alpha}$ level was used for all tests of significance.

Size

Rating. Table 1 presents mean rating of decision difficulty as a function of size ratio, with the corresponding standard deviations. The rating data are collapsed across conditions to give an overall difficulty rating for each ratio. As subjects in each condition were treated identically throughout the rating phase, there is no problem with collapsing the data across conditions. Ratings were comparable in all conditions.

An increase in rating indicates that the relevant comparisons are considered easier than others with lower difficulty ratings. There is a slight increase in rating as the size ratio increases. A 1x5 repeated measures analysis of variance was performed on the rating data. The results of the ANOVA indicate that size ratio had a strong effect on

Table 1. Mean rating of decision difficulty (on a seven-point scale with 1 representing an extremely difficult decision) as a function of size ratio or difference in word length for conditions in Experiment 1. Standard deviations are in parentheses.

	SIZE RATIO				
	1.5	2.0	2.5	3.0	3.5
MEAN RATING OF DECISION DIFFICULTY	5.93 (.79)	6.57 (.53)	6.65 (.41)	6.78 (.40)	6.82 (.37)

	DIFFERENCE IN WORD LENGTH (NUMBER OF LETTERS)				
	1	2	3	4	5
MEAN RATING OF DECISION DIFFICULTY	5.22 (1.19)	5.99 (.91)	6.53 (.75)	6.77 (.48)	6.87 (.25)

difficulty rating [$\underline{F}(4,236)=10.4$, MSe=7.71], despite the fact that the differences in ratings are not large in absolute terms. A $\underline{t}$-test for comparison of the means was performed and indicates several significant differences among the size ratio ratings [1.5 ($\overline{X}$ = 5.93) < all others; 2.0 ($\overline{X}$ = 6.57) < 3.0 ($\overline{X}$ = 6.78); 2.5 ($\overline{X}$ = 6.65) < 3.5 ($\overline{X}$ = 6.82)]. The rating data serve to confirm that the size ratio manipulation has an effect on decision difficulty, although, as mentioned above, the differences were not large in absolute terms.

Recognition. Table 2 presents the mean number of words correctly recognized as a function of size ratio, along with the corresponding standard deviations, for the size and semantic similarity conditions. Probability of correct recognition is also provided for each ratio and condition. The downward trend in correct recognition as a function of an increase in size ratio is apparent for each of the two conditions. A 2x5 mixed design analysis of variance shows the effect of type of distractor (size, semantic similarity) to be nonsignificant ($\underline{F}<1$) and the effects of ratio [$\underline{F}(4,152)=37.3$] and the interaction of ratio and type of distractor [$\underline{F}(4,152)=2.86$, MSe=0.90] to be significant. A $\underline{t}$-test comparison of the means reveals that recognition performance is significantly lower at the 3.0 and 3.5 size ratios than at the 1.5 ratio. False alarm rates are approximately 5% for each condition.

One may conclude from the lack of significance of the distractor variable that type of distractor has no systematic effect on recognition in this task. The significance of the interaction suggests that type of distractor affects recognition in some manner, the exact nature of which is not determinable from the data. Size distractors seem to have a more

Table 2. Mean rating of words correctly recognized (of a possible 6) as a function of size ratio or difference in word length for each recognition condition in Experiment 1. Standard deviations are in parentheses. Probability of correct recognition for each value is indicated by an asterisk.

	SIZE RATIO				
Condition	1.5	2.0	2.5	3.0	3.5
Size	4.95 (1.19) *.82	4.65 (1.27) *.78	4.80 (0.95) *.80	3.40 (1.19) *.57	3.50 (1.73) *.58
Semantic Similarity	5.15 (0.81) *.86	4.40 (1.14) *.73	4.30 (1.26) *.72	4.30 (1.49) *.72	3.95 (1.05) *.66
Average Across Conditions	5.05 *.84	4.53 *.76	4.55 *.76	3.85 *.65	3.73 *.62

	DIFFERENCE IN WORD LENGTH (NUMBER OF LETTERS)				
Condition	1	2	3	4	5
Physical Similarity	2.70 (1.30) *.45	2.40 (1.47) *.40	2.15 (1.14) *.36	2.15 (1.09) *.36	2.83 (1.41) *.47
Semantic Similarity	2.60 (1.05) *.43	2.85 (1.42) *.48	2.20 (1.47) *.37	2.85 (1.27) *.48	2.35 (1.79) *.39
Average Across Conditions	2.65 *.44	2.63 *.44	2.18 *.37	2.50 *.42	2.59 *.43

detrimental effect on recognition of items from large size ratios than from smaller ratios, relative to the semantic similarity distractors. Such a result is consistent with the nature of the decision task, which requires attending specifically to the referent size of presented items. When the comparison being made is an easy one, size of the referent may be the sole distinctive information considered. In such a case, size information is quite sufficient to distinguish the pair members and thus there is no need to engage in more detailed meaningful processing of the items. Detailed size information is not sufficient for recognizing presented items on a retention test which includes distractors matched with targets on referent size.

From the results it is concluded that the memory traces formed during the decision task include a size as well as a meaning component. The inclusion of a size component is consistent with the notion that the nature of the encoding required to perform a task affects the memory trace of the event. The specificity of the encoding and of the resulting memory trace must be considerable, as many of the semantic similarity distractors are not greatly discrepant in size from the corresponding targets, making the distractors similar in both size and meaning to the targets.

The overall decrease in recognition as a function of increasing difference in size between pair members is as expected. The discrepant pairs are not memorable, because they were easily distinguished in the original comparisons, requiring no specific contrasts to be drawn to distinguish between them. The ease of comparison allowed analysis to be performed at a general level without loss of accuracy. The size-similar distractors have little effect on recognition for small-ratio items, for which relatively specific encoding of the pair was required to make a correct decision. However, the size-similar distractors appear to depress recognition in the larger ratios. It may be that for the large ratio items, little more than general meaning and comparative size information are encoded during the comparison task, making the traces relatively indistinct.

Recall. Table 3 presents the mean number of words correctly recalled from each size ratio for each of the three recall conditions. In the recall portion of the cue recognition plus associate recall tests, size and semantic similarity, a word is scored as correctly recalled only if it is paired with the cue originally presented with it. It is important to note that the items used for analysis of correct recall in the free recall condition are the to-be-recalled items from the cued recall conditions, so that all measures have the same maximum recall value.

Inspection of results for the two cued recall conditions reveals that a decrease in recall accompanies an increase in size ratio, suggesting an effect of specificity. The effect appears to persist when recall is conditionalized on recognition. Table 4 presents values for correct recall conditional on recognition for each of the cued recall conditions. Values were calculated for each subject and then averaged across subjects. The free recall condition shows recall levels to be equivalent for all size ratios, showing no effect of similarity on the availability of individual items.

A 3x5 mixed design analysis of variance confirms the significant effects of recall condition [$F(2,57)=22.4$, $MSe=1.82$], ratio [$F(4,228)=4.91$], and the interaction of the two variables [$F(8,228)=$

Table 3. Mean number of words correctly recalled (of a possible 6) as a function of size ratio or difference in word length for each of the test conditions in Experiment 1. Standard deviations are in parentheses. Probability of correct reccall for each value is indicated by an asterisk.

	SIZE RATIO				
Test Condition	1.5	2.0	2.5	3.0	3.5
Size	1.15 (.99) *.19	1.45 (1.1) *.24	0.55 (.60) *.09	0.65 (.99) *.11	0.75 (1.39) *.12
Semantic Similarity	1.20 (1.15) *.20	0.75 (.85) *.12	0.35 (.59) *.06	0.35 (.59) *.06	0.30 (.47) *.05
Free Recall	1.85 (1.14) *.31	1.75 (.85) *.29	1.90 (.85) *.32	1.75 (.91) *.29	1.85 (1.23) *.28
Degree of Association in Free Recall (Probability of recall of one member of pair given recall of other)	.33	.49	.09	.16	.12

	DIFFERENCE IN WORD LENGTH (NUMBER OF LETTERS)				
Test Condition	1	2	3	4	5
Physical Similarity	0.10 (.31) *.02	0.15 (.37) *.025	0.10 (.31) *.02	0.15 (.37) *.025	0.20 (.41) *.03
Semantic Similarity	0.15 (.37) *.025	0.05 (.22) *.01	0.10 (.31) *.02	0.10 (.31) *.02	0.00 0 *0
Free Recall	0.45 (.69) *.075	0.50 (.51) *.08	0.50 (.69) *.08	0.75 (.79) *.125	0.70 (.66) *.12

Table 4. Conditional probability of correct recall given recognition as a function of size ratio for the associate recall conditions in Experiment 1.

	SIZE RATIO				
Condition	1.5	2.0	2.5	3.0	3.5
Size	.23	.32	.11	.17	.16
Semantic Similarity	.24	.17	.09	.07	.09

2.04, MSe=0.66]. A t-test comparison of the means shows each condition to be distinct from each other, with free recall being highest, size intermediate, and semantic similarity lowest in overall recall performance. Recall for 1.5 ratio items is marginally greater than for 3.0 ratio items.

Ratio appears to have no effect in the free recall condition. The semantic similarity condition shows a clear effect of ratio, with recall being higher for items from low ratios than for those from high ratios. The size condition shows basically the same effect, with some perturbations. For cued recall, specificity of encoding affects the correct recall of an item with its original pair member as the cue.

The superiority of recall in the size condition over that in the semantic similarity condition is explained easily. The semantic distractors appear to affect recall performance at lower ratios than do size distractors. As mentioned previously, the semantic distractors are not greatly discrepant in size from the pair members with which they are matched. Therefore, incidentally, semantic distractors are similar to presented items in size and meaning, thereby increasing their confusability with presented items. The systematic effect of semantic distractors on recognition is mirrored in cued recall. The semantic distractors may serve to reduce the distinctiveness of the pair trace. The contrasts used in the initial processing of the pair may be applicable to a pairing of the distractor with the response as well as to the target with its response. Such overlap of contrasts would diminish the distinctiveness of the original pair trace more than would be the case when only size contrasts are shared between the target and its corresponding distractor.

The equivalence of free recall performance for different ratios suggests that individual items from all size ratios are equally available in memory. A task that does not require reinstatement of the original pair structure is performed quite adequately. The equivalence of recall for all ratios argues against incomplete encoding of greatly discrepant pairs, although the members of the pair need not be contrasted precisely. Table 3 also presents the degree of association between pair members in free recall, presenting probability of recall of the second member of a pair given recall of the first, for each ratio. Degree of association values were calculated for each subject, then averaged across subjects. Despite perturbations, it is clear that, as was the case in cued recall, pairs are more frequently recalled from small than from large ratios, even though individual words from all ratios are recalled equally frequently. This finding indicates that the pair is not being contacted as a unit for dissimilar pairs, for which specific encoding was not required for successful performance in the comparison task. The pattern for degree of association in free recall mirrors the data presented in the cued recall conditions.

There is a ready explanation for the superiority of free recall to cued recall in the size conditions. The structure imposed by cued recall makes the task more difficult than free recall. Subjects must recognize the target and reinstate its pair membership by providing the appropriate response. Cue recognition was affected by specificity of encoding, resulting in fewer opportunities for recall of items from large ratios. There is no outlet for subjects to indicate items recalled without recognition of the appropriate targets from the presentation list. Cued recall in the present context is necessarily a matter of recall of the pair rather than of individual words.

The cued recall results as discussed depend on a stage model of cued recall. The cue must be encoded, and recognized as having occurred previously. Cue recognition entails making contact with the original trace of the pair. Contact with the original trace allows response production to occur if the trace contains adequate response information (Begg, 1978). Both stages of cued recall were affected by specificity of encoding in the size comparisons. Cue recognition and the subsequent response production both occurred with increasing frequency as the specificity required by the task increased.

The data as a whole suggest that two aspects of memory are affected by specificity of encoding in the size comparison task. Recognition of an item as having been presented is affected, with recognition decreasing as the similarity of the pair members decreases. Also, the effectiveness of a cue for recall, when the cue is a member of the presented pair, is reduced as the initial decision requires less specificity of encoding. The subject cannot use the cue effectively to retrieve the pair member required. The failure may be caused by a difference in the specificity of the pair trace for discrepant pairs. Response availability per se is not affected by specificity of encoding in the size comparison task. The free recall results demonstrate that individual items, but not pairs, are equally available in memory for all levels of specificity of encoding tested. Differences only appear if the retention test requires discrimination finer than that initially performed.

Length

Rating. Table 1 presents the mean rating of decision difficulty as a function of the difference in word length, with corresponding standard deviations. The rating data are collapsed across conditions to give an overall difficulty rating for each difference in word length. Ratings for each condition are comparable. The upward trend is apparent, although small in absolute terms; as the difference in word length increases, so does the ease of making the length comparison. A 1x5 repeated measures analysis of variance substantiates the effect of difference in word length on difficulty rating [$F(4,236)=89.8$, $MSe=11.1$]. A t-test comparison of the means indicates that each difference in word length is significantly different from any other, except for 4 and 5. The length manipulation thus had an effect on perceived decision difficulty.

Recognition. Table 2 presents the mean number of words correctly recognized for each difference in word length between pair members, along with the standard deviations, for the physical similarity and semantic similarity conditions. The probability of correct recognition is also provided for each difference in word length and condition.

Recognition is considerably lower in the length conditions than in the size conditions, for which the probability of correct recognition ranged from .58 to .86. For length conditions, the probability of correct recognition ranges from .36 to .48. False recognition is high, approximately 15% for the length conditions.

There appears to be no systematic effect of difference in word length on correct recognition. A 2x5 mixed design analysis of variance revealed no significant effects [type of distractor, $F<1$; difference in word length, $F(4,152)=1.13$; the interaction of difference and type of distractor, $F(4,152)=2.07$, MSe=1.4].

The lack of effect of specificity of encoding on recognition is as predicted. Comparisons of word pairs on the basis of word length do not require that members of a pair be meaningfully contrasted to determine which word is longer. Attention to the structural elements of the word pair does not provide distinctive information, except incidentally, useful in a standard recognition task. The pattern of results for recognition following the word length comparisons indicates that it is not merely cognitive effort, or difficulty *per se*, that results in higher retention. Rather, the crucial factor is the resulting distinctiveness of the memory trace relative to alternatives considered during the recognition test.

Recall. Table 3 presents the mean number of words correctly recalled, as well as the probability of correct recall, as a function of difference in word length for each of the three recall conditions. Correct recall is defined as for the size conditions. In the cued recall conditions, recall of a word is scored as correct only if it is paired with the word originally presented with it. In free recall, the to-be-recalled items from cued recall are the items used for analysis. The probability of recall is extremely low, less than .10 overall for any condition.

A 3x5 mixed design analysis of variance reveals the significant factors to be those of recall condition [$F(2,57)=23.3$, MSe=0.32] and the interaction between condition and difference in word length [$F(8,228)=3.0$, MSe=0.18]. Difference in word length itself is not a significant factor ($F<1$). The effect of condition is attributable to the superiority of free recall to the cued recall conditions. The less rigid structure of free recall has been discussed in the recall results for size ratio; the argument holds as well for the recall results for difference in word length. The level of recall in the cued recall conditions is so low as to call into question any explanation of the cause for the interaction effect. It is possible that the cued recall results represent a floor effect.

The lack of effect on recall of difference in word length underlines the fact that the specificity of encoding argument presented relies on the distinctiveness of the memory trace causing the effect on memory. In a structural or nonsemantic task such as the length task, no meaningful analysis is necessary. The context at recall is greatly different from the context at encoding and requires use of trace features which may not exist following the structural task. The incongruity of context between encoding and testing serves to hinder recall in a standard recall task.

Comparison of the results for the size and length conditions provides a means of distinguishing two explanations of the memory effects attributable to decision difficulty, namely specificity of encoding and cognitive effort. I shall argue that the results of this experiment are consistent with the tenets of a specificity of encoding hypothesis while refuting a standard cognitive effort hypothesis.

Recognition and cued recall data for the conditions making referent size comparisons indicate that specificity of encoding, as measured by decision difficulty, has an effect on memory performance in standard tasks. Both recognition and cued recall levels are higher for pairs requiring an initially difficult comparison than for pairs requiring a relatively easy one. No such effect of specificity is found in the corresponding data for the conditions making comparisons of word length. Specific analysis relevant to the comparison task is beneficial to memory of the stimulus only if the resulting contrastive information is of a nature to be reinstated at the time of retrieval. The distinction between size and length comparisons in terms of demonstrating an effect of specificity is expected from the specificity of encoding argument. Some meaningful processing of length pairs may occur; the typical purpose of reading words is to process them meaningfully. However, a very general encoding for meaning is more than sufficient for the length comparison task; consequently, only a minimal contrastive analysis may be performed on the pairs in such a task. Such a minimal encoding for meaning is insufficient to facilitate recognition or recall during a standard memory test.

The mean rating for overall decision difficulty is 6.55 for the size comparisons and 6.28 for the length comparisons; the length task is perceived as being more difficult than the size task. The retention effects are opposite to the between-task difference in difficulty, contrary to the thesis of a simple effort hypothesis.

A 2x5 mixed design analysis of variance indicates that dimension of comparison [$\underline{F}(1,118)=8.05$, MSe=49.9] has a significant effect on difficulty rating. Decision difficulty for high to intermediate degrees of pair similarity is perceived as being greater for length than for size comparisons.

Rated difficulty may be regarded as an indicator of the cognitive effort expended during a task. The significantly more effortful task of length comparisons does not result in higher, but rather lower, recall. Free recall and cued recall probabilities for length versus size conditions respectively are as follows: p.=.1 and .02 for length, p.=.3 and .13 for size. The results are contrary to the predictions generated from a cognitive effort hypothesis.

One could argue that rated difficulty is not an accurate indicator of the actual cognitive effort involved in a task. Such an argument must provide a strong rationale for disregarding the ratings. Otherwise the argument remains circular; if more effort is involved, higher recall results. Using an indicator distinct from the recall data is preferable to a post hoc ordering of tasks by recall level and declaring the ordering consistent with effort, unless the indicator is necessarily invalid. The validity of rated difficulty as an indicator of effort has not been called into question. In addition, rated difficulty is internally consistent, reflecting the similarity structure of the stimulus lists.

The size and length data provide solid support for a specificity of encoding hypothesis, with its emphasis on contrastive analysis and the relative distinctiveness of resulting memory traces. Size comparisons result in meaningful, distinctive traces relative to length comparisons; recognition and recall levels are higher for size than for length conditions. Difficult size comparisons require more specific meaningful encoding of words relative to easier size comparisons; recognition and cued recall levels reflect the difference in specificity. Difficult length comparisons require a more detailed structural encoding of words than do easier length comparisons, but the comparisons do not provide contrasts that make the memory traces distinctive during the retrieval task. Recognition and recall levels are no higher for items from difficult than from easy length comparisons.

The first experiment has served two purposes. The more important of these was the demonstration that specificity of encoding does a better job of predicting and explaining levels and patterns of retrieval obtained after comparisons that differ in difficulty than does the hypothesis of cognitive effort. The second purpose of the experiment was to establish a picture of the effects that specificity of encoding has on memory performance. When the encoding task emphasizes meaningful analysis of stimuli, greater specificity of encoding results in a memory trace that is readily retrievable, with the contents of the trace being recognized and recalled with greater frequency. Specificity of encoding also results in a memory trace that is more cohesive, with the contents being generated as a unit more frequently.

The second experiment explores further the nature of the effect that specificity of encoding has on the nature and retrievability of the memory trace. Critical aspects of the dimension of referent size, such as objectivity and its perceptual nature, are considered by examining a dimension that differs with respect to those aspects, the dimension of synonymy.

CHAPTER 3

EXPERIMENT 2

The first experiment demonstrated effects of specificity of encoding on associative recall and recognition after a task requiring comparisons of referent size. The size comparisons entail judgments on a semantic dimension that has a specific objective basis. The second experiment examined the extent to which specificity of encoding affects memory following a task requiring judgments of a more inferential nature. Synonymy was chosen as an appropriate dimension for investigation. As in the first experiment, specificity of encoding was manipulated by varying the similarity of meaning within comparisons.

Synonymy represents a special aspect of semantic analysis. In most cases of meaningful comparison, a fairly general sense of a word is sufficient to distinguish it from alternatives. A comparison of functional synonyms requires that analysis proceed from the general meaning of a word to a specific meaning of the word, "stemming from restrictions on the use of the word for precise communication" (Herrmann, 1978, p. 473). Connotative as well as denotative aspects of meaning restrict the relative applicability of synonyms in a particular context and serve to distinguish synonyms from each other.

The use of synonymy as a dimension of comparison represents a case in which the pre-experimental association between pair members is quite high. Is it possible to give subjects a comparison task which will affect this strong association? Begg (1978) demonstrated that examining related items for differences in meaning results in higher recall than does examining them for similarity in meaning. (See also

Hunt and Einstein, 1981). The comparison task in this experiment uses a similar comparison. The task draws attention to differences between members of a pair by asking subjects to decide which member has the more specific reference. The task should highlight differences in the range of applicability of pair members. When the decision as to specificity of reference is relatively simple, the contrasts made during the comparison should be relatively general and imprecise; serving to trivialize the encoding relative to comparisons in which the distinction is more difficult.

The encoding task requires that comparisons of a highly specific nature be drawn. Retention levels should be high for all levels of synonymy tested. The relationship between level of synonymy and decision difficulty in the comparison task is not apparent. Synonymy and specificity of reference are not identical. However, the greater distinctiveness resulting from the more difficult of the synonymy comparisons should yield a higher level of recall. If pairs from one level of synonymy are perceived as being more difficult to distinguish as to specificity of reference, then items from those pairs should be recognized and recalled more frequently, relative to items from a level of synonymy that requires less difficult distinctions. Internal consistency must be considered the primary test of whether distinctiveness in the synonymy comparisons results in more retrievable memory traces.

Method

Subjects. Twenty subjects participated in each of three conditions. The 60 subjects were introductory psychology students who participated in the experiment as a course requirement. Subjects were tested in group sessions, with each group randomly assigned to a condition.

Materials. (See Appendix 2 for a complete listing of stimulus pairs and distractors used in the experiment.)

Presentation List. Forty pairs of words were chosen from Whitten, Suter, and Franks' (1979) synonymy ratings for noun pairs. Pairs were chosen so that no pair was synonymous with any other used. All of the pairs in the Whitten et al. (1979) ratings are synonyms, rated on a seven-point scale with 7 representing virtually perfect synonymy. The 40 pairs selected represent mean synonymy ratings of 6.5, 5.5, 4.5, 3.5, and 2.5 (e.g., VICTOR-WINNER, POSTURE-STANCE, WINDOW-PORTHOLE, SOLID-MASS, SPIKE-NEEDLE), with equal representation from each level of synonymy. Thirty of the chosen pairs were used as critical pairs, with six pairs representing each level of synonymy. The remaining pairs served as buffers. The synonymy task differs from others in the present investigation in that the pairs used are more abstract than are the stimulus pairs constructed for the other dimensions of comparison.

The stimulus pairs were recorded on videotape so that a single pair appeared on the screen at a time. A five-second presentation rate was used. Five buffer pairs preceded the critical pairs and five buffer pairs followed. A random order of presentation within the critical and buffer pairs was predetermined.

Recognition Test Lists. Two recognition test lists were constructed. Each list consisted of an equal number of targets and distractors, representing each critical pair in the presentation list. A distractor was chosen individually for each target, with personal judgment determining the appropriateness of a distractor. The targets and distractors were positioned on the test sheet in a random order.

The recognition lists were constructed so that distractors varied minimally from their targets in terms of physical similarity (e.g., harm, harem) or semantic similarity (e.g., harm, injury) of the words. Particular care was taken to ensure that a semantic distractor was synonymous with both members of the appropriate stimulus pair, so that no alternative meaning for a word was stressed in the recognition task. Half of the targets were the first member of a pair as presented, and half were the second.

Procedure. Subjects were told they were participating in an experiment examining the difficulty of decisions. They would be shown 40 pairs of words; when a pair appeared on the screen they were to indicate the word with the more specific meaning by choosing left or right, corresponding to its position on the screen. After making a decision, subjects were to indicate the difficulty of that decision by rating it on a seven-point scale, with 1 indicating an extremely difficult decision.

After subjects completed the encoding task, they were given one of three retention tests. Subjects in the free recall condition were asked to write down as many words as possible from the presentation list. Subjects in the semantic similarity or physical similarity conditions were given the appropriate recognition test list and asked to circle the 30 words which had appeared on the presentation list, and to write in the word which had appeared with each in the list. Subjects were given five minutes to complete each memory test.

Results and Discussion.

The .05 α level was used for all tests of significance.

Rating. Table 5 presents mean rating of decision difficulty, and the corresponding standard deviation, as a function of level of synonymy. As test conditions were treated identically during the rating task, the rating data are collapsed across test condition to provide an overall rating of difficulty. There were no explicit expectations as to the differences in specificity required by varying levels of synonymy. It appears that the extreme levels of synonymy tested require the most difficult decisions on the basis of specificity of reference.

A 1x5 repeated measures analysis of variance performed on the data reveals that level of synonymy had a significant effect on rated decision difficulty [$F(4,216)=2.72$, $MSe=15.4$]; the differences in rating are not large in absolute terms. A t-test comparison of the means indicates that pairs from the 5.5 and 4.5 levels of synonymy are rated as requiring easier decisions than are pairs from the 6.5, 3.5, and 2.5 levels.

A simple prediction concerning specificity might be that difficulty ratings would decrease with a decrease in level of synonymy. Pairs of nearly-perfect synonyms must be extremely close in meaning, suggesting a high degree of similarity in scope of reference. Looser synonyms have less overlap in terms of shared features and more contrastive features, making them easier to distinguish. Such a straightforward pattern of difficulty did not occur; synonymy and specificity of reference are not identical. Nevertheless, if distinctiveness resulting from the more specific encoding given difficult pairs is effective in producing more retrievable traces, the

Table 5. Mean rating of decision difficulty as a function of level of synonymy in Experiment 2. Standard deviations are in parentheses.

	LEVEL OF SYNONYMY				
<u>Test Condition</u>	6.5	5.5	4.5	3.5	2.5
MEAN RATING OF DECISION DIFFICULT	5.25 (1.25)	5.55 (1.21)	5.50 (1.14)	5.24 (1.11)	5.38 (1.18)

patterns found in the rating data should be paralleled in the memory data. That is, pairs from levels of synonymy perceived as requiring more difficult decisions in the encoding task should be recognized and recalled with greater frequency than are pairs from less difficult decisions, requiring less specificity of encoding.

Recognition. Table 6 presents mean number of words correctly recognized as a function of level of synonymy, along with standard deviation and probability of correct recognition, for each recognition condition. Words from the 5.5 and 4.5 levels of synonymy, those rated as requiring less difficulty for the initial comparisons, are less frequently recognized than are words from the other levels of synonymy.

A 2x5 mixed design analysis of variance reveals that while type of distractor was not a significant variable ($F<1$), level of synonymy was [$F(4,162)=.60$, $MSe=0.94$]. A t-test comparison of the means indicates that words from the 6.5, 3.5, and 2.5 levels of synonymy were recognized more frequently than were targets from the 5.5 level of synonymy. The 4.5 level of synonymy, while not significantly so, did result in lower recognition than did those levels of synonymy rated as requiring more difficult comparisons in the encoding task.

The pattern of results found in the recognition data parallels the pattern of the difficulty ratings. Targets from levels of synonymy that required greater specificity of encoding are recognized more frequently than are targets from levels of synonymy requiring less specificity.

The degree of specificity required by the initial comparison task results in highly specific memory traces. The specific contrasts required by the comparison task are so precise that they virtually

Table 6. Mean number of words correctly recognized (of a possible 6) as a function of level of synonymy for each of the recognition conditions in Experiment 2. Standard deviations are in parentheses. Probability of correct recognition is indicated by an asterisk.

	LEVEL OF SYNONYMY				
Test Condition	6.5	5.5	4.5	3.5	2.5
Physical Similarity	4.45 (1.00) *.74	4.20 (1.10) *.70	4.45 (0.89) *.74	4.50 (1.05) *.75	4.85 (1.39) *.81
Semantic Similarity	5.05 (1.00) *.84	4.25 (1.02) *.71	4.35 (1.35) *.73	4.85 (0.99) *.81	4.75 (0.97) *.79
Averaged across Recognition Conditions	4.75 *.79	4.23 *.70	4.40 *.73	4.68 *.78	4.80 *.80

preclude confusion of targets and distractors. The semantic distractors, while synonymous with the corresponding targets, have not been matched with presented items for the level of synonymy represented in the original pairs. The pair traces contain specific information that render distractors ineffective.

Recall. Table 7 presents mean number of words correctly recalled as a function of level of synonymy for each of the three recall conditions, with the corresponding standard deviation and probability of correct recall values. Correct recall was scored as described in Experiment 1.

The cued recall conditions resulted in recall levels superior to those exhibited by the free recall condition. The superiority of cued recall to free recall is a reversal of the trends found for recall conditions involving size or length comparisons.

There is a high degree of similarity among recall patterns in the various test conditions. Such strong similarity of recall patterns is unique to the synonymy results. The basic pattern established by the rating data for synonymy is repeated in the recall data; recall levels are higher for those levels of synonymy that required more difficult decisions during the comparison task.

A 3x5 mixed design analysis of variance performed on the recall data indicates that recall condition [$F(2,57)=56.3$, MSe=3.06] and level of synonymy [$F(4,228)=11.9$, MSe=0.91] had significant effects on recall performance. Recall in the semantic similarity condition is higher than in the physical similarity condition. Both cued recall conditions exhibited higher levels of recall than did the free recall condition. The specific differences between conditions and levels of synonymy were

Table 7. Mean number of words correctly recalled (of a possible 6) as a function of level of synonymy for each test condition in Experiment 2. Standard deviations are in parentheses. Probability of correct recall is indicated by an asterisk.

	LEVEL OF SYNONYMY				
Test Condition	6.5	5.5	4.5	3.5	2.5
Physical Similarity	3.60 (1.31) *.60	2.95 (0.83) *.49	2.35 (1.27) *.39	3.55 (1.36) *.59	3.40 (1.60) *.57
Semantic Similarity	3.85 (1.39) *.64	3.20 (1.06) *.53	2.70 (1.45) *.45	3.90 (0.79) *.65	3.60 (1.19) *.60
Free Recall	1.00 (1.03) *.17	0.75 (0.72) *.12	0.80 (0.77) *.13	1.55 (1.10) *.26	1.15 (1.04) *.19

confirmed by t-test comparisons of the means. The 5.5 and 4.5 levels of synonymy resulted in lower recall levels than did the other levels of synonymy. The effect of level of synonymy in cued recall persists when conditionalized on recognition.

The cues for recall from the synonymy list provide strong clues to the items being sought. The semantic distractors may act as additional cues, or they may serve to eliminate possible alternative response candidates. The structure of the presentation list is so obvious that a memory search must be restricted to candidates with high similarity in meaning. The structure provides an advantage to either cued recall condition. Subjects in the free recall condition must carry out a memory search without clues to the targets.

All three recall conditions demonstrate the same effect of level of synonymy. Recall in the free recall condition as well as in the cued recall conditions was affected by specificity of encoding. The pattern of results suggests that items from various levels of synonymy are not equally available, contrary to the findings for the size free recall condition in Experiment 1. The availability of individual items in free recall was affected by specificity of encoding.

The results for the synonymy conditions are internally consistent. For both recognition and recall, performance is higher for pairs involved in difficult initial comparisons. Within the constraints of such a restricted range, an increase in the specificity required to complete the task successfully results in higher memory for the items involved in the task. Specificity of encoding affected the distinctivess of the memory traces for pairs involved in synonymy comparisons, even if specificity was not predictable on a priori grounds.

The results indicate that several aspects of memory are affected by specificity of encoding in the synonymy task. Recognition of an item as having been presented, the effectiveness of a cue for recall, and response availability are all affected by specificity of encoding, with each measure demonstrating a decrease as specificity of encoding decreases.

There is evidence to suggest that the effect in recall attributable to differences in specificity of encoding is primarily an effect of less contrastive encoding being detrimental to a standard encoding for meaning. The probability of correct recall in free recall is relatively low (p.=.12 to .26) given the highly salient relationship of stimulus pairs. This type of encoding task, directing attention to a single aspect of synonymy, may actually operate to depress the contrastive analysis given synonymous pairs in a less structured task; comparisons may focus solely on differences in specificity of reference and not consider other contrastive information. The net effect of such a narrow focus would be to make it difficult to locate the intended trace; once it has been located, the highly specific information of the trace makes it easy to reject alternative response candidates. That such an effect may be operating is supported by the fact that subjects do not falsely recognize distractors and by the superiority of cued recall to free recall.

The results of the synonymy comparisons extend the applicability of the specificity of encoding hypothesis to include relatively small differences in specificity, when the basic similarity in meaning is already quite high. The results also indicate that specificity of encoding affects trace distinctiveness when the comparison is more abstract and inferential in nature than had been established. In addition, the results establish the fact that encoding may be specific without being complete.

The generality of the results with synonymy may be limited to cases in which the comparisons focus on a specific aspect of synonymy. When the comparisons allow synonyms to be examined in a less restricted manner, the *a priori* similarity in meaning of synonyms may override potential effects of specificity of encoding. A later study not reported in the thesis seems to confirm that synonymy has limited utility as a dimension of comparison. Because of such potential problems, synonymy is given no further consideration for the course of this investigation.

CHAPTER 4

EXPERIMENT 3

The third experiment expanded the investigation of specificity of encoding by introducing a new variable of potential significance, congruity of encoding. The dimension of comparison in the rating task was varied so that the rating task was either congruous with the structure of the stimulus list (based on referent size or pleasantness) or incongruous with the list structure but requiring a specific semantic comparison (e.g., rating the size list on the dimension of pleasantness).

The earlier experiments in the thesis established that, to be effective, the specificity of encoding must be referential in nature; specific contrasts about nominal aspects of the stimulus generally do not result in effective memory traces (cf. Hunt and Elliott, 1981). Quite small differences in specificity of encoding may have a substantial effect on the discriminability, accessibility, and cohesion of the memory trace. Specificity is effective for extensive ranges of encoding, from the very general such as correcting the spelling of a word (Jacoby et al., 1979, Exp. 1) to the highly specific task of choosing the more specific referent from synonyms (Exp. 2 of the thesis). Thios (1975) even demonstrated that the specificity of the verb influenced the level and accuracy of subsequent recall of sentences. Clearly, however, specificity is not the sole factor of the importance in determining the distinctiveness of a memory trace, and specificity cannot be considered in isolation. The experiments thus far presented in the thesis have kept a crucial factor, that of task congruity, constant across the various dimensions of comparison.

The concept of congruity as used here refers to a correspondence in character between the encoding context and the targets for encoding, or compatibility of a stimulus structure and the encoding task. In general, such congruity results in higher retrieval than does incongruity of task and stimulus. Goldman and Pellegrino (1977) and Moscovitch and Craik (1976, Exp. 1) found congruity of encoding question and target resulted in higher retrieval levels than did incongruity.

Schulman (1971) had subjects scan a list of words, searching for targets congruous with the task dimension, for one of several different dimensions, among them, the letter A present, a living thing, or a geographical location. After each of the tasks mentioned, the recognition level was higher if a word had been a target of the search criterion, congruous with the task. Schulman (1971) attributes the superiority to the more distinctive processing a congruous target receives while being categorized.

Schulman (1974) examined congruity between subject and descriptor in questions such as "Is a bubble a sphere?" or "Is a chapter slippery?". He presented subjects with 100 such questions and required quick responses. A congruous descriptor was a defining attribute or superordinate of the subject of a question. Schulman found congruous questions resulted in higher cued recall, free recall, and recognition levels than did incongruous questions. He suggested that a congruous question fostered a relational encoding between subject and descriptor while an incongruous question encouraged an independent encoding.

Auble et al. (1979, Exp. 2) present evidence related to an examination of congruity and decision difficulty. They presented subjects with sentences that were easy or difficult to comprehend, followed by appropriate or inappropriate postcues. As an example, consider the sentence "The notes were sour because the seam split" in the context of an inappropriate cue, "wet paint." The sentence remains virtually meaningless when the cue is inappropriate to the intended meaning of the sentence. Now consider the same sentence in the context of the appropriate cue, "bagpipes." The congruous context provides a means for increasing the specificity of the encoding by establishing the relationships intended in the sentence. Recall of difficult sentences was higher if the sentences were followed by appropriate rather than inappropriate cues. No such effect of congruity was found for recall of the easy sentences. Morris et al. (1977, Exps. 2 and 3) also found a beneficial effect of the congruity of targets and sentence frames on subsequent recall.

Mathews (1977, Exp. 2) found that incongruity of encoding context and stimulus structure depressed subsequent recall, even when the words paired during encoding were members of a single category. Apparently the focus created during the encoding task affects the later utility of semantic relations.

Schnur (1977, Exp. 1) examined recognition for words following category decisions. Congruous decisions involved high-, intermediate-, or low-ranking exemplars of a category, while incongruous decisions involved such exemplars from a different category. For congruous decisions, the exemplar level should influence the difficulty of a category decision. Recognition data confirmed that for congruous decisions, recognition level increased with a decrease in exemplar ranking. Data for the incongruous decisions revealed no such effect of exemplar ranking.

Hall and Geis (1980) examined the effects of sentence complexity and congruity on recall following a sentence completion task. Congruous targets were recalled more frequently than incongruous targets for all levels of complexity tested. An increase in sentence complexity resulted in higher recall only when the sentence and target were congruous (Exp. 2). Hall and Geis attribute the superiority in recall of congruous targets to greater integration of frame and target.

The evidence cited suggests that congruity between stimulus and task has a powerful effect on the integration and specificity of an encoding. The impact on the memory trace of a factor such as specificity of encoding may be altered by a lack of congruity between stimulus and task structures. Congruity in the context of the present experiment may be defined as the correspondence between stimulus list structure and the relevant dimension of comparison employed in the rating task.

In addition to the theoretical issues addressed by a consideration of specificity in conjunction with congruity, the experiment varied procedure to clarify some empirical issues. Two different encoding tasks were used. Some subjects chose one member of a stimulus pair as being larger or more pleasant and rated the difficulty of the decision, as in the first two experiments. Other subjects rated the difference in referent size or pleasantness between the members of a pair, as in Jacoby et al. (1979, Exp. 2). The variation of rating task allows a comparison of the effect of specificity for the two tasks. It might be argued that the difficulty task, by requiring that subjects choose between members of a pair, may undermine the specificity of encoding if the pair is easily distinguished on the dimension of comparison. If the difference-rating task yields a similar pattern of memory performance, then the argument is unsubstantiated. The dissimilar pairs must be encoded to a considerable degree to assign a rating of the degree of difference. However, specificity of encoding would still be higher to ascertain the relevant difference between more similar pairs.

The data from the retention tests used, cued recall and matching, should provide a clearer picture of the effect that specificity of encoding has on memory performance. Cued recall, with no distractors to intefere with cue recognition, should be a truer measure of response production as influenced by specificity of encoding. The matching task, by providing both pair members, yields a measure of the pair association for a stimulus pair without problems caused by failure in other aspects of cued recall, such as cue recognition or response generation.

Pleasantness was chosen to be one of the dimensions of comparison used in this experiment. A brief review of the literature suggests that pleasantness is a salient aspect of meaning, a promising candidate to affect specificity (see Appendix 3 for a pilot experiment using pleasantness). Paivio (1978, Exp. 1) established that subjects required more time to choose between two words highly similar in rated pleasantness than to choose between words less similar in pleasantness. Decision difficulty and, by extension, specificity of encoding, seem to be affected by similarity in pleasantness. Several investigators have found higher recall levels following pleasantness ratings than following other semantic rating tasks (e.g., Hyde, 1973; Hyde & Jenkins, 1973; Postman & Kruesi, 1977; Walsh & Jenkins, 1973). Such evidence suggests that pleasantness is a salient aspect of meaning.

Packman and Battig (1978) found that words rated for pleasantness were recalled more frequently than words rated on several other semantic dimensions (e.g., meaningfulness, categorizability, imagery, familiarity, concreteness). Pleasantness ratings had the smallest correlation with the other dimensions tested. Packman and Battig argue that it is the unrelated nature of pleasantness relative to other semantic dimensions that results in higher memory; items rated for pleasantness are encoded more distinctively than are items rated on a dimension more related to other semantic dimensions.

Although evidence suggests that similarity in pleasantness affects specificity of encoding, it is important to note that specificity in affective comparisons may not affect memory in the direct manner found for size comparisons. The affective component may not form an essential part of the memory trace, thereby limiting the potential distinctiveness of the trace. There is no obvious mental affective continuum along which to compare individual referents; a word may be rated for pleasantness only if the word has a strong affective component (i.e., the pleasantness or unpleasantness is a salient aspect of the word's meaning). Johnson-Laird et al. (1978) suggest that "Tasks such as rating items as pleasant or unpleasant tap genuine semantic components for only a small proportion of words" (p. 372). If difficult pleasantness comparisons do not require specific analysis, or if the result of the comparison does not in part constitute the trace, no differential effects of specificity of encoding will result.

Method

Subjects. Fifteen subjects participated in each of 16 conditions, eight conditions with each of two presentation lists. The 240 subjects were introductory psychology students who participated in the experiment as a course requirement. Subjects were tested in groups, with each group randomly assigned to a condition.

Materials. Two presentation lists were constructed, one each for the dimensions of size and pleasantness. Pairs in the size list were composed of nouns from Paivio's (1975) size norms. Pairs in the pleasantness list were composed of nouns from Paivio's (1979, Reference Note 1) pleasantness norms. (See Appendix 4 for a complete listing of the stimulus pairs used in Exp. 3).

For each set of norms used in constructing the lists, the set was broken down into deciles (small to large, unpleasant to pleasant). Three ranges of decile separation were chosen to represent similar, intermediate, and dissimilar pairs on the relevant dimension. Members of a stimulus pair were from the same decile (0) or differed by 3 or 6 deciles (e.g., for the size list, SNAIL-PEANUT, EAGLE-HORSE, PRUNE-DESK). Eight pairs were constructed for each decile separation.

For the similar pairs, the members of a pair were from the same decile of the appropriate norms list. One pair of items was chosen from each of deciles 1 through 4 and 7 through 10. Items chosen for a pair were nearly identical in normed rating on the relevant dimension.

The intermediate pairs in the lists were items representing a 3-decile separation in the norms. Two pairs were constructed of items from deciles 1 and 4, and two were constructed of items from deciles 7 and 10, to maintain equal representation from the extremes of the relevant dimension. One pair was constructed of items from each of decile combinations 2 and 5, 3 and 6, 5 and 8, and 6 and 9.

Dissimilar pairs were items representing a 6-decile separation in the norms. Two pairs were constructed from items from each of decile combinations 1 and 7, 2 and 8, 3 and 9, and 4 and 10.

Ten buffer pairs were constructed for each list with 3 pairs representing similarity in size or pleasantness, 4 representing intermediate similarity, and 3 representing dissimilarity between members of a pair.

Each presentation list was arranged so that five buffer pairs preceded presentation of the critical pairs, and five buffer pairs followed. The critical pairs were ordered in four randomized blocks of six pairs each, two pairs from each level of similarity represented. The blocks provided early exposure to the entire range of similarity present in the list. Each block of buffer pairs contained at least one pair from each level of similarity.

The presentation lists were videotaped at a presentation rate of five seconds per pair.

A cued recall list was constructed for each presentation list. One member from each critical pair was presented as a cue for recall. Half of the cues had been the left-hand member of the presented pair, and half had been the right-hand member. Half of the cues were the larger or more pleasant member of a presented pair, and half of the cues were the smaller or less pleasant.

The matching sheets were constructed by placing the cues for the cued recall test list on the left side of the page, and placing the target members of presented pairs, in a different random order, on the right side of the page, numbered as to order on the page.

<u>Experimental Conditions</u>. The combination of the factors of presentation list (2), dimension of comparison (2), rating task (2), and memory test (2) creates 16 experimental conditions.

The dimensions of comparison used were referent size and pleasantness. Subjects rated pairs in the size or pleasantness presentation list along one of the two dimensions. The rating task required that subjects either choose one member of a pair as being the larger or more pleasant and rate the difficulty of the decision, or that they rate the degree of difference between pair members on the assigned dimension. Following completion of the assigned rating task, subjects were given one of two memory tests, cued recall or matching.

<u>Procedure</u>. Subjects were shown one of two presentation lists.

Subjects were told they would be shown 34 pairs of words, at a presentation rate of five seconds per pair. Then they were given the set of instructions appropriate to their rating task condition.

Subjects in the <u>difficulty</u> rating conditions were told to choose the larger (or more pleasant) member of a presented pair. Then they were to rate the difficulty of the decision from 1 for a very difficult decision to 7 for an extremely easy one. Subjects were asked to make a decision and rating for each pair as it appeared.

Subjects in the difference rating conditions were asked to rate each word pair as it appeared for the degree of difference in referent size (or pleasantness). The rating scale they were to use ranged from 1 for virtually no difference between pair member to 10 for a vast difference.

Following the completion of a rating task, subjects were given instructions and the appropriate test sheets for one of the two memory tests.

Subjects given the cued recall test were asked to write in the word that had been paired with each of the cues during presentation. Subjects were encouraged to recall as many words as possible.

Subjects given the matching test were instructed to reconstruct pairs from the presentation list by placing the number of a word from the right side of the sheet next to the appropriate word on the left side. Subjects were encouraged to recreate as many pairs as possible.

Subjects were given five minutes to complete a test.

Results and Discussion

Results are reported separately for each of the presentation lists. Comparisons across lists will also be made.

The .05 α level was used for all tests of significance. Scheffé tests were used to confirm those differences determining significant effects.

Size List

Rating. Table 8 presents mean rating of pair difference and mean rating of decision difficulty as a function of pair separation (inversely related to similarity in size) for each dimension of

Table 8. Mean rating of decision difficulty or pair difference as a function of decile separation for the dimensions of comparison involving the size list in Experiment 3. Standard deviations are in parentheses.

DECISION DIFFICULTY (Scale 1-7)

Comparison Dimension	DECILE SEPARATION 0	3	6
Congruous (Size)	4.03 (1.40)	5.79 (0.89)	6.43 (0.66)
Incongruous (Pleasantness)	4.23 (0.89)	5.08 (0.78)	4.52 (0.95)

PAIR DIFFERENCE (Scale 1-10)

Comparison Dimension	DECILE SEPARATION 0	3	6
Congruous	3.40 (0.99)	5.76 (1.25)	8.31 (1.41)
Incongruous	4.96 (1.69)	4.73 (1.71)	5.43 (1.66)

comparison. Values for difficulty and difference ratings are not directly comparable, as the rating scales are not of equal range. Both types of rating demonstrate a clear rise in rating (greater pair difference or less difficult decision) with a decrease in similarity for the corresponding size-congruous condition. Rating patterns of the incongruous conditions do not correspond in form.

Rating results were analyzed separately for the two types of rating task. A 2x3 mixed design analysis of variance performed on the data for rating decision difficulty found that task congruity [$F(1,58)=17.7$, $MSe=105$], similarity in size [$F(2,116)=85.9$, $MSe=26.2$] and their interaction [$F(2,116)=41.4$] all had significant effects on difficulty rating. Congruous decisions were rated as being easier than were incongruous ones. Pairs that were similar in size were rated as being more difficult to distinguish than were pairs that were intermediate or dissimilar. When the dimension of comparison was congruous with the list structure, each level of similarity was perceived as being distinct in degree of difficulty. When the dimension of comparison was incongruous with the list structure, members of intermediate-similarity pairs were perceived as being less difficult than were members of pairs from the extremes of similarity in size.

A 2x3 mixed design analysis of variance performed on the data for difference rating reveals that congruity [$F(1,58)=2.59$, $MSe=637$] had no significant effect on rating, but that similarity in size [$F(2,116)=153$, $MSe=40.4$] and the interaction of similarity and congruity [$F(2,116)=132$] had significant effects on rating of pair difference. Pairs from each level of similarity were rated as being distinct in degree of difference, with rating of difference increasing as similarity decreased. Items from similar pairs were rated as being less different when the dimension of comparison was congruous rather than incongruous with the list structure; however, the items from intermediate and dissimilar pairs were perceived as being more different when the dimension of comparison was congruous with the list structure. Pairs of intermediate similarity were rated as being less different than dissimilar pairs, regardless of the basis of comparison. The pattern is unlike that for the difficulty ratings.

The results for the two rating tasks are highly similar when the dimension of comparison is congruous with the list structure. Each level of similarity in size is perceived as being distinct in terms of pair difference or decision difficulty, with the pattern of rating reflecting the similarity directly.

Cued Recall. The results for cued recall following difference or difficulty rating are presented in Table 9. Overall, the tasks of rating decision difficulty and pair difference result in similar patterns of recall. Recall in the congruous conditions decreases as similarity in size decreases, but the decrease in recall is sharper for the difficulty task than for the difference task. Recall in the incongruous conditions is higher for items from the extremes of similarity in size than for items of intermediate similarity.

A 2x2x3 mixed design analysis of variance was performed on the recall data. The between-subject variables of congruity [$F(1,56)=5.56$, $MSe=1.52$] and rating task [$F(1,56)=9.51$] had significant effects on recall, while their interaction ($F<1$) did not. Congruous rating tasks resulted in higher levels of recall than did incongruous rating tasks. The task of rating pair difference resulted in a higher level of correct

Table 9. Mean number of words correctly recalled (of a possible 8) as a function of decile separation for congruous and incongruous conditions following rating for decision difficulty or pair difference of the size list in Experiment 3. Standard deviations are in parentheses. Probability of correct recall is indicated by an asterisk.

DECISION DIFFICULTY

Comparison Dimension	DECILE SEPARATION 0	3	6
Congruous	2.20 (1.72) *.28	0.47 (0.62) *.06	0.53 (0.50) *.07
Incongruous	0.73 (0.77) *.09	0.33 (0.47) *.04	0.60 (0.61) *.08

PAIR DIFFERENCE

Comparison Dimension	DECILE SEPARATION 0	3	6
Congruous	1.73 (1.00) *.22	1.60 (1.25) *.20	1.33 (0.79) *.17
Incongruous	1.00 (0.89) *.13	0.73 (1.18) *.09	1.87 (1.09) *.23

recall than did the task of rating decision difficulty. Similarity in size [$F(2,112)=7.92$, $MSe=0.76$], as well as the interactions of similarity with congruity [$F(2,112)=9.74$] and with rating task [$F(2,112)=6.92$], significantly affected recall. The interaction of all three variables was marginally significant [$F(2,112)=3.01$, $p.=.053$].

Items from pairs that were similar in size were recalled more frequently than were items from intermediate or dissimilar pairs, when the dimension of comparison was congruous with the list stucture. When comparisons were made on an incongruous basis, items from dissimilar pairs were recalled more frequently than items of intermediate similarity.

Although recall was higher when pairs had been rated for difference rather than for difficulty, the effect of similarity on pattern of recall in congruous conditions was greater when difficulty had been rated. Recall was equivalent at all levels of similarity following rating of difference in size.

The results for cued recall confirm the predictions of a specificity hypothesis. Congruous encoding of stimulus pairs, with processing directed toward the salient quality of similarity in size, resulted in a higher level of recall. The superiority of congruous encoding occurred in the condition requiring difficulty rating, even though it had been rated as easier than the corresponding incongruous difficulty rating. The difficulty that subjects encountered in rating size pairs on the dimension of pleasantness was not of a beneficial nature; the contrasts made in differentiating members of a pair did not result in a distinctive memory trace of the pair.

Specificity of encoding affected the pattern of recall in the expected manner when the dimension of comparison focused on similarity in size. The pattern of recall following a congruous encoding was stronger when the rating task called for a rating of difficulty rather than of pair difference. It is suggested that the difficulty task emphasizes the differential similarity in size between pairs from different similarity levels, thereby emphasizing the relative distinctiveness of comparisons. The more specific judgments (across similarity levels) called for in the difference task are reflected in a higher overall level of recall, and an attenuated effect of similarity in size.

The cued recall data are highly consistent with the results reported in Experiment 1, enhancing the stability of the effect of specificity on memory when the encoding task creates a specific, objective semantic focus.

<u>Matching</u>. Table 10 presents the mean number of pairs correctly recreated as a function of pair separation for congruous and incongruous conditions following the task of rating pair difference or decision difficulty. The congruous conditions show a clear effect of similarity in size on correct matching, with probability of correct matching decreasing as similarity decreases. Although only very slightly, incongruous conditions show items from the extremes of similarity being matched more frequently than are pairs of intermediate similarity. Considerably more pairs are reconstructed by the subjects who rated pairs for difference rather than for difficulty.

A 2x2x3 mixed design analysis of variance was performed on the matching data. The between-subject factor of type of rating, difficulty

Table 10. Mean number of pairs correctly matched (of a possible 8) as a function of decile separation for congruous and incongruous conditions following rating for decision difficulty or pair difference of the size list in Experiment 3. Standard deviations are in parentheses. Probability of correct pair reinstatement is indicated by an asterisk.

DECISION DIFFICULTY

Comparison Dimension	DECILE SEPARATION 0	3	6
Congruous	3.20 (1.83) *.40	1.00 (1.15) *.13	1.00 (1.10) *.13
Incongruous	1.67 (1.62) *.21	1.60 (1.02) *.20	1.80 (1.42) *.23

PAIR DIFFERENCE

Comparison Dimension	DECILE SEPARATION 0	3	6
Congruous	3.07 (1.53) *.38	2.40 (1.54) *.30	1.80 (1.22) *.23
Incongruous	2.40 (1.58) *.30	2.20 (1.42) *.28	2.53 (1.50) *.32

or difference, [$F(1,56)=4.80$, $MSe=4.45$] had a significant effect on correct matching of pair items, but neither congruity nor the interaction of congruity and type of rating (F's<1) affected matching significantly. Rating pair difference resulted in greater success at matching than did rating decision difficulty. Similarity in size [$F(2,112)=11.8$, $MSe=1.06$] affected matching success, with pairs of high similarity in size being matched correctly more frequently than pairs less similar in size. The interaction of similarity and congruity [$F(2,112)=13.0$] had a significant effect; the effect of more specific encoding is confined to conditions in which the dimension of encoding is congruous with the list structure. The interaction of similarity and rating task [$F(2,112)=1.80$] and the interaction of all three factors [$F(2,112)=2.46$] did not affect correct matching.

Clearly, specificity of encoding affected the correct recreation of stimulus pairs. The similar pairs, those requiring a more specific comparison, were matched correctly more frequently than were less similar pairs. Difference ratings resulted in a higher overall level of correct matching. However, the two rating tasks resulted in highly similar patterns of correct matching.

Again, predictions are confirmed. Patterns of matching (pair association) performance are consistent with the patterns found in cued recall, for congruous and incongruous encoding conditions. The cohesion of the pair trace was affected by specificity of encoding in the congruous conditions, confirming the results reported for free recall in Experiment 1.

Pleasantness List

Rating. The results for rating of pair difference and of decision difficulty for the pleasantness list are presented in Table 11. Both types of rating display a clear increase in rating (increased difference, decreased difficulty) with a decrease in pair similarity, when the dimension of comparison is congruous with the list structure. When the task and structure are incongruous, pairs are rated as being slightly more different or less difficult when the pairs are from the extremes of similarity in pleasantness.

A 2x3 mixed design analysis of variance was performed on the data for rating of decision difficulty. The factors of congruity [$F(1,58)=39.6$, MSe=104], similarity in pleasantness [$F(2,116)=21.5$, MSe=13.3], and their interaction [$F(2,116)=9.02$] all had significant effects on rating of decision difficulty. Congruous decisions were rated as being more difficult than were incongruous ones. Recall that the reverse was true with the size list, suggesting that pleasantness comparisons are somewhat difficult to perform. The level of decision difficulty as rated conformed with the similarity of the pair, with each level being rated as of distinct difficulty. Subjects in the incongruous conditions did not perceive the various levels of similarity as requiring decisions of differential difficulty.

A 2x3 mixed design analysis of variance performed on the data for rating pair difference reveals that congruity [$F(1,58)=4.88$, MSe=414], similarity [$F(2,116)=22.7$, MSe=42.3], and their interaction [$F(2,116)=15.6$] all significantly affected rating of pair difference. Incongruous decisions were rated as being more different than were congruous decisions. When task and structure were congruous, the

Table 11. Mean rating of decision difficulty of pair difference as a function of decile separation for the dimensions of comparison involving the pleasantness list in Experiment 3. Standard deviations are in parentheses.

DECISION DIFFICULTY

Comparison Dimension	DECILE SEPARATION 0	3	6
Congruous (Pleasantness)	3.90 (0.79)	4.43 (1.06)	6.18 (0.78)
Incongruous (Size)	5.50 (0.92)	5.40 (1.06)	5.75 (0.87)

PAIR DIFFERENCE

Comparison Dimension	DECILE SEPARATION 0	3	6
Congruous	4.99 (1.49)	5.60 (1.28)	6.55 (1.41)
Incongruous	6.78 (1.38)	6.01 (1.61)	6.86 (1.59)

dissimilar pairs were perceived as being more different than pairs from other levels of similarity. When task and structure were incongruous, pairs from the extremes of similarity were perceived as being more different than pairs of intermediate similarity.

The results for the two rating tasks are highly similar. The ratings of difficulty and of difference correspond with the similarity of the pairs when the dimension of comparison is congruous with the list. The ratings for incongruous conditions suggest that pairs from the extremes of similarity are perceived as being more different than are pairs of intermediate similarity. Congruous decisions are perceived as being more difficult than are incongruous ones. Subjects appear to find the task of comparing items on the affective dimension of pleasantness more difficult than comparing the same items on the objective dimension of size.

Cued Recall. The results for cued recall of the pleasantness list are presented in Table 12. Recall following difference rating increases very slightly with a decrease in pair similarity regardless of whether the dimension of comparison was congruous or incongruous with the list structure. Following difficulty rating, items from the extremes of similarity are recalled more frequently than intermediate items following congruous encoding, while the pattern is reversed following incongruous encoding.

A 2x2x3 mixed design analysis of variance was performed on the recall data. The rating task significantly affected recall [$F(1,56)=10.8$, $MSe=2.52$]; correct recall was higher following difference rather than difficulty rating. Neither congruity nor its interaction with rating task affected recall significantly (F's<1). Similarity in

Table 12. Mean number of words correctly recalled as a function of decile separation for congruous and incongruous conditionings following rating of decision difficulty or pair difference of the pleasantness list in Experiment 3. Probability of correct recall is indicated by an asterisk.

DECISION DIFFICULTY

	DECILE SEPARATION		
Comparison Dimension	0	3	6
Congruous	0.53 (0.72) *.07	0.27 (0.44) *.03	0.53 (0.72) *.07
Incongruous	0.33 (0.60) *.04	0.87 (0.81) *.11	0.60 (0.95) *.08

PAIR DIFFERENCE

	DECILE SEPARATION		
Comparison Dimension	0	3	6
Congruous	1.13 (1.50) *.14	1.27 (1.29) *.16	1.60 (1.82) *.20
Incongruous	1.07 (1.00) *.13	1.13 (1.20) *.14	1.60 (0.88) *.20

pleasantness [$F(2,112)=2.75$, MSe=0.56] had no significant effect on recall; neither did the interaction of similarity and congruity ($F<1$), of similarity and rating task [$F(2,112)=1.32$] or the interaction of all three variables [$F(2,112)=1.38$] affect recall.

Matching. The results for matching following difficulty or difference rating are presented in Table 13. None of the conditions give indications of similarity affecting the cohesion of the pair trace. A 2x2x3 mixed design analysis of variance confirms that the task of rating pairs for difference rather than difficulty [$F(1,56)=14.0$, MSe=3.57] resulted in a higher level of success at matching pair members. Congruity [$F(1,56)=2.13$] did not affect pair matching, neither did the interaction of congruity and rating task ($F<1$). Similarity in pleasantness ($F=1$) had no effect on matching pairs; neither did any interaction of similarity with congruity and/or rating task (F's<1) affect matching.

The lack of significant effects of congruity and similarity, alone and in combination, make it implausible to assert that the task of rating difficulty has resulted in a more defined effect of specificity of encoding on trace distinctiveness. A conservative conclusion is that the task of comparing pair members in terms of pleasantness is not performed consistently by subjects. The absolute pleasantness value of an item would seem to be more salient in encoding and in memory than is the relative pleasantness of pair members. A specific comparative decision does not necessarily yield a cohesive pair trace relative to a less specific comparison. That pairs of low similarity in pleasantness are not always recalled more frequently than more similar pairs suggests that specificity of encoding may operate to some extent in pleasantness

Table 13. Mean number of pairs correctly matched as a function of decile separation for congruous and incongruous conditions following rating for decision difficulty of pair difference of the pleasantness list in Experiment 3. Probability of correct pair reinstatement is indicated by an asterisk.

DECISION DIFFICULTY

	DECILE SEPARATION		
Comparison Dimension	0	3	6
Congruous	1.00 (0.89) *.13	0.80 (0.98) *.10	1.33 (1.01) *.17
Incongruous	1.53 (1.26) *.19	1.20 (1.76) *.15	1.13 (1.26) *.14

PAIR DIFFERENCE

	DECILE SEPARATION		
Comparison Dimension	0	3	6
Congruous	2.20 (1.28) *.28	1.86 (1.19) *.23	1.87 (1.26) *.23
Incongruous	2.53 (1.63) *.32	2.47 (1.36) *.31	2.53 (1.78) *.32

comparisons. The dissimilar pairs necessarily represent items of extreme pleasantness more often than do more similar pairs. The relative ease with which such disparate items are compared undermines the distinctiveness such extreme items typically receive individually when encoded for pleasantness.

Congruity in the context of these comparison tasks does not necessarily result in a higher level of recall than does incongruity, but it does result in a distinct pattern of recall. The difference and difficulty tasks show similar patterns of recall; either encoding task shows a similar effect of specificity of encoding when the list varies similarity in referent size. The retention tests of cued recall and matching show similar effects of specificity of encoding. These factors considered together suggest a strong effect of specificity of encoding when not masked by the vagaries of pleasantness comparisons.

It may be argued that the task of rating pairs for degree of difference results in a higher overall level of specificity for pairs of intermediate and low similarity. Because subjects are not choosing one member as representing the larger or more pleasant of the pair, noting a clear disparity may not suffice in making a rating. Instead, more detailed contrasts may be used to examine the pair for degree of difference. Such an increase in overall specificity would have a twofold effect on subsequent memory for items from the list. First, the task of rating difference would result in higher overall retention because the more specific encoding of dissimilar items yields more distinctive memory traces, relative to the difficulty task. Difference rating resulted in higher retention levels than did difficulty rating for both lists and memory tests. Second, the effect that similarity has on memory would be mitigated to some extent by the more specific encoding that pairs of intermediate and low similarity receive when compared for degree of difference. Such an attenuation of the effect of specificity was significant only for the cued recall data with the size list, although the matching data from the size list are consistent with an attenuation.

Experiment 3 examined congruity of task and structure; such congruity resulted in distinct patterns of recall. The level of retention for pairs encoded incongruously was not necessarily lower than for pairs encoded congruously. As the dimension of comparison could be applied meaningfully, though not distinctively, when incongruous with the structure, the pairs are encoded fairly completely (see McClelland, Rawles, and Sinclair, 1981, and Roenker, Wenger, Thompson, and Watkins, 1978, for other demonstrations that incongruous encoding does not necessarily result in a lower level of recall.) The encoding process did not result in anomaly or a very general level of semantic encoding. That specificity of encoding appears to influence memory only when the stimulus is congruous with the encoding context is examined further in the fourth experiment.

CHAPTER 5

EXPERIMENT 4

Experiment 3 demonstrated that congruity of encoding task and stimulus structure resulted in a distinct pattern of retention. Specificity of encoding seems to influence the distinctiveness of the memory trace only when the stimulus structure is congruous with the encoding task. Such a finding is somewhat unusual because incongruity in the context of Experiment 3 did not result in an anomalous or superficial encoding of stimuli. The fourth experiment was conducted to explore more fully the influence that incongruity has on specificity of encoding. An additional level of specificity is considered to clarify the joint examination of congruity and specificity. The dimension of comparison in the encoding task was varied so that the rating task was congruous with the structure of the stimulus list, incongruous but requiring comparisons on a specific semantic dimensions, or general in nature (e..g, understanding, pronunciation).

Aside from the addition of general encoding dimensions, the experiment is similar to Experiment 3. The stimulus lists, varied as to similarity in referent size or pleasantness, were those used in Experiment 3. The difference rating task was used, as in Experiment 3. As the difference task resulted in higher levels of recall, but a similar pattern, relative to the rating of decision difficulty, the use of the difference task might allow for more distinction of relative levels of recall.

The utility of the concept of specificity of encoding in explaining human memory would be hampered by a demonstration that specificity of encoding affects the distinctiveness of the memory trace only when the encoding task corresponds directly with the underlying stimulus structure. Comparisons of specific encoding conditions with general encoding conditions should negate the possibility that specificity of encoding is a useful concept only when nearly perfect congruity of stimulus and encoding exists. It is expected that the specific encoding conditions should result in higher levels of recall than general encoding conditions, even when the specific encoding is incongruous with the stimulus structure. Incongruity in the context of this experiment undermines specificity of encoding in relatively minor ways. When the specificity of encoding does not correspond with the stimulus structure, a conflict exists. The potential distinctiveness of a specific encoding is undermined by the salient characteristics of the underlying structure. Aspects of meaning irrelevant to the encoding task affect the distinctiveness of the memory trace. Despite this conflict between task and structure, the greater specificity of encoding called for by a specific rather than general comparison should result in more specific encoding of stimuli and more distinctive memory traces. The higher level of recall following specific tasks might indicate that specificity of encoding is still operative in some incongruous situations, but that its potential effect is masked to some extent by the lack of correspondence with stimulus structure.

Method

Subjects. Twenty subjects participated in each of eight conditions, four conditions with each of the two stimulus lists. The 160 subjects were introductory psychology students who participated in the experiment as a course requirement. Subjects were tested in groups, with each group randomly assigned to a condition.

Materials. The two stimulus lists were those used in Experiment 3. One list varied pairs of words as to similarity in referent size; the other varied pairs as to similarity in pleasantness.

The cued recall test lists from Experiment 3 were used for the recall tests.

Rating Dimensions and Experimental Conditions. Subjects in a condition rated the degree of difference between pair members (of one stimulus list) along one of four dimensions, the dimensions being referent size, pleasantness, ease of understanding, or ease of pronunciation.

The combination of stimulus list and rating task results in eight conditions. For each stimulus list, one rating task represents congruity of list structure and rating task (e.g., rating stimuli in the size list for differences in referent size), one condition represents incongruity, but specificity, of structure and task (e.g., rating stimuli in the size list for difference in pleasantness), and two conditions represent general processing (i.e., rating stimuli in either list for difference in ease of understanding or pronunciation).

Procedure. Subjects were shown one of the two presentation lists.

Subjects were told they would be shown 34 pairs of words. When a pair appeared on the screen they were to indicate the amount of difference between the pair members, along a given dimension, by giving the difference a rating on a 10-point scale. On the scale 1 represented virtually no difference and 10 represented a vast difference.

The dimension along which subjects rated difference in pair members differed across conditions. According to the assigned condition, subjects were told to base their difference rating on difference in referent size, pleasantness, difficulty in understanding, or difficulty in pronunciation.

After they completed the appropriate rating task, subjects were given an unexpected cued recall test. Subjects were given 5 minutes to complete the recall test.

Results and Discussion

Results are reported separately for each of the presentation lists. Comparisons will later be made across lists.

The .05 α level was used for all tests of significance. Scheffé tests were used to confirm the specific differences accounting for significant effects.

Similarity in Size

Rating. Table 14 presents mean rating of pair difference as a function of similarity in size (decile separation) for each of the dimensions of comparison. The size-congruous condition exhibits a sharper increase in difference rating with decreased similarity than do the other conditions. The distinct pattern of the congruous condition demonstrates a high degree of sensitivity to the composition of the list when rating is done on a dimension congruous with the structure of the list. The specific-incongruous condition (pleasantness) demonstrates a slight rise in difference rating with a decrease in similarity, a rise not found in Experiment 3. The general encoding conditions (understanding, pronunciation) result in lower ratings of pair

Table 14. Mean rating of pair difference as a function of decile separation in size for each of the dimensions of comparison involving the size list in Experiment 4. Standard deviations are in parentheses.

Comparison Dimension	DECILE SEPARATION IN SIZE 0	3	6
Congruous			
Size	4.29 (1.41)	5.83 (0.98)	8.19 (0.98)
Incongruous: Specific			
Pleasantness	5.91 (1.98)	6.26 (1.42)	6.94 (1.78)
Incongruous: General			
Understanding	3.48 (1.51)	3.89 (1.90)	4.23 (2.12)
Pronunciation	4.09 (1.37)	4.04 (1.37)	4.28 (1.50)

difference, with the understanding condition displaying a slight increase in rating with a decrease in similarity. Rated difference in the pronunciation condition shows no effect of similarity in size.

A 4x3 mixed design analysis of variance was performed on the rating data for the size list. The factors of rating dimension [$F(3,76)=16.0$, MSe=405], pair similarity [$F(2,152)=62.3$, MSe=44.9] and the interaction of dimension and similarity [$F(6,152)=20.0$] all significantly affected difference rating. Difference ratings for one specific semantic condition was distinct from the rating of the other; the specific-incongruous condition resulted in pairs being perceived as more different than did the congruous condition. Pair members were rated as being more different from each other in the specific conditions than in the general encoding conditions.

Decreased similarity in size resulted in higher difference ratings. The relationship between similarity and rating was particularly strong for the congruous condition, suggesting that the difference ratings are more sensitive when the dimension of comparison is appropriate to the list structure.

Cued Recall. Table 15 presents mean number of words correctly recalled, as well as the probability of correct recall, as a function of similarity in size for each of the four rating conditions using the size list. The congruous condition demonstrates a decrease in correct recall with a decrease in specificity of encoding. The decrease in recall is confined to the difference in recall for the similar and intermediate items, as the items from dissimilar pairs are recalled as frequently as are intermediate items. The remaining semantic conditions display

Table 15. Mean number of words correctly recalled (of a possible 8) as a function of decile separation in size (inversely related to similarity in size) for each of the dimensions of comparison involving the size list in Experiment 4. Standard deviations are in parentheses. Probability of correct recall for each value is indicated by an asterisk.

	DECILE SEPARATION IN SIZE		
Comparison Dimension	0	3	6
Congruous			
Size	1.95 (2.20) *.24	1.35 (1.68) *.17	1.45 (1.77) *.18
Incongruous: Specific			
Pleasantness	1.40 (1.11) *.18	0.70 (1.14) *.09	1.50 (1.07) *.19
Incongruous: General			
Understanding	1.45 (1.43) *.18	1.05 (0.97) *.13	1.30 (1.27) *.16
Pronunciation	0.35 (0.57) *.04	0.30 (0.46) *.04	0.60 (0.73) *.08

higher recall for items from the extremes of similarity in size than for intermediate items.

Frequency of correct recall in the specific-incongruous condition is not directly related to difference rating, as a measure of specificity of encoding, or to similarity in size. The relationship between specificity of encoding and recall for the congruous condition is consistent with a specificity of encoding hypothesis.

A 4x3 mixed design analysis of variance was performed on the recall data. The results of the analysis indicate that rating dimension [$F(3,76)=3.70$, MSe=3.98] and similarity in size [$F(2,152)=6.55$, MSe=0.67], but not their interaction [$F(6,152)=1.41$] are significant variables. The significance of rating dimension is attributed to cued recall being higher for the congruous condition than for understanding or pleasantness. All semantic dimensions resulted in higher recall than did the phonemic dimension. Specific semantic encoding does not necessarily result in higher recall than does a general semantic encoding, as evidenced by the equivalence of recall for the pleasantness and understanding conditions.

The congruous and incongruous conditions display different patterns of recall for items from intermediate and dissimilar pairs. The similarity structure is clearly reflected in the pattern of recall when the rating task is congruous with the dimension of similarity. Unlike the pattern for the congruous condition, recall for the incongruous condition does not correspond with the pattern suggested by the difference ratings. The more similar pairs, as indicated by smaller difference ratings, are not recalled more frequently than are less similar pairs. Rather, items from the extremes of similarity in size are recalled more frequently. Although ratings reflect the similarity structure of the list, the effect of similarity is not reflected in recall when the dimension of comparison focuses encoding away from the similarity structure. It appears that specificity of encoding along a dimension is salient, resulting in a distinctive trace, only when the encoding task focuses on the appropriate structure.

Similarity in Pleasantness

Rating. Table 16 presents mean rating of pair difference as a function of similarity in pleasantness for each dimension of comparison. The pleasantness rating task results in a pattern of rating distinct from that of any other condition involving the pleasantness list. The pleasantness condition shows a steady rise in difference rating with a decrease in similarity. When ratings are based on the dimension of size or ease of understanding, the extremes of pair similarity are rated as being more different than are items intermediate in similarity. Such an effect was also found for size ratings of the pleasantness list in Experiment 3. Ratings for the pronunciation condition suggest that pairs are perceived as being less different as they decrease in similarity, although such an effect is slight.

The pattern of ratings for the size and understanding conditions suggests that similarity in pleasantness has an effect on perceived differences between pair members on another semantic dimension. However, the effect does not mirror the underlying list structure, in contrast to the results with the size list.

A 4x3 mixed design analysis of variance was performed on the rating data. All factors investigated had significant effects on rating: rating dimension [$F(3,76)=8.86$, MSe=388), similarity in pleasantness

Table 16. Mean rating of pair difference as a function of decile separation in pleasantness for each of the dimensions of comparison involving the pleasantness list in Experiment 4. Standard deviations are in parentheses.

	DECILE SEPARATION IN SIZE		
Comparison Dimension	0	3	6
<u>Congruous</u>			
Pleasantness	5.69 (1.62)	5.87 (1.67)	6.66 (1.29)
<u>Incongruous: Specific</u>			
Size	6.58 (1.00)	5.83 (1.32)	6.84 (0.91)
<u>Incongruous: General</u>			
Understanding	4.68 (1.68)	4.31 (1.75)	4.85 (1.91)
Pronunciation	4.84 (1.41)	4.68 (1.44)	4.39 (1.61)

[$\underline{F}(2,152)=10.8$, $MSe=31.5$], and their interaction [$\underline{F}(6,152)=5.54$]. Each specific semantic condition resulted in a distinct level of perceived pair difference. Pairs were rated as being more different when rated for difference in size than when rated for difference in pleasantness. Both specific encoding conditions resulted in higher difference ratings than did the understanding and pronunciation conditions, which did not differ from each other.

Similarity in pleasantness had different effects on rating for each of the specific semantic conditions with the expected effect of rating reflecting similarity for the congruous condition. Size ratings display an end effect with pairs at the extremes of similarity being rated as more different than are the intermediate pairs.

Cued Recall. Table 17 presents mean number of words correctly recalled, along with the corresponding probability of correct recall, as a function of similarity in pleasantness for each of the comparison dimensions.

The pattern of recall for the congruous rating condition is different from the pattern for the other conditions. In the pleasantness condition, intermediate items are recalled more frequently than are items from either extreme of similarity. There is no clear relationship between specificity of encoding and subsequent recall for the congruous condition. (The pattern is consistent with the pattern found in free and cued recall data with semantic distractors in the pilot experiment in Appendix 3.)

The pattern of recall demonstrated by the remaining semantic conditions is a reversal of the pattern exhibited by the congruous condition. Items from the extremes of pleasantness similarity are

Table 17. Mean number of words correctly recalled (of a possible 8) as a function of decile separation in pleasantness (inversely related to similarity in pleasantness) for each of the dimensions of comparison involving the pleasantness list in Experiment 4. Standard deviations are in parentheses. Probability of correct recall for each value is indicated by an asterisk.

	DECILE SEPARATION IN SIZE		
Comparison Dimension	0	3	6
<u>Congruous</u>			
Pleasantness	1.35 (1.01) *.17	1.95 (1.94) *.24	1.45 (1.16) *.18
<u>Incongruous: Specific</u>			
Size	1.55 (1.02) *.19	1.30 (1.14) *.16	1.80 (1.25) *.23
<u>Incongruous: General</u>			
Understanding	0.90 (1.34) *.11	0.75 (1.37) *.09	0.95 (1.28) *.12
Pronunciation	0.15 (0.36) *.02	0.20 (0.51) *.03	0.30 (0.71) *.04

recalled more frequently than are intermediate items. In the incongruous condition, items from the dissimilar pairs are recalled slightly more frequently than are items from similar pairs.

A 4x3 mixed design analysis of variance was performed on the recall data. The results of the analysis suggest that rating dimension [$F(3,76)=8.56$, $MSe=2.95$] has a significant effect on recall, but that the interaction of dimension and similarity [$F(6,152)=1.60$, $MSe=0.68$] and pleasantness similarity itself ($F<1$) do not. Recall for each of the specific semantic dimensions is higher than for the understanding condition. Recall in the understanding condition is higher than in the pronunciation condition, a difference supporting the levels-of-processing hypothesis.

Any specific semantic rating of the pleasantness list produces a higher level of recall than does a general semantic rating. The specificity of the trace resulting from a specific encoding is greater than for a general encoding. Specific aspects of meaning are utilized to make the initial comparisons and the resulting traces more detailed and precise, hence more distinctive.

The patterns in recall for the specific conditions show an effect of congruity. The pattern for the incongruous condition closely resembles the pattern for the incongruous encoding of the size list, even though the rating patterns do not so correspond.

The pattern of recall in the congruous condition using the pleasantness list does not resemble the pattern of recall for the corresponding condition with the size list. The memory trace for a pleasantness encoding may not be developed in the manner assumed for purposes of predictions. Pleasantness may not form a salient or retrievable aspect of the memory trace except when items of extreme pleasantness are encoded.

Pollio (1968) presents evidence in support of an end anchor effect in verbal tasks. When word sets that comprise a sequential structure (e.g., freshman, sophomore, junior, senior) are randomly presented in a verbal learning task, the structure of the word set is reflected in the resulting learning curve. That is, the typical serial position curve results when performance is plotted as a function of the sequential structure. Words at each extreme of the stucture are learned most quickly. As the position of a word in the structure approaches the middle of the sequence, the number of trials needed to learn the appropriate response increases. Similarly, the more extreme a word is along a relevant semantic dimension, the easier the word is to learn and the more likely it is to be recalled. The data for the pleasantness dimension fit a pattern predicted for an end anchor effect. Items from the extremes of the pleasantness ratings are recalled more frequently than are neutral items. The list structure, emphasized by the comparison task, may have influenced recall to an extent which could not be overridden by the potential effects of specificity of encoding.

If the relative pleasantness of a pair is less substantial or less permanent than are aspects of a more objective nature such as referent size, the presentation of a recall cue may not provide ready access to the pair trace; pleasantness information may not be accessible in most cases without presentation of both members of a pair. Pleasantness judgments may even divert attention from those aspects of meaning that are especially useful to the creation of a specific, distinctive memory trace for a stimulus pair.

The contrasts between the dimensions of size and pleasantness are emphasized by the effects that similarity along each dimension has on difference ratings and patterns of recall for congruous and incongruous conditions. Similarity along the specific, objective dimension of size is directly reflected in difference ratings for the congruous condition, and less strongly reflected for the incongruous conditions. Patterns of recall parallel perceived and actual differences in similarity when the size list is encoded along the congruous dimension, as in Experiments 1 and 3. Patterns of recall are distinct when the dimension of encoding is semantic but incongruous; items from the extremes of similarity in size are recalled more frequently than are items of intermediate similarity. Similarity along the specific, but subjective, dimension of pleasantness is reflected in difference ratings for the congruous condition, but is not reflected in ratings for incongruous conditions; similarity in pleasantness does not appear to be a salient quality when encoding is directed away from it. Patterns of recall are distinct for congruous and incongruous conditions; when pleasantness is the relevant dimension, items from pairs of intermediate similarity are recalled more frequently than are items from the extremes of similarity. The pattern of recall is reversed in the incongruous conditions, and resembles the pattern found with incongruous tasks for the size list.

An attempt should be made to explain the pattern of retention results generally found in the incongruous conditions. The results for cued recall in the present experiment suggest that list structure is affecting the pattern of retention in incongruous conditions. The pattern is similar to the end-anchor effect (Pollio, 1968) mentioned in the context of the pleasantness comparisons.

Items from the extremes of the dimensional structure represented in a stimulus list are remembered more frequently than are items from the middle of the dimension. The structure of the presentation list creates a range of pair difference rather than of individual words along the structural dimension. Although incongruous decisions do not require explicit use of the list structure, contrasts along the structural dimension could provide salient distinction among pairs that are relatively indistinct on the dimension of comparison. The salient or distinctive structure provided by the irrelevant dimension of the list structure is unrelated to specificity of encoding. The extremes of this irrelevant structure are more salient than are intermediate pairs, as the extremes provide the organization of the structure, establishing the range of comparison under consideration.

While congruity of encoding affected the pattern of recall, specificity of encoding affected the level of recall. For either list one or both of the specific encoding conditions resulted in higher levels of recall than did the general semantic (understanding) condition. When encoding was made even more general, and nonsemantic, by focusing on phonemic aspects of the stimulus, recall was still lower.

Clearly the situation represented by congruity and specificity of encoding may be more complex than had been indicated by previous studies of congruity. In most such studies, incongruity of encoding and stimulus had a devastating effect on the level of recall. Granted, incongruity in the present context is a somewhat different case;

incongruous in the present context means inappropriate to the structure of the list, but incongruity does not result in an encoding remaining anomalous. This experiment has not enabled a comparison of congruity and incongruity for general as well as specific dimensions of encoding. That question is addressed in the next experiment.

CHAPTER 6

EXPERIMENT 5

The experiments presented thus far demonstrate that beyond a certain level of specificity, greater specificity of encoding, as measured by pair similarity and by decision difficulty, results in more distinctive memory traces only when the encoding dimension is congruous with the list structure. When dimension and structure are incongruous, aspects of the stimulus structure (e.g., extreme values of similarity) other than specificity of encoding have demonstrable effects on the pattern of recall. Congruity generally did not affect the level of recall, but rather the pattern of recall, with respect to specificity of encoding.

The experiments demonstrate the manner in which incongruity between encoding and structure may alter the potential effects of specificity of encoding on the distinctiveness of the memory trace. The demonstration is limited by the fact that subjects used a single encoding dimension. Although it was possible to determine the effects of congruity and incongruity along a specific dimension, and the range of specificity within that dimension, it was not reasonable to compare the effects of congruity and incongruity along a more general dimension.

The fifth experiment extended the investigation by varying the congruity and specificity of encoding for a particular stimulus pair. The encoding task called for subjects to make category judgments, with or without specific descriptive information, about presented pairs.

Four stimulus lists were constructed so that each stimulus pair occurred with equal frequency in the context of general or specific, congruous or incongruous category questions. Consider these examples. A specific congruous encoding question would be "Are these flowers used in corsages?" for the stimulus pair "orchid-carnation." A general congruous encoding question would be "Are these sciences?" for the stimulus pair "chemistry-geology." A specific incongruous encoding question for the pair "train-streetcar" would be "Are these chemical elements that glow?." "Are these footwear?" would be a general incongruous encoding question for the pair "potato-onion."

Subjects were presented with encoding questions, each followed by a word pair. They decided whether the encoding question was an appropriate description of the pair that followed. After the categorization, they were given a cued recall test. The cues were members of pairs presented in the stimulus list.

The different encoding contexts allow for a determination of the way in which congruity and specificity of encoding affect the cohesion of the memory trace. The members of a pair are always members of the same category, and share a specific characteristic. Therefore, presentation of one member as a cue for the recall of the other provides a strong indication of the nature of the item being sought. The search for the other member of a pair should be restricted to a small subset of items consistent with the nature of the cue.

It is expected that congruity will affect subsequent recall. Congruous encoding in the context of a particular encoding question for each stimulus requires a more specific analysis than does an incongruous encoding (Craik & Tulving, 1975; Hall & Geis, 1980; Mathews, 1977). An examination of the examples provided earlier may convince the reader that such is the case. The greater specificity of encoding should result in a more distinctive and cohesive trace. Greater specificity in the encoding task should also enhance retention, but only when the context is such as to make the specificity of some contrastive value. In an incongruous context, greater detail does not distinguish the encoded pair from other pairs which also are incongruous with respect to a specific encoding question. From the examples given earlier, the question "Are these chemical elements that glow?" applies to "train-streetcar" no better than to "fathom-knot" or "bee-wasp" or any number of other related pairs. Therefore, specificity in encoding should serve to increase distinctiveness and enhance retention only when the word pair is encoded in a congruous context.

Method

Subjects. Ten subjects participated in each of four experimental conditions, one with each of four presentation lists. In addition, ten subjects took part in a control (generate) condition. The 50 subjects were introductory psychology students who participated in the experiment as a course requirement. Subjects were tested in group sessions, with each group randomly assigned to a condition.

Materials

Presentation Lists. Four presentation lists were constructed. The use of four lists allowed each critical pair of items to be represented in all combinations of congruity (congruous-incongruous) and specificity (specific-general).

Forty critical pairs of items were chosen to represent 40 different categories from the Battig and Montague (1969) category norms. (See Appendix 5 for a listing of the 40 pairs and how they are described in a specific congruous encoding question.) For each list, one-quarter of the pairs were presented in the context of a specific congruous question, one-quarter in that of a general congruous, one-quarter in that of a specific incongruous, and one-quarter in the context of a general incongruous question. Across lists, specificity and congruity were varied so that each critical pair was presented once in each of the four contexts. Pairs that were specific in List 1 were general in List 2 (and vice versa). Pairs that were in a congruous context in List 1 were in an incongruous context in List 3. Both specificity and congruity were reversed from List 1 to List 4.

In addition to the critical pairs, 40 filler pairs were selected. Ten of these pairs were chosen from the Battig and Montague (1969) norms. (The remaining six categories in the norms were considered inappropriate for inclusion, either because of extensive overlap with a critical category or because the categories seemed inappropriate to the task requirements.) The additional 30 filler categories and pairs were constructed by the experimenter.

The lists were arranged so that the first 12 and last 12 encoding items were filler items. The interior 56 stimulus positions were assigned randomly to critical or filler pairs. The assignment of filler pairs to positions was constant across lists. The critical pairs were assigned to positions across lists so that if a given position represented a specific congruous pair in List 1, it would represent each of the other three contexts in one of the remaining three lists. The assignment of a critical pair to position was random within the constraints of balancing context across lists. The balancing of position with type of context was not accomplished by keeping the same critical pair in a position and changing the encoding question.

For each presentation list, the 80 questions and pairs were recorded on videotape from a computer (Apple II) display. An encoding question was presented for 7.5 seconds, then immediately followed by a word pair for 2.5 seconds. A 2.5 second inter-trial interval followed the presentation of each word pair.

Test Lists. The cued recall test given to subjects following each presentation list consisted of 40 cues. The cues were the first member of each of the critical pairs. The cues were arranged on the test sheet in random order.

Procedure. Subjects in the experimental conditions were shown one of four lists. They were told that they were participating in an experiment examining categorization of word pairs. For 80 pairs of words, they would see a question followed by a pair of words. Their task was to answer the question as it applied to the pair of words that followed it. Two examples, one requiring a positive and one requiring a negative response, were given. Subjects were given response sheets on which to record their answers.

After subjects completed the categorization task, they were given recall sheets and asked to write down for each of the cues on the sheet the word that had been paired with the cue during the categorization task. Subjects were given seven minutes to complete the recall task.

Subjects in the generate condition were given the cued recall sheet and asked to write down for each word another word that was related to the first. The generate task was self-paced.

Results and Discussion

The .05 α level was used for all tests of significance. Scheffé comparisons were used to confirm those differences accounting for significant effects.

The probability of correctly generating a response without prior exposure to a presentation list was .008. The ten subjects generated only three correct responses. Therefore, the discussion considers only the data from the experimental conditions.

Table 18 presents the mean number of words correctly recalled as a function of congruity and specificity of encoding, along with the corresponding standard deviation and probability of correct recall for each value. As there were no differences in performance across lists, the data have been collapsed across lists.

The effects of congruity and specificity are apparent. Congruity of encoding had a strong impact on recall of targets. A 2x2 repeated measures analysis of variance confirms that the factors of congruity [$F(1,117)=352$, $MSe=1.71$], specificity [$F(1,117)=14.1$] and the interaction of congruity and specificity [$F(1,117)=15.0$] all significantly affected correct recall of target items. Items that had been encoded in a specific context were recalled much more frequently than items from a general context, but only when the items had appeared in a specific congruous context.

Table 18. Mean number of words correctly recalled (of 10 possible) as a function of the encoding context in Experiment 5. Standard deviations are in parentheses. Probability of correct recall is also given for each value.

Encoding Context	$\bar{X}$ (σ)	p.
congruous specific	7.65 (1.58)	.77
congruous general	6.08 (1.62)	.61
incongruous specific	2.98 (1.66)	.30
incongruous general	3.00 (1.99)	.30

The use of one member of the encoded pair as a cue for the other provides a measure of the cohesion of the memory trace for the pair. The cohesion of the trace is influenced strongly by the specificity of the encoding context, but only when that context is consistent with the meaning of the word encoded. When the context is inappropriate to the meaning, further specificity does not influence the precision of the encoding required to make the decision called for in the categorization task. It takes no more detailed processing to determine that "ear" and "toe" are not "<u>blossoming</u> trees" than to determine that they are not "trees."

Subjects averaged 3.2 intrusions in recall. Extralist intrusions accounted for 93% of all errors. The intrusions represented items with a high degree of similarity to the intended targets. Again, the specificity of encoding appeared to have an effect when encoding was congruous. When encoding had been specific, intrusions closely resembled targets in specific aspects encoded as well as category membership; when encoding had been general in nature a somewhat higher proportion of intrusions shared category membership but not specific attributes of meaning with the intended targets.

The results are exactly as expected. Further specificity in the encoding task enhances recall only when the task calls for a congruous processing of the stimuli. An incongruous encoding allows the decision task to be completed with a more superficial encoding of the meaning of the stimuli. Such a cursory encoding results in relatively low cohesion between the elements of the trace.

The situation is very different when the encoding context is congruous with the stimuli being examined. To make a decision regarding distinction as well as category membership requires that the subject encode the word pair more fully and specifically. It is not sufficient to encode "ear" and "toe" solely as "parts of the body" if you are also being asked if they "can be wiggled." The further specificity of encoding called for by the question results in a more detailed and precise description, or trace, that is more cohesive than is the trace resulting from a more general encoding question.

It has been established that in a congruous context, specificity of encoding affects the cohesion of the memory trace. A reasonable next step is to examine the influence of specificity of encoding on the likelihood of making contact with the memory trace and to determine the impact that specificity at the time of test has in conjunction with specificity of encoding on the likelihood of recall and cohesion of the trace. The next experiment was conducted to address those issues.

CHAPTER 7

EXPERIMENT 6

The final experiment was conducted to examine the nature of the effect of specificity of encoding and of test context on the memory trace. The experiment resembled the fifth experiment in the encoding phase. However, instead of receiving one member of the word pair as a cue for recall of the other, subjects received cues of different levels of specificity. Different measures of recall, as well as of intrusions in recall, provide an estimate of access to the memory trace and the discriminability and cohesion of the trace.

Subjects made category decisions about pairs of words. The encoding question for a particular pair was either specific or general, and congruous or incongruous. The presentation lists were those used in Experiment 5. After encoding, subjects received recall cues that varied, between subjects, in specificity. Half of the subjects received cues which consisted of the full descriptive part of a specific encoding question (e.g., "sports that use poles"). The other half of the subjects received cues that represented appropriate general category information (e.g., "sports").

The combination of degree of specificity at time of encoding and at time of test permits a determination of the relative impact that specificity has on the precision of the trace, during formation, and that either type of specificity has on the accessibility of the trace. In addition, it is possible to determine whether either type of specificity affects the cohesion of the pair trace.

An auxiliary experiment was conducted at the same time to determine whether the category and descriptive information present in the specific cue have equivalent effects, qualitatively and quantitatively, on the accessibility and discriminability of the memory trace. To address the issue, additional groups of subjects were given as cues for recall the descriptive information from specific encoding questions (these cues hereafter are referred to as descriptor cues). Such cues were constructed by deleting the category label from the specific cues (e.g., instead of the cue "chemical elements that glow in the dark," subjects receiving descriptor cues were given the cue "glow in the dark").

Clearly it is expected that cases in which encoding and the recall cue are specific should result in the highest level of recall. The specific encoding induces subjects to encode the word pair more specifically, resulting in a more distinctive memory trace, which should be readily discriminable from alternative traces. Such a trace should be quite cohesive, as the contrasts developed during encoding are highly specific to that trace. In addition to the benefits of such specific encoding, the specific context at time of recall should provide the best possible conditions for locating the intended trace. The specific context at time of test aids in two ways. First, the context provides a specific clue to the contents of the trace, making confusion with alternative traces unlikely. Second, the context at time of test is highly similar to that at time of encoding, in a sense recreating the encoding context. Thus the accessibility of the trace is not hampered by a change of context from encoding to test.

The relative levels of recall for the remaining combinations of context at encoding and test are less predictable. The distinctive trace fostered by specific encoding should be located with somewhat more difficulty when the general cue is used, so the recall level would be lowered relative to the condition in which both contexts are specific. However, it is likely that the general cue will be sufficient to provide contact with the intended trace to a considerable extent.

The nature of the memory trace is in no way altered by the test context. The distinctiveness of the memory trace may be masked by differences in context between encoding and test because the test context may not recreate the encoding context to an extent necessary to indicate the memory trace being sought. Once the appropriate memory trace is located, both members of the encoding should be recalled to the same extent regardless of the test context. The cohesion of the trace should be unaffected by the change of context at the time of test. The specificity of encoding, but not of test context, affects the cohesion of the memory trace. The specific encoding context, by requiring a specific comparison of the stimulus pair being encoded, results in a more cohesive trace, regardless of the context at the time of test.

It is predicted that level of recall and the cohesion of the memory trace will be lower following general encoding, regardless of the context at time of test. It is not possible *a priori* to determine whether the similar context or the specificity of the cue will have a stronger effect on the likelihood of making contact with the trace. Again, specificity of the cue should have no impact on the cohesion of a trace, once it has been discriminated from alternative traces. The opposing effects of specific context and same context at time of test should be explained.

Tulving and Thomson (1973) demonstrated that a change in context from time of encoding to time of test seriously impairs recall. Any change in context that encourages alternative aspects of meaning not stressed during encoding hampers retrieval. In the present experiment, the specific cue may make explicit a similarity between pair members not considered by the subject during a general encoding. If such is the case, subjects may not recognize the intended trace when they have located it. The change in context may or may not be offset by the greater specificity of the cue, which more precisely indicates the nature of the items being sought.

There is, however, precedent for a specific cue proving a more successful cue for recall than a cue which is explicitly part of the encoding. Anderson, Pichert, Goetz, Schallert, Stevens, and Trollip (1976) demonstrated that an implicitly specific context at time of encoding might encourage a more precise identification of the subject noun than the general encoding explicit in the context. Such a precise identification is known as instantiation. Particular terms under some circumstances are better cues for recall than are the presented general terms. For example, subjects presented with the sentence "The woman worked in the theatre" may recall the sentence more frequently when given the word "actress" as a cue than when provided with the cue "woman," the word actually used in the sentence.

While it is not possible to predict whether the effect of specificity or similarity of test context will have equivalent, though opposing, effects on recall, it is clear that the general encoding context results in less specific memory traces. Those traces should be more difficult to access, more difficult to discriminate, and less cohesive once isolated from alternative traces. The general encoding should result in a higher relative level of intrusions in recall. These intrusions may resemble intended targets only to the extent of sharing category membership, but not specific characteristics, when the test cue is general. Subjects should restrict response candidates to those consistent with the descriptive information available to them, from both encoding and test contexts.

For the auxiliary experiment, it is predicted that recall following general encoding may be particularly low when the descriptor cues are provided. The descriptor cues provide relational information not made explicit during the encoding. A more specific instantiation of the pairs is possible during general encoding, but the more specific encoding is not a necessary aspect of the task. If a trace is located within the context of the descriptor cue, it should be relatively cohesive. Little confusion with possible alternatives should occur subsequent to locating the intended trace.

Method

Subjects. Ten subjects participated in each of eight experimental conditions, two with each of four presentation lists. In addition, ten subjects provided control data for each of the cued recall tests used. An additional 50 subjects provided data for the auxiliary experiment, ten with each of the four presentation lists and an additional ten as controls. The 150 subjects were introductory psychology students who participated in the experiment as a course requirement. Subjects were tested in group sessions, with each group randomly assigned to a condition.

Materials. Four presentation lists and two cued recall tests, along with the descriptor cued recall test, were used.

Presentation Lists. The four presentation lists were those used in Experiment 5.

Cued Recall Tests. The general recall test list consisted of the category labels (e.g., "vegetables') appropriate to the 40 critical pairs represented in the stimulus lists. The specific recall test list consisted of the full description of a specific encoding question (e.g., "vegetables that grow underground"). The descriptor recall test list consisted of the specific cues with the deletion of the category label (e.g., "grow underground"). The cues for each list were positioned on the test sheet in random order.

Procedure. Subjects were instructed as in Experiment 5 prior to being shown a presentation list. After they completed the encoding task, they were given one of the recall tests. They were instructed that each of the 40 descriptions on the sheet was appropriate to one of the pairs they had categorized. Their task was to write down as many of those words and pairs of words as possible next to the appropriate descriptions. Subjects were given seven minutes to complete a recall test.

Subjects in the control (generate) conditions were given the appropriate recall sheet and asked to write down for each description a pair of words which fit the description. The generate tasks were self-paced.

Results and Discussion

The .05 α level was used for all tests of significance. Scheffé comparisons were used to confirm those differences accounting for significant effects.

Table 19 presents mean number of words correctly recalled as a function of specificity of encoding and specificity of cue along with the corresponding standard deviations. Probability of correct recall is provided for each value. The data have been collapsed across the four presentation lists.

The data presented for recall of items encoded in an incongruous context represent items that were produced in response to the description appropriate to, but not presented with, a stimulus pair. That is, the pair "deer-sheep" would be scored as correct when produced in response to the cue "animals that have hooves" even though the pair had been encoded in the context of "periodical reading material." (Only three subjects produced four correct responses when the data were scored according to the presented encoding context ($p = .002$).) Notice that the standard deviations are extremely high, indicating extreme variability in performance. An analysis of variance performed on the recall data for items encoded incongruously reveals no signficant factors. Neither specificity of the cue ($F<1$), of the encoding [$F(1,78)=1.34$, $MSe=3.93$], nor the interaction of encoding and cue specificity ($F<1$) affected recall significantly. Experiment 5 demonstrated that specificity of encoding had no effect on the recall of incongruously encoded stimuli. The present experiment confirms that finding and also demonstrates that the specificity of the recall context has no effect on recall of such incongruously encoded items.

Table 19. Mean number of words correctly recalled as a function of specificity of encoding and specificity of recall cue for each experimental condition in Experiment 6.

Specificity of Encoding/ Specificity of Test Context	$\bar{X}$ (σ)	p.
	Congruous Encoding	
Specific / Specific	16.53 (2.75)	.83
Specific / General	13.38 (3.62)	.67
General / Specific	11.50 (3.98)	.58
General / General	10.83 (3.04)	.54
Specific / Descriptor	14.28 (4.11)	.71
General / Descriptor	5.45 (3.03)	.27
	Incongruous Encoding	
Specific / Specific	4.88 (13.5)	.24
Specific / General	5.28 (9.35)	.26
General / Specific	4.48 (12.1)	.22
General / General	4.95 (13.2)	.25
Specific / Descriptor	3.30 (2.80)	.17
General / Descriptor	3.70 (3.29)	.19

Examination of the recall data for congruous encoding suggests that specificity affects the likelihood of making contact with a memory trace. The specificity of the recall cues seems to affect the likelihood of recall only when encoding has been specific.

A 2x2 mixed design anlaysis of variance confirms that specificity of cue [$F(1,78)=9.76$, MSe=15.6], specificity of encoding [$F(1,78)=94.2$, MSe=7.89] and the interaction of specificity at time of encoding and test [$F(1,78)=7.77$] all had significant effects on the probability of correct recall. Specificity of encoding enhanced recall regardless of the type of cue. Specificity of the recall cue enhanced recall only when the encoding task was specific.

Specificity of encoding and of test context are the best possible combination for making contact with the memory trace. The context at time of test is consistent with the context at encoding, in accord with the concept of encoding specificity. Because the initial encoding context was specific, the resulting memory trace is more precise and more readily accessible. The specific context at test also enhances the likelihood of contacting the trace, by making it more readily discriminable from other memory traces.

Once specificity is lessened at either time of encoding or of test, the accessibility of the memory trace is depressed. Specificity of encoding is the crucial factor, making the discriminability of a memory trace higher.

It is reasonable to assume that the equivalence of recall for cuing conditions following general encoding stems from two factors acting in opposition. When the cue is general, an encoding specificity effect may be operating to benefit recall; the cue re-creates the context at encoding. However, recall may be depressed because the general cue does not offer the potential benefit of greater specificity to make it easier to discriminate the trace. When the cue is specific, the benefit of a specific context (instantiation) at time of test exists, but the cue does not recreate the encoding context.

Table 20 presents the average number of intrusions as well as the proportion of total recall output represented by intrusions, as a function of encoding and test contexts for items that had been encoded congruously. An analysis of variance on the total number of intrusions revealed that specificity of the cue significantly affected the number of intrusions produced in recall [$F(1,78)=4.00$, MSe=1.70]. Specificity of encoding [$F(1,78)=1.46$, MSe=1.23] and its interaction with type of cue [$F(1,78)=1.48$] had no effect on the number of intrusions produced. However, subjects produced a higher percentage of intrusions to total output when both encoding and cue were general.

Intrusions resembled intended targets in terms of category and intended specific characteristics for 61% of intrusions produced when both encoding and cue were specific. A decrease in specificity at time of encoding or time of test resulted in equivalent drops in similarity (48% vs. 50%). When both encoding and cue were general, only 26% of intrusions were highly similar to intended targets, although virtually all intrusions did come from the same categories as intended targets.

The analysis of total recall data provides an estimate of the likelihood of gaining access to a particular memory trace for the different encoding and cue conditions. An analysis of the likelihood of producing the second member of the pair given the first member has been produced provides an estimate of the cohesion of the memory trace that

Table 20. Mean number of intrusions produced in recall as a function of specificity of encoding and specificity of cue for conditions in Experiment 6. The proportion of total recall output that is represented by intrusions is also provided for each experimental condition.

Specificity of Encoding/ Specificity of Cue	Mean number of Intrusions (s.d.)	Proportion of Recall Output
Specific / Specific	0.78 (0.89)	.04
Specific / General	0.98 (1.38)	.07
General / Specific	0.78 (1.00)	.06
General / General	1.40 (1.48)	.11
Specific / Descriptor	0.60 (0.84)	.04
General / Descriptor	0.68 (1.16)	.11

has been accessed. Cohesion of the trace should not be affected by the specificity of the recall cue; rather, it should be affected solely by the context at encoding. Such would clearly be the case here; when encoding was specific, *P* (Recall of second member|Recall of first) was .77 when the cue was specific and .74 when the cue was general. These values are highly similar to the recall probability following specific congruous encoding in Experiment 5 (p =.77). When encoding was general *P*(Recall of Second member|Recall of first) was .64 when the cue was specific and .58 when the cue was general. Again, the values are highly similar to the corresponding condition in Experiment 5 (p =.61). A 2x2 mixed design analysis of variance confirmed that specificity of cue [$F(1,78)=1.25$, MSe=.056] had no significant effect on the cohesion of the trace, while specificity of encoding [$F(1,78)=21.3$, MSe=.037] significantly affected the cohesion of the pair trace. The interaction of the two factors had no effect on the cohesion of the trace ($F<1$).

Specificity of encoding appears to affect more than one aspect of retention. The more specific trace developed during the specific encoding task is more readily discriminable, more accessible, and more cohesive. Specificity at the time of test enhances access to the memory trace, and may make the trace more discriminable from other traces by more fully specifying the required context, but it has no influence on the cohesion of that accessible memory trace.

The data from the auxiliary experiment considered alone and in combination with the data from the main experiment provide additional information regarding specificity of encoding and the completeness of the recall cue. Table 19 presents the mean number of words correctly recalled as a function of specificity of encoding with the descriptor cues as cues for recall. Specificity of encoding has a dramatic effect on the likelihood of making contact with the memory trace. When encoding was general, the descriptor cue rarely made contact with the memory trace.

A 1x2 repeated measures analysis of variance confirms that specificity of encoding had a significant effect on correct recall [$\underline{F}(1,39)=93.9$, MSe=12.3]. It is hardly surprising that the descriptor cue made contact with the trace rarely following general encoding. The information contained in the descriptor cue is precisely that information that explicitly differentiated the specific and general encodings. The low likelihood of contacting the general trace with the descriptor cue seems to indicate that instantiation is not a major contributing factor in the encoding of this general task, or at least that such instantiation does not make use of the specific attributes used by the experimenter. Had subjects encoded the general pairs more specifically, they should have found it easy to gain access to the trace when provided with the descriptor cue at time of recall.

A 3x2 mixed design analysis of variance was performed to compare recall using a descriptor cue with that of recall using a specific or general cue. The descriptor cue does not enhance recall over the general cue when encoding was specific. When encoding was general, the descriptor cue resulted in lower recall than did the general cue. The analysis confirmed these observations. The effects of specificity of cue [$\underline{F}(2,117)=37.4$, MSe=14.6] and of specificity of encoding [$\underline{F}(1,117)=194.5$, MSe=9.37] were strong in their influence on the likelihood of making contact with the memory trace. The interaction of specificity of cue and encoding had no effect on recall ($\underline{F}<1$). Recall was equivalent for subjects using descriptor and general cues following specific encoding, significantly lower than for subjects using specific cues. Following a general encoding, recall was significantly lower when descriptor cues rather than general cues were used.

The equivalence of general and descriptor cues in conjunction with specific encoding suggests that the loss of either category or descriptive information impairs recall. The context at recall no longer recreates the context at encoding, making the traces less accessible or discriminable. Contact with the trace is seriously impaired following general encoding when the context at recall presents an implicit rather than explicit relationship between the encoding context and the stimulus pair being considered.

Intralist intrusions were much more frequently produced in the context of descriptor cues than were such intrusions with either other type of cue. No analysis of such intrusions was performed after it ws determined that many such intrusions were consistent with the descriptor cues, even though the cues more closely fit the intended targets. The descriptor cues appear to reduce the discriminability of memory traces, thus affecting the likelihood of making contact with the intended trace.

The descriptor cues resemble general cues after specific encoding in terms of the effect on access to the memory trace. However, the descriptor cues do result in comparable measures of trace cohesion to those found for specific and general cues. When encoding was specific, $\underline{P}$(Recall of second member | Recall of first) = .82 (compare to .77 and .74). When encoding was general, the comparable measure of trace cohesion was .73 (compare to .64 and .58). Discriminability of the trace and access to the trace are impaired by the descriptor cue relative to the specific cue, but once the trace has been located, both members of a stimulus pair are usually produced. The higher estimate of pair cohesion from the descriptor cues probably reflects a recall situation in which only the more cohesive memory traces are being contacted by the descriptor cue, a situation which leads to an inflated estimate of cohesion.

The analysis of variance performed on the likelihood of both members being represented in recall given one had been confirmed the significance of type of cue [$F(2,117)=4.11$, MSe=.044], and of specificity of encoding [$F((1,117)=25.8$, MSe=.036], but no effect of the interaction ($F<1$). The greater cohesion of trace may be partially a result of the type of cue *per se*. In the *generate* conditions, the likelihood of producing both members given one is .21 for the specific cues and .22 for the descriptor cues, but only .02 for the general cues. The category information seems to lead to a greater range of items being represented in *generation* or recall. The relationship between pair members is less specific, resulting in a less cohesive memory trace, or a greater diversity of items generated when the appropriate category has been determined. (Naturally, the specific cues resulted in greater overall accuracy ($p.=.38$) at generating intended targets than did descriptor ($p.=.14$) or general ($p.=.10$) cues in the *generate* conditions).

The descriptor cues do not enhance access to the memory trace relative to the general cue. In fact, when encoding has been general, the descriptor cue greatly impairs recall by changing the context to one not explicit during encoding. The discriminability of the trace has been reduced, but the cohesion of the trace has not been depressed by the change in context.

The results as a whole indicate that specificity at the time of encoding is critical to the formation of a distinctive memory trace that is readily accessible, discriminable, and cohesive. Reducing the completeness of the cue to general or descriptive information impairs the likelihood of contacting and discriminating the appropriate trace, but does not affect the cohesion of that trace.

CHAPTER 8

SUMMARY, CONCLUSIONS, AND GENERAL DISCUSSION

Specificity of encoding has a striking impact on human memory. The present investigation has established some critical limits to the effect of specificity of encoding on the distinctiveness and retrievability of the memory trace. Those findings are summarized here.

Experiment 1 established that the specificity of encoding required by the encoding task should center on a semantic rather than a structural analysis of the stimulus. Greater specificity in encoding structural aspects of the stimulus does not usually result in distinctive information useful in discriminating stimuli from alternatives. When the specificity of encoding concerns referential aspects of the stimulus, the memory trace resulting from the encoding is recognized, recalled, and cohesive to an extent corresponding to the relative specificity of encoding. Experiment 1 also demonstrated that the memory effects attributed to specificity of encoding are more accurately predicted from the specificity of the memory trace than from differential degrees of cognitive effort.

Experiment 2 extended the generality of specificity of encoding with a consideration of synonymy. A relative increase in specificity of encoding is effective in enhancing memory even when the general level of specificity, as determined by pair similarity, is quite high. However, there are indications that specificity of encoding under such unusual circumstances as comparing synonyms for specificity of reference may not truly reflect the distinctiveness of the memory trace. Encoding may be specific without being complete; if such is the case, the resultant memory trace may be relatively inaccessible, but quite cohesive.

Experiments 3 and 4 examined specificity of encoding in the context of another critical factor, congruity of encoding. The experiments showed that specificity of encoding within an encoding task has a demonstrable effect on the retrievability of the stimulus only when the specificity of encoding is congruous with the stimulus structure. When specificity is incongruous with the stimulus structure, any potential effect of specificity is undermined by salient aspects of the encoding context or the underlying stimulus structure. However, the overall specificity of encoding required by the encoding task may affect recall independently of congruity. An incongruous but specific encoding task may result in a higher level of memory performance than does a more general task. Congruity of encoding necessarily results in a higher level of memory performance ONLY if incongruity results in anomaly, with the encoding context presenting a meaning foreign to that of the stimulus being encoded. The results taken together support the idea that the specificity of encoding necessary to fulfill the purpose of an encoding affects the retrievability of the memory trace.

Experiment 5 examined specificity of encoding in conjunction with a different type of incongruity. When an incongruous encoding context stresses meaning that does not apply to stimuli being encoded in that context, the incongruous context encourages only a superficial meaningful encoding of the stimuli. Stimuli in such a context are recalled with far less frequency than stimuli that are congruously encoded. Increased specificity of encoding in the incongruous context does not result in a more specific encoding of the meaning of the stimulus. The experiment confirms that specificity of encoding within a task is effective only when the encoding context is congruous with the stimulus.

Experiment 6 compared the effects of specificity at the time of encoding and the time of memory test. Specificity at time of encoding and test results in the best possible context for a memory search. The memory trace is distinctive, and the test context recreates the context of the encoding. The memory trace is accessible, discriminable, and cohesive. Any loss of specificity or completeness from the encoding context to the test context undermines the accessibility of the memory trace. The cohesion of the memory trace is determined by the specificity of encoding and is unaffected by a change in context at the time of retreival.

The findings of the thesis demonstrate that specificity of encoding is not the sole determinant in the formation of a distinctive memory trace. Encoding may be highly specific without being complete, resulting in a cohesive, but not accessible memory trace. Along with completeness, congruity of encoding is critical to specificity of encoding resulting in a distinctive memory trace. An incongruous encoding task may result in stimuli being encoded only superficially for meaning. Incongruity may allow salient aspects of the encoding context to predominate, overwhelming the potential effects of specificity of encoding. Any change in completeness or specificity of context from the encoding to the memory test may lower the accessibility of the memory trace, and thereby limit the demonstrated effect of specificity of encoding.

Although a relatively clear picture of the significance of specificity of encoding has emerged in the course of the thesis, the significance of congruity remains somewhat more elusive. The experiments presented have demonstrated situations in which incongruity interferes with the potential effects of specificity of encoding, with or without a corresponding decrease in overall memory performance. In contrast, Hall and Geis (1980, Exp. 1) demonstrated that when subjects are given sufficient time, specificity of encoding was reflected in patterns of recall when the encoding was incongruous. Hall and Geis manipulated specificity of encoding by varying the complexity of sentence frames. However, in such a case, the effect of increasing specificity may be to allow subjects to make a simple transformation to render the stimuli comprehensible (e.g., "hog-bed" in the sentence "I made my hog so that it would be neat"). The net effect of increased sentence complexity may be to nullify the incongruity by making the "true" referent of the sentence more apparent. Thus, if specificity of encoding can erase the anomaly or ambiguity of an incongruous encoding context, specificity of encoding will affect memory performance. Generally, however, the effect of incongruity is to encourage a relatively superficial encoding. This effect of incongruity suggests that the encoding process is quite efficient, making contrasts at whatever level of meaning and specificity is sufficient to fulfill the requirements of the task.

Some issues raised in the thesis have not been answered fully. The nature of contrasts developed during encoding was examined by the use of distractors in the first two experiments. The distractors contributed little additional information. A more careful matching of distractor and target might provide valuable information about the specific nature of encoding in the comparison tasks. Asking subjects for a report about the specific nature of items they fail to recall (as in investigations of the tip-of-the-tongue phenomenon, e.g., Brown & McNeill, 1966) might provide insight into the quality of information used in encoding, as influenced by the task.

It might be possible to discern the basic direction of the effect of specificity. It may be the case that the effect stems basically from more superficial or less extensive encoding within a task hampering the development of a distinctive trace relative to a "typical" meaningful encoding. Alternatively the primary action of the effect may come from a specific experimental encoding proving more distinctive (and memorable) than a typical encoding of meaning. Asking subjects to rate the difference between members of a pair on a particular dimension of comparison could be contrasted with simply asking them to study the word pairs for meaning. If the overall level of recall is higher following the "study" task, then the effect of distinctiveness may be attributed to the superficial encoding that pairs receive when they are readily distinguishable on the dimension of comparison.

Findings presented in the thesis underscore the significance that perceived task demands have on the nature of the encoding process. It may be possible for a task to require specific comparisons, so that encoding is relatively specific, without necessitating a comparably complete encoding. Consider Experiment 2, in which subjects determined the more specific referent of synonymous pairs. The task as presented did not seem to require an extensive set of contrasts between pair members to distinguish them from other members of the stimulus set.

Within the set, differential specificity resulted from the more specific and extensive processing required to complete the task for difficult pairs. However, subjects recalled relatively few of the stimuli presented when they were not given cues to the contents of a memory trace. These findings suggest that within the context of the encoding task, specificity of encoding is typically accompanied by relatively complete encoding. That level of completeness may be inadequate when the context at retrieval supposes a more complete or specific memory trace.

It might prove enlightening to investigate the issue of completeness of encoding. A comparison of tasks might demonstrate a great change in the specificity and completeness of encoding from a subtle change in the encoding context. Simply increasing the intralist similarity of some synonymous pairs might encourage a more complete encoding of those pairs. Determining the more specific referent would require the same degree of specificity demonstrated; the increased overlap among some stimulus pairs might influence subjects to contrast pair members more fully to distinguish them from similar stimuli. If the context did influence the encoding as outlined, then the memory trace should prove more accessible in a free recall test. A cued recall measure could determine that the increased output did not occur at the expense of the cohesion of the pair trace, from increased intralist confusion. Such a demonstration would provide compelling support for the idea that encoding is an efficient process, with stimulus analysis being carried out to the level of specificity and completeness necessary to fulfill the purpose of the encoding.

Aside from theoretical considerations, congruity and specificity may have practical implications for effective education. If teachers fail to establish the context in which a topic should be considered, then they may sacrifice the possible benefits of the most closely argued lecture. Their students do not have a grasp of how to regard the information being given them, and consequently the students often fail to connect the information in the most meaningful form. For example, if no mention is made of the time frame of Goya's artwork, the Spanish Civil War, students may remain puzzled that Goya seems obsessed with the ugly and bestial throughout much of his work. Comparisons with the court painter Velásquez, because of a shared nationality, are far less appropriate than are comparisons with the gloomy period of Van Gogh's work or Picasso's Guernica. Within the appropriate context the added detail and contrasts enhance the memory and understanding students have for a topic. If the context is somewhat inappropriate, students gain little from increased specificity, unless they can determine the appropriate context for themselves.

REFERENCES NOTE

1. Paivio, A. Unpublished list of normed pleasantness ratings for 260 nouns. Department of Psychology, University of Western Ontario, London, Ontario, Canada.

REFERENCES

Anderson, R.C., Pichert, J.W., Goetz, E.T., Schallert, D.L., Stevens, K.V., and Trollip, S.R. Instantiation of general terms. Journal of Verbal Learning and Verbal Behavior, 1976, 15, 667-679.

Arbuckle, T.Y. and Katz, W.A. Structure of memory traces following semantic and nonsemantic orientation tasks in incidental learning. Journal of Experimental Psychology: Human Learning and Memory, 1976, 2, 362-369.

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

Auble, P.M., Franks, J.J. and Soraci, S.A., Jr. Effort toward comprehension: Elaboration of "aha!"? Memory and Cognition, 1979, 7, 426-434.

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

Banks, W.P. and Flora, J. Semantic and perceptual processes in symbolic comparisons. Journal of Experimental Psychology: Human Perception and Performance, 1977, 3, 278-290.

Begg, I. Similarity and contrast in memory for relations. Memory and Cognition, 1978, 6, 509-517.

Bransford, J.D., Franks, J.J., Morris, C.D., and Stein, B.S. Some general constraints on learning and memory research. In L.S. Cermak and F.I.M. Craik (Eds.), Levels of Processing in Human Memory. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1979, 331-354.

Brown, R. and McNeill, D. The tip-of-the-tongue phenomenon. Journal of Verbal Learning and Verbal Behavior, 1966, 5, 325-337.

Coltheart, V. Recognition errors after incidental learning as a function of different levels of processing. Journal of Experimental Psychology: Human Learning and Memory, 1977, 3, 437-444.

Cooper, E.H. and Pantle, A.J. The total-time hypothesis in verbal learning. Psychological Bulletin, 1967, 68, 221-234.

Craik, F.I.M. and Jacoby, L.L. Elaboration and distinctiveness in episodic memory. In L.G. Nilsson (Ed.), Perspectives on memory research: Essays in honor of Uppsala University's 500th Anniversary. Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1979, 145-166.

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

Davies, G. and Cubbage, A. Attribute coding at different levels of processing. Quarterly Journal of Experimental Psychology, 1976, 28, 653-660.

Elias, C.S. and Perfetti, C.A. Encoding task and recognition memory: The importance of semantic encoding. Journal of Experimental Psychology, 1973, 99, 151-156.

Epstein, M.L., Phillips, W.D. and Johnson, S.J. Recall of related and unrelated word pairs as a function of processing level. Journal of Experimental Psychology: Human Learning and Memory, 1975, 104, 149-152.

Eysenck, M.W. Age differences in incidental learning. Devlopmental Psychology, 1974, 10, 936-941.

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

Eysenck, M.W. Levels of processing: A reply to Lockhart and Craik. British Journal of Psychology, 1978b, 69, 177-178.

Eysenck, M.W. Depth, elaboration, and distinctiveness. In L.S. Cermak and F.I.M. Craik (Eds.), Levels of Processing in Human Memory. Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1979, 89-118.

Eysenck, M.W. and Eysenck, M.C. Processing depth, elaboration of encoding, memory stores, and expended processing capacity. Journal of Experimental Psychology: Human Learning and Memory, 1979, 5, 472-484.

Eysenck, M.W. and Eysenck, M.C. Effects of processing depth, distinctiveness and word frequency on retention. British Journal of Psychology, 1980, 71, 263-274.

Fisher, R.P. Interaction between encoding distinctiveness and test conditions. Journal of Experimental Psychology: Human Learning and Memory, 1981, 7, 306-310.

Fisher, R.P. and Craik, F.I.M. Interaction between encoding and retrieval operations in cued recall. Journal of Experimental Psychology: Human Learning and Memory, 1977, 3, 701-711.

Frase, L.T. and Kammann, R. Effects of search criterion upon unanticipated free recall of categorically related words. Memory and Cognition, 1974, 2, 181-184.

Gardiner, J.M., Craik, F.I.M. and Bleasdale, F.A. Retrieval difficulty and subsequent recall. Memory and Cognition, 1973, 1, 213-216.

Goldman, S.R. and Pellegrino, J.W. Processing domain, encoding elaboration, and memory trace strength. Journal of Verbal Learning and Verbal Behavior, 1977, 16, 29-43.

Griggs, R.A. and Shea, S.L. Integrating verbal quantitative information in linear orderings. Memory and Cognition, 1977, 5, 287-291.

Hall, D.M. and Geis, M.F. Congruity and elaboration in free and cued recall. Journal of Experimental Psychology: Human Learning and Memory, 1980, 6, 778-784.

Herrmann, D.J. An old problem for the new psychosemantics: Synonymity. Psychological Bulletin, 1978, 85, 490-512.

Holyoak, K.J., Dumais, S.T. and Moyer, R.S. Semantic association effects in a mental comparison task. Memory and Cognition, 1979, 7, 303-313.

Holyoak, K.J. and Walker, J.H. Subjective magnitude information in semantic orderings. Journal of Verbal Learning and Verbal Behavior, 1976, 15, 287-299.

Hunt, R.R. and Einstein, G.O. Relational and item-specific information in memory. Journal of Verbal Learning and Verbal Behavior, 1981, 20, 497-514.

Hunt, R.R. and Elliott, J.M. The role of nonsemantic information in memory: Orthographic distinctiveness effects on retention. Journal of Experimental Psychology: General, 1980, 109, 49-74.

Hunt, R.R. and Mitchell, D.B. Specificity in nonsemantic orienting tasks and distinctive memory traces. Journal of Experimental Psychology: Human Learning and Memory, 1978, 4, 121-135.

Hyde, T.S. Differential effects of effort and type of orienting task on recall and organization of highly associated words. Journal of Experimental Psychology, 1973, 97, 111-113.

Hyde, T.S. and Jenkins, J.J. Differential effects of incidental tasks on the organization of recall of a list of highly associtaed words. Journal of Experimental Psychology, 1969, 82, 472-481.

Hyde, T.S. and Jenkins, J.J. Recall for words as a function of semantic, graphic, and syntactic orienting tasks. Journal of Verbal Learning and Verbal Behavior, 1973, 12, 471-480.

Jacoby, L.L. and Craik, F.I.M. Effects of elaboration of processing at encoding and retrieval: Trace distinctiveness and recovery of initial context. In L.S. Cermak and F.I.M. Craik (Eds.), Levels of Processing Human Memory. Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1979, 1-21.

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

Johnson-Laird, P.N., Gibbs, G. and de Mowbray, J. Meaning, amount of processing, and memory for words. Memory and Cognition, 1978, 6, 372-375.

Johnston, C.D. and Jenkins, J.J. Two more incidental tasks that differentially affect associative clustering in recall. Journal of Experimental Psychology, 1971, 89, 92-95.

Kerst, S.M. and Howard, J.H., Jr. Mental comparisons for ordered information on abstract and concrete dimensions. Memory and Cognition, 1977, 5, 227-234.

Klein, K. and Saltz, E. Specifying the mechanisms in a level-of-processing approach to memory. Journal of Experimental Psychology: Human Learning and Memory, 1976, 2, 671-679.

Krinsky, R. and Nelson, T.O. Task difficulty and pupillary dilation during incidental learning. Journal of Experimental Psychology: Human Learning and Memory, 1981, 7, 293-298.

Lane, D.M. and Robertson, L. The generality of the levels of processing hypothesis: An application to memory for chess positions. Memory and Cognition, 1979, 7, 253-256.

Lesgold, A.M. and Goldman, S.R. Encoding uniqueness and the imagery mnemonic in associative learning. Journal of Verbal Learning and Verbal Behavior, 1973, 12, 193-202.

Martin, E. Generation-recognition theory and the encoding specificity principle. Psychological Review, 1975, 82, 150-153.

Mathews, R.C. Semantic judgments as encoding operations: The effects of attention to particular semantic categories on the usefulness of interitem relations in recall. Journal of Experimental Psychology: Human Learning and Memory, 1977, 3, 160-173.

McClelland, A.G.R., Rawles, R.E. and Sinclair, F.E. The effects of search criteria and retrieval cue availability on memory for words. Memory and Cognition, 1981, 9, 164-168.

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

Moscovitch, M. and 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.

Moyer, R.S. Comparing objects in memory: Evidence suggesting an internal psychophysics. Perception and Psychophysics, 1973, 13, 180-184.

Moyer, R.S. and Bayer, R.H. Mental comparison and the symbolic distance effect. Cognitive Psychology, 1976, 8, 228-246.

Moyer, R.S. and Dumais, S.T. Mental comparison. In G.H. Bower (Ed.), The Psychology of Learning and Motivation. New York: Academic Press, 1978, Vol. 12.

Nelson, D.L., Wheeler, J.W., Jr., Borden, R.C. and Brooks, D.H. Levels of processing and cuing: Sensory versus meaning features. Journal of Experimental Psychology, 1974, 103, 971-977.

Norman, D.A. and Bobrow, D.G. Descriptions: An intermediate stage in memory retrieval. Cognitive Psychology, 1979, 11, 107-123.

Packman, J.L. and Battig, W.F. Effects of different kinds of semantic processing on memory for words. Memory and Cognition, 1978, 6, 502-508.

Paivio, A. Perceptual comparisons through the mind's eye. Memory and Cognition, 1975, 3, 635-647.

Paivio, A. Mental comparisons involving abstract attributes. Memory and Cognition, 1978, 6, 199-208.

Paivio, A. and Marschark, M. Comparative judgments of animal intelligence and pleasantness. Memory and Cognition, 1980, 8, 39-48.

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

Parkin, A.J. Specifying levels of processing. Quarterly Journal of Experimental Psychology, 1979, 31, 175-195.

Pollio, H.R. Associative structure and verbal behavior. In T.R. Dixon and D.L. Horton (Eds.), Verbal Behavior and General Behavior Theory. Englewood Cliffs, New Jersey: Prentice-Hall Inc., 1968, 37-66.

Postman, L. and Kruesi, E. The influence of orienting tasks on the encoding and recall of words. Journal of Verbal Learning and Verbal Behavior, 1977, 16, 353-369.

Postman, L., Thompkins, B.A. and Gray, W.D. The interpretation of encoding effects in retention. Journal of Verbal Learning and Verbal Behavior, 1978, 17, 681-705.

Potts, G.R. Information processing strategies used in the encoding of linear orderings. Journal of Verbal Learning and Verbal Behavior, 1972, 11, 727-740.

Potts, G.R. Bringing order to cognitive structures. In F. Restle, R.M. Shiffrin, N.J. Castellan, H.R. Lindman and D.B. Pisoni (Eds.), Cognitive Theory. Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1975, Vol. 1, 247-270.

Potts, G.R., Banks, W.P., Kosslyn, S.M., Moyer, R.S., Riley, C.A. and Smith, K.H. Encoding and retrieval in comparative judgments. In N.J. Castellan, Jr. and F. Restle (Eds.), Cognitive Theory. Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1978, Vol. 3, 243-308.

Roenker, D.L., Wenger, S.K., Thompson, C.P. and Watkins, B. Depth of processing: When the principle of congruity fails. Memory and Cognition, 1978, 6, 288-295.

Schnur, P. Testing the encoding elaboration hypothesis: The effects of exemplar ranking on recognition and recall. Memory and Cognition, 1977, 5, 666-672.

Schulman, A.I. Recognition memory for targets from a scanned word list. British Journal of Psychology, 1971, 62, 335-346.

Schulman, A.I. Memory for words recently classified. Memory and Cognition, 1974, 2, 47-52.

Seamon, J.G. and Murray, P. Depth of processing in recall and recognition memory: Differential effects of stimulus meaningfulness and serial position. Journal of Experimental Psychology: Human Learning and Memory, 1976, 2, 680-687.

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

Smith, A.D. Adult age differences in cued recall. Developmental Psychology, 1977, 13, 326-331.

Stein, B. The effects of cue-target uniqueness on cued recall performance. Memory and Cognition, 1977, 5, 319-322.

Stein, B.S. Depth of processing reexamined: The effects of the precision of encoding and test appropriateness. Journal of Verbal Learning and Verbal Behavior, 1978, 17, 165-174.

Stein, B.S., Morris, C.D. and Bransford, J.D. Constraints on effective elaboration. Journal of Verbal Learning and Verbal Behavior, 1978, 17, 707-714.

Thios, S.J. Memory for general and specific sentences. Memory and Cognition, 1975, 3, 75-77.

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

Tyler, S.W., Hertel, P.T., McCallum, M.C. and Ellis, H.C. Cognitive effort and memory. Journal of Experimental Psychology: Human Learning and Memory, 1979, 5, 607-617.

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

Whitten, W.B., II, Suter, W.N. and Franks, M.L. Bidirectional synonym ratings of 464 noun pairs. Journal of Verbal Learning and Verbal Behavior, 1979, 18, 109-127.

Winograd, E. Elaboration and distinctiveness in memory for faces. Journal of Experimental Psychology: Human Learning and Memory, 1981, 7, 181-190.

APPENDIX 1

Stimulus pairs used in the size comparisons in Experiment 1.

Size Ratio 1.5

ICEBERG	TENT
CAMEL	DISHWASHER
QUAIL	CHAIR
SALT	CATERPILLAR
BLENDER	APPLE
MULE	WINDOW

Size Ratio 2.0

RIBBON	RACCOON
VIOLIN	FREEZER
CRUMB	RAT
TRUCK	TRICYCLE
WOLF	BOOK
BED	RABBIT

Size Ratio 2.5

FROG	BICYCLE
LOBSTER	DOT
TURTLE	KANGAROO
AIRPLANE	PORCUPINE
THUMB	HAIRBRUSH
SNAKE	GIRAFFE

Size Ratio 3.0

SNAIL	BEAVER
THIMBLE	DOG
SPOON	BATHTUB
ZEBRA	BANANA
TEAPOT	ELEPHANT
VASE	CAR

Size Ratio 3.5

TREE	HAND
PENCIL	LION
WALLET	MOOSE
STOVE	QUARTER
GOAT	RAISIN
TRUMPET	ANT

Targets and distractors used in recognition conditions following the Size list in Experiment 1.

Target	Similar in Size	Similar in Meaning
Caterpillar	Louse	Worm
Blender	Toaster	Mixer
Quail	Duck	Partridge
Mule	Donkey	Donkey
Camel	Desk	Horse
Tent	Chest	Hut
Lobster	Knife	Crayfish
Frog	Mouse	Toad
Snake	Cat	Lizard
Thumb	Cigar	Finger
Turtle	Flute	Tortoise
Porcupine	Goose	Hedgehog
Trumpet	Pillow	Trombone
Raisin	Marble	Grape
Stove	Table	Oven
Pencil	Watch	Pen
Wallet	Cup	Billfold
Hand	Foot	Foot
Crumb	Roach	Morsel
Wolf	Sheep	Coyote
Rabbit	Skunk	Hare
Tricycle	TV	Wagon
Ribbon	Cord	Twine
Violin	Monkey	Viola
Snail	Bee	Slug
Spoon	Orange	Fork
Zebra	Elk	Antelope
Car	Bear	Van
Teapot	Kettle	Kettle
Thimble	Cherry	Needle

Stimulus Pairs used in the length comparisons in Experiment 1.

Length Diff. 1

WORKHOUSE	REPTILE
EARTH	MAIDEN
CORD	FLOOD
AIR	BOSS
SHOTGUN	PRAYER
PROPERTY	RACKETEER

Length Diff. 2

BEVERAGE	INSECT
LEOPARD	ANTITOXIN
CHASM	FOX
ROSIN	SPINACH
STAGECOACH	BACTERIA
SEASON	DOVE

Length Diff. 3

DAYBREAK	ELBOW
SEA	FRIEND
KETTLE	BEE
PUDDING	IRON
FABRIC	AVALANCHE
AX	JELLY

Length Diff. 4

TABLESPOON	MEADOW
NUN	SKILLET
MOSQUITO	CELL
BARNACLE	MULE
KEG	NURSERY
WOMAN	ALLIGATOR

Length Diff. 5

HOOF	FURNITURE
TOWER	LETTERHEAD
INFIRMARY	LAKE
LAD	BUNGALOW
UMBRELLA	FUR
MARRIAGE	ARM

Targets and distractors used in recognition conditions following the length list in Experiment 1.

Target	Physical Similarity	Semantic Similarity
Workhouse	Workshop	Sweatshop
Earth	Dearth	Soil
Flood	Floor	Deluge
Air	Ear	Oxygen
Prayer	Praise	Plea
Racketeer	Musketeer	Gangster
Insect	Insult	Spider
Leopard	Leotard	Panther
Chasm	Charm	Cleft
Rosin	Resin	Wax
Bacteria	Backbone	Bacillus
Dove	Cove	Pigeon
Daybreak	Daydream	Dawn
Friend	Fiend	Ally
Kettle	Nettle	Teapot
Iron	Icon	Steel
Fabric	Facade	Cotton
Jelly	Belly	Jam
Meadow	Mellow	Pasture
Nun	Bun	Sister
Cell	Bell	Jail
Mule	Mare	Donkey
Keg	Leg	Barrel
Woman	Roman	Female
Hoof	Roof	Foot
Tower	Bower	Spire
Lake	Cake	Pond
Lad	Lap	Boy
Fur	Cur	Hair
Arm	Arc	Leg

APPENDIX 2

Stimulus pairs used in the synonymy comparisons in Experiment 2.

Rating 6.5

ANSWER	REPLY
VICTOR	WINNER
TEACHER	INSTRUCTOR
SURPLUS	EXCESS
PRESENT	GIFT
RESULT	OUTCOME

Rating 5.5

PISTOL	GUN
HARM	DAMAGE
PICTURE	PHOTOGRAPH
BOLDNESS	DARING
POSTURE	STANCE
CAPTURE	SEIZURE

Rating 4.5

PRESENCE	EXISTENCE
LIGHTNING	THUNDERBOLT
LAUGHTER	CHUCKLE
PORTHOLE	WINDOW
RECOVERY	RESCUE
ENTERPRISE	PROJECT

Rating 3.5

TOWN	CITY
VETO	PROHIBITION
SOLID	MASS
SENTENCE	PHRASE
MOUTAIN	HILL
WORRY	TORMENT

Rating 2.5

THUNDER	CLAP
SUBURB	NEIGHBORHOOD
NEEDLE	SPIKE
PATIENT	INVALID
VISIT	CHAT
MUZZLE	NOSE

Targets and distractors used in recognition conditions in Experiment 2.

Target	Physical Similarity	Semantic Similarity
Reply	Replay	Response
Victor	Vintner	Champion
Teacher	Teamster	Educator
Excess	Excise	Glut
Gift	Gilt	Talent
Result	Resin	Consequence
Pistol	Piston	Firearm
Harm	Harem	Injury
Photograph	Phonograph	Snapshot
Daring	Darning	Bravery
Posture	Postern	Carriage
Capture	Capsule	Arrest
Presence	Prescience	Subsistence
Lightning	Lighting	Firebolt
Laughter	Lavender	Snicker
Window	Windmill	Portal
Rescue	Residue	Salvage
Project	Protest	Venture
Town	Towel	Village
Prohibition	Projection	Embargo
Solid	Solute	Lump
Sentence	Senate	Clause
Hill	Hail	Peak
Torment	Torpor	Torture
Clap	Clip	Peal
Suburb	Subway	Vicinity
Spike	Spice	Nail
Patient	Patent	Sufferer
Chat	Cheat	Gossip
Muzzle	Nuzzle	Snout

APPENDIX 3

This pilot experiment was conducted to examine the effect of decision difficulty on comparisons of pleasantness. Pleasantness was chosen as a dimension of comparison because pleasantness involves inferential subjective decisions. Thus, pleasantness could be used as a contrast with referential size, which involves judgments that are direct and objective in nature.

The experiment was not presented in the main body of the thesis, because it does not demonstrate an interpretable pattern of results. It was discovered that some bias in the selection of stimulus pairs made interpretation of the data somewhat suspect. However, the experiment does serve to demonstrate why pleasantness was used as a dimension of comparison in Experiments 3 and 4 of the thesis.

A stimulus list was constructed. The pairs of concrete nouns used by Paivio (1978, Experiment 1) in an investigation of the symbolic distance effect were used as the critical pairs in the list. The pairs had been rated for pleasantness and represented mean pleasantness ratios of 1.5, 2.5, and 3.5, with an equal number of pairs for each ratio.

Subjects decided for each pair of words in the list which member of the pair was more pleasant and rated the difficulty of each decision. After completion of the decision and rating task, subjects were given one of three retention tests, free recall or one of two recognition plus cued recall tests.

The dimension of pleasantness provides contrasts with the size dimension in Experiment 1. Both size and pleasantness are specific attributes of meaning. Size comparisons are based on objective aspects of stimuli; pleasantness comparisons are subjectively based.

Two different types of distractors were chosen for each of the targets for recognition. Type of distractor was defined as in Experiment 1 for the length list. Each distractor was chosen to vary minimally from a target word in terms of semantic or physical features. If any effect of distractor type occurs with the pleasantness comparisons, the effect should be that semantic distractors are more readily confused with targets than are physical distractors, which share little semantic overlap with the appropriate targets. Semantic distractors should resemble targets in pleasantness as well as general components of meaning.

Method

Subjects. Twenty subjects participated in each of three conditions. The 60 subjects were introductory psychology students who participated in the experiment as a course requirement. Subjects were tested in group sessions, with each group randomly assigned to a condition.

Materials

Presentation List. The pleasantness list was constructed using the 18 pairs of concrete words from a study by Paivio (1978, Experiment 1). The words comprising the pairs had been rated for pleasantness. The pairs represented mean pleasantness ratios of 1.5, 2.5, and 3.5 (e.g., FACTORY-BUCKLE, SNAKE-OASIS, NOOSE-FLOWER), with six pairs of words representing each of the three ratios. Ten buffer pairs were constructed by choosing twenty concrete nouns from Paivio, et al.'s (1968) norms and randomly pairing them. As will become apparent in the discussion, there are selection problems in the pairs adopted from Paivio (1978).

The pairs were recorded on videotape so that a single word pair appeared on the screen at a time. A five-second presentation rate was used. Five buffer pairs preceded the critical pairs and five buffer pairs followed. A random order within the structure of critical and buffer pairs was predetermined.

Recognition Test Lists

Two recognition test lists were constructed. Each test list consisted of an equal number of targets and distractors, representing each critical pair in the presentation list. A distractor was chosen individually for each target, with personal judgment determining the appropriateness of a distractor. The targets and distractors were positioned on the test sheet in random order.

Half of the target items were the more pleasant word in a presented pair, and half were the less pleasant. Distractors for the physical similarity test list were chosen to differ minimally in physical appearance from the corresponding targets from the presentation list (e.g., noose, nose). Distractors for the semantic similarity test list were chosen to be as similar as possible in general meaning to the corresponding targets (e.g., noose, rope). No manipulation of the pleasantness of a distractor was possible, as no pleasantness norms were available at the time the experiment was conducted.

Procedure

Decision and Rating. Subjects were told they were in an experiment examining decision difficulty. They would be shown 28 pairs of words; when a pair appeared on the screen they were to indicate the more pleasant referent by indicating left or right, corresponding to its position on the screen. After making a decision, subjects were to indicate the difficulty of the decision by rating it on a seven-point scale, with 1 representing an extremely difficult decision.

Memory Tests. After subjects completed the decision and rating task, they were given one of three retention tests. Subjects in the free recall condition were asked to write down as many words as possible from the presentation list. Subjects in the semantic similarity or physical similarity conditions were given the appropriate recognition test list and asked to circle the 18 words on the sheet which had appeared in the presentation list, and write in the word which had appeared with each during presentation. Subjects in each test condition were given five minutes to complete the task.

Results and Discussion

The .05 *a* level was used for all tests of significance.

Rating. Table 1 presents mean rating of decision difficulty as a function of pleasantness ratio along with the corresponding standard deviations. The rating data are collapsed across conditions because it was determined that the ratings did not differ across conditions.

The trend in difficulty rating is apparent; as pleasantness ratio increases, the perceived difficulty of making a comparison decreases. A 1x3 repeated measures analysis of variance performed on the data indicates that pleasantness ratio had a strong effect on difficulty rating [$F(2,118)=109$, $MSe=8.154$]. The results of a *t*-test comparison of the means indicates that difficulty ratings for the 1.5

Table 1. Mean rating of decision difficulty (on a seven-point scale with 1 representing an extremely difficult decision) as a function of pleasantness ratio. Standard deviations are in parentheses.

PLEASANTNESS RATIO		
1.5	2.5	3.5
5.58	6.41	6.84
(0.78)	(0.70)	(0.28)

and 3.5 ratios differ significantly. The manipulation of pleasantness ratio had an effect on decision difficulty.

Recognition. Table 2 presents the mean number of words correctly recognized, along with the corresponding standard deviation and probability of correct recognition, as a function of pleasantness ratio for the semantic similarity and physical similarity conditions. A 2x3 mixed design analysis of variance performed on the recognition data reveals no significant main effect (F's for condition and ratio < 1). The interaction of ratio and type of distractor was significant [$F(2,76)=4.21$, $MSe=0.70$]. Overall, recognition levels were approximately 85%, with false recognition rates of approximately 5%.

As with size comparisons, type of distractor had no systematic effect on recognition in the pleasantness conditions. The two types of distractor affected recognition in different ways, as evidenced by the significant interaction of ratio and distractor. Semantic similarity distractors had a more detrimental effect on recognition of items from large rather than small ratios, relative to physical similarity distractors.

Recognition in the semantic similarity condition follows the general difficulty pattern of a decrease in correct recognition with an increase in ratio. The pattern is reversed in the physical similarity condition; recognition increases as ratio increases.

An *ad hoc* interpretation of the conflicting patterns in the recognition data may fit a distinctiveness explanation. The detailed, precise contrasts required for the small-ratio comparisons enable the subject to reject semantically similar distractors as not possessing the contrasts present in the pair trace. As the pleasantness comparisons

Table 2. Mean number of words correctly recognized (of a possible 6) as a function of pleasantness ratio for each of the recognition conditions. Standard deviations are in parentheses. Probability of correct recognition is indicated by an asterisk.

	PLEASANTNESS RATIO		
Test Condition	1.5	2.5	3.5
Physical Similarity	4.75 (1.07) *.79	4.95 (1.15) *.83	5.30 (0.92) *.88
Semantic Similarity	5.40 (0.88) *.90	5.30 (0.80) *.88	4.90 (0.91) *.82
Averaged across Test Conditions	5.08 *.85	5.13 *.86	5.10 *.85

become less difficult, with an increased difference in pleasantness between pair members, semantic distractors are less easily discerned from presented items. Pleasantness becomes a dominant factor in the pair encoding as it provides sufficient information for making the comparison. However, pleasantness does not suffice to distinguish semantic distractors from their corresponding targets.

Subjects in the physical similarity condition may rely on absolute pleasantness judgments to distinguish targets from distractors for the larger ratio items. The physical distractors are not similar to corresponding targets in terms of pleasantness, and hence readily discernible. As pleasantness is an increasing dominant aspect of the encoding trace with an increase in ratio, correct recognition may increase with an increase in pleasantness ratio.

<u>Recall</u>. Table 3 presents mean number of words correctly recalled, with corresponding standard deviation and probability of correct recall values, as a function of pleasantness ratio for each of the recall conditions. Correct recall for cued recall and free recall conditions was scored as described in Experiment 1.

The free recall condition exhibits higher recall levels than do the cued recall conditions. The patterns of recall in the cued recall conditions are consistent with the patterns found in the corresponding recognition data. Any differences in recall level as a function of pleasantness ratio are very slight in the cued-recall conditions.

The results of a 3x3 mixed design analysis of variance demonstrate that all factors are significant: condition [$F(2,57)=17.3$, $MSe=1.3$)], pleasantness ratio [$F(2,114)=5.59$] and their interaction [$F(4,114)=4.28$, $MSe=0.99$]. The effect of condition is attributable to

Table 3. Mean number of words correctly recalled (of a possible 6) as a function of pleasantness ratio for each of the test conditions. Standard deviations are in parentheses. Probability of correct recall is indicated by an asterisk.

	PLEASANTNESS RATIO		
Test Condition	1.5	2.5	3.5
Physical Similarity	1.10 (1.12) *.18	1.40 (1.23) *.23	1.50 (0.95) *.25
Semantic Similarity	1.30 (1.17) *.22	1.45 (0.89) *.24	1.05 (0.76) *.18
Free Recall	1.45 (1.14) *.24	3.05 (1.05) *.51	2.60 (0.99) *.43

the superiority of free recall over the cued recall conditions; the source of the effect was confirmed by a t-test comparison of the means. Recall for items from the 1.5 ratio is significantly lower than for items from the other ratios, with the free recall condition again accounting for the effect.

The pattern of recall in the free recall conditions does not correspond with the pattern of free recall found in the size condition in Experiment 1. Memory traces apparently are not equally available for items from the different pleasantness ratios. The 1.5 ratio items are particularly poorly recalled. An affective dimension such as pleasantness may not lend itself to precise relative comparisons, individual absolute judgments being more salient than the comparisons.

Packman and Battig (1978) state that words from the extremes of pleasantness are more likely to be recalled than are neutral words. Post hoc examination of the stimuli used in the experiment, after receipt of a set of pleasantness norms thanks to Marc Marschark and A. Paivio, reveals that the 2.5 and 3.5 ratio items are all from the extreme thirds of the norms; only half of the 1.5 ratio items are from the extreme thirds, that half being paired with words from the neutral third of the norms. Sixty-five percent of the 1.5 ratio items produced in free recall come from the extremes of the norms. Such a disparity or bias calls into question the validity of effects found in the recognition and recall data. Other experiments reported here were designed to eliminate the possibility of extreme pleasantness values differentially affecting some degrees of decision difficulty.

Stimulus pairs used in the experiment.

Pleasantness Ratio 1.5

CACTUS	SCORPION
FACTORY	BUCKLE
PEACH	BOWL
WAGON	STRAWBERRY
FIREPLACE	BARREL
HAMMER	APPLE

Pleasantness Ratio 2.5

BULLET	PIANO
SNAKE	OASIS
DAFFODIL	CRUTCH
WITCH	MOUNTAIN
HARP	DEVIL
EYE	SPIDER

Pleasantness Ratio 3.5

FLOWER	NOOSE
BED	MOSQUITO
KU KLUX KLAN	BABY
GUN	DOVE
ROSE	HANDCUFFS
EXECUTIONER	BUTTERFLY

Targets and distractors used in the recognition tests.

Target	Physical Similarity	Semantic Similarity
Cactus	Canvas	Thistle
Buckle	Chuckle	Clasp
Hammer	Hamper	Mallet
Barrel	Barrow	Keg
Strawberry	Raspberry	Raspberry
Peach	Beach	Pear
Devil	Evil	Demon
Spider	Spinet	Insect
Snake	Snare	Asp
Witch	Watch	Wizard
Piano	Paint	Organ
Crutch	Crust	Cane
Baby	Baboon	Infant
Noose	Nose	Rope
Gun	Gum	Rifle
Executioner	Executive	Hangman
Rose	Rosin	Thorn
Bed	Belt	Cot

APPENDIX 4

Stimulus pairs used in Experiments 3 and 4.

Size List

Pair Separation 0		Pair Separation 3		Pair Separation 6	
THIMBLE	ERASER	EAGLE	HORSE	TUBA	ANT
WINDOW	FOX	SQUIRREL	MOTH	TOMATO	BED
CRAB	MOUSE	DUCK	TRUNK	MAN	CATERPILLAR
LADDER	TABLE	BANANA	PILLOW	CAMEL	MITTEN
MOOSE	BOAT	BUTTON	SNAKE	WOLF	BEE
SNAIL	PEANUT	SAUCER	CHERRY	SHOE	GARAGE
FREEZER	STALL	BEAR	DOG	PRUNE	DESK
LOBSTER	TURTLE	PIANO	GOAT	FROG	TIGER

Pleasantness List

Pair Separation 0		Pair Separation 3		Pair Separation 6	
DEER	EYE	TIRE	BULLET	CLOWN	SKUNK
COBRA	BRUTE	WEB	LEMON	CIGAR	ACORN
MARBLE	PEAR	BOTTLE	PLUM	HOUSE	BOWL
DAISY	BOOK	BARREL	WASP	DRUM	MOON
TWEEZERS	FIRE	FLAG	PIANO	CELERY	RAZOR
CACTUS	ANVIL	ROSE	CARROT	LEAF	WAGON
DONKEY	SCISSORS	CANARY	THUMB	YAM	FLUTE
KEY	SHORTS	PIPE	SANDAL	FIST	CRADLE

APPENDIX 5

List of specific encoding descriptions and the word pair appropriate to each description, used in Experiments 5 and 6. The portion underlined represents the descriptive cue in Study 1 of Experiment 6. (All of the statements when presented were preceded by "Are these..." in the encoding task.)

types of synthetic cloth	NYLON	RAYON
periodical reading material	NEWSPAPER	MAGAZINE
high-ranking members of the clergy	CARDINAL	BISHOP
types of alcohol made from grain	GIN	RYE
footwear worn on the beach	SANDALS	THONGS
precious stones that are milky in appearance	OPAL	PEARL
physical sciences	CHEMISTRY	GEOLOGY
military ships	CRUISER	SUBMARINE
contagious diseases	MEASLES	MUMPS
destructive weather phenomena	TORNADO	CYCLONE
exploding weapons	BOMB	GRENADE
parts of a building used for storage	ATTIC	CLOSET
toys that spin	TOP	YO-YO
flowers used in corsages	ORCHID	CARNATION
fish that live in the ocean	SHARK	SALMON
buzzing insects	BEE	WASP
types of classical music	CONCERTO	SONATA
appointed professions	JUDGE	SENATOR

sports that use poles	SKIING	FISHING
vehicles that run on rails	TRAIN	STREETCAR
violent crimes	MURDER	ASSAULT
long (intervals) units of time	EON	ERA
temporary dwellings	TENT	HOTEL
predatory birds	EAGLE	HAWK
chemical elements that glow	PHOSPHORUS	NEON
vegetables that grow underground	POTATO	ONION
animals that have hooves	DEER	SHEEP
metals used in construction	COPPER	STEEL
earth formations deeply cut into the ground	CANYON	RAVINE
nautical measures of distance	FATHOM	KNOT
poisonous snakes	COPPERHEAD	COBRA
carpenter's tools used to loosen things	PLIERS	WRENCH
military titles of non-commissioned officers	SERGEANT	CORPORAL
blossoming trees	DOGWOOD	CHERRY
Latin countries	SPAIN	ITALY
beverages that contain stimulants	TEA	COKE
Fuels that burn	COAL	WOOD
parts of the body that can be wiggled	EAR	TOE
fruits that peel	BANANA	ORANGE