71-7280

HOCHHAUS, Larry, 1942-
EDITING PROCESSES IN MEMORY.

Iowa State University, Ph.D., 1970
Psychology, experimental

University Microfilms, Inc., Ann Arbor, Michigan

EDITING PROCESSES IN MEMORY

by

Larry Hochhaus

A Dissertation Submitted to the
Graduate Faculty in Partial Fulfillment of
The Requirements for the Degree of
DOCTOR OF PHILOSOPHY

Major Subject: Psychology

Approved:

Signature was redacted for privacy.

In Charge of Major Work

Signature was redacted for privacy.

Head of Major Department

Signature was redacted for privacy.

Dean of Graduate College

Iowa State University
Ames, Iowa

1970

TABLE OF CONTENTS

INTRODUCTION AND LITERATURE REVIEW

"Editing" in memory usually means a process which occurs somewhere between the implicit responses to a stimulus and the final overt response (Melton, 1967). Furthermore, the concept seems to be intimately related to "*knowing*" about the correctness or incorrectness of a response.

Several important behavioral phenomena appear to point directly to underlying editing processes, yet because most of these phenomena are well known and commonplace and because of the covert nature of editing, there has been little in the way of systematic documentation for these behaviors.

In an early investigation of multiple list learning McKinney and McGeoch (1935) noted frequent interlist intrusions as well as several cases of words from other lists written down and (incompletely) scratched out. This scratching out of responses may be viewed as evidence for an editing mechanism which is based on the *S*'s identification of the incorrectness of his responses. Likewise, the intrusions can be considered as cases of failure of the editing mechanism.

Similarly, other interference theorists have long used the concept of a "selector mechanism," which operates through knowledge of list membership, to explain the omission of competing responses (Underwood and Schulz, 1960). According to this view, the *S*s generate implicit responses, recognize these when wrong, and suppress responding.

Additionally, "tip of the tongue" behaviors are closely related to the process of editing. Commonly Ss will make overt errors which they quickly reject as incorrect, seemingly aware that they are near the correct answer, but not able to immediately produce it. For example, in trying to recall a person's name the S's responses might be, "Pearson, Peirce, Patterson, Patten, Peters, Peterson" (correct). Many authors have given descriptive accounts of the "tip of the tongue" phenomenon (e.g., Freud, 1954; James, 1950), and recently the behavior has been investigated under laboratory conditions.

By reading dictionary definitions of rare words to Ss, Brown and McNeill (1966) were able to produce several "tip of the tongue states." They found that in such states Ss were variously able to indicate parts of the missing target words (letters, prefixes) as well as information about the abstract form of the word sought (number of syllables, syllabic stress, meanings).

The retention of meanings in the absence of specific item identification has also been shown by Yavuz and Bousfield (1959). Their Ss learned to respond to paralogs (nonsense syllables presented under the guise of "Turkish words") with English "translations." Even in the absence of recall, semantic-differential ratings of the paralogs indicated retention of connotative meanings appropriate to the absent English words. Apparently, emotional reaction may be an attribute of memory which is stored as a component of the whole item and may be retained independently.

In tip of the tongue states *S*s report that they have strong feelings that they know the item sought. Hart (1965) emphasized the actual accuracy of these feelings. He began by reading general information questions (e.g., "Who wrote 'The Tempest'?"), and when answers were not forthcoming, asked *S*s to predict their ability to recognize the correct answer from among other incorrect alternatives. They were able to do so with an accuracy significantly above chance.

Aware that *S*s may sometimes withold responses from recall due to a general conservatism (response bias) and that such responses could be the source of feelings of knowing, Hart (1966) replicated the original study using forced responding. This procedure reduced the ability to predict recognition, but performance was still significantly above chance. (Note, however, that this factor was not controlled in the Yavuz and Bousfield study.)

Recent work by Hart has dealt with the "feeling of knowing" phenomenon under more controlled conditions as well. Here control was exercised over the input phases of memory in addition to testing. In a paired-associate task, Hart (1967) used words as stimuli and consonant trigrams (CCC's) as responses. One, two, or three presentations of the 42-pair list were given followed by a recall test. Employing his usual "RJR" paradigm (recall-judgement-recognition), Hart found that *Ss* could also accurately predict the recognition of the trigrams which they couldn't recall.

In the 1967 report Hart indicated that for his measures to be maximally accurate in terms of "memory monitoring" ability, *Ss* should predict recognition in the same proportion as they are actually accurate in recognizing. That is, the difficulty of the recognition test is an important factor in the measured ability to predict, and, by the same reasoning, so are biases for or against reporting feelings of knowing. The confounding of response bias with the dependent variable represents a strong argument for the use of alternative methodologies.

On the other hand, a measurement technique directly relevant to confidence rating accuracy has been developed. Although originally interpreted as an index of memory (Murdock, 1966), the "Type II" *d*' of Signal Detection Theory is now usually interpreted as a measure of correct versus incorrect response discriminability (Hochhaus, 1970; Murdock, 1970). The logic of the analysis is quite simple; *Ss* use confidence ratings to indicate the certainty of their responses. The degree of match between high ratings and correct responses is given by the *d*' measure; the tendency to under- or overuse particular ratings is given by β (response bias). The *d'* and β measures are statistically independent.

By way of summarization, the Hart research has shown that *Ss* are able to accurately predict recognition even when recall memory is absent. Hart has proposed a "memory monitoring (MEMO)" process to summarize this talent, yet has left the dynamics of the MEMO process completely unspecified; he implies that the MEMO system operates prior to or in the absence of retrieval (Hart, 1966, 1967).

Another interpretation of the "feeling of knowing" phenomenon has been given by Adams (1967) in terms of a two-trace memory system. The "closed-loop" theory of Adams is based on a feedback model which was first proposed by Mowrer (1960); in Adams' words, "the response of the system is fed back and compared with a *reference level* which defines the correct value for the system. If there is an error difference between the response and the reference level, the system undertakes an adjustive correction to lessen the error, compares the resultant response with the reference level again, adjusts again, and so on," (pp. 291-292). Editing processes are handled by treating recall as the joint effect of independent retrieval and recognition (reference level) processes.

When recall attempts produce a close match with the independently stored recognition traces, *Ss* will report "yes" as their feeling of knowing (as long as the recognition traces are strong ones). On the other hand, negative feelings of knowing are presumed to arise from far matches as a result of weak recognition traces. Following the judgement stage of the feeling of knowing experiment Ss take the recognition test; since Adams' theory states that performance on the recognition test is based on the recognition traces, the result is a correlation between feeling of knowing and recognition. In the present review, however, it will be argued that the two process views of Hart and Adams are perhaps not parsimonious with regard to an alternative, simpler conception of the feeling of knowing phenomenon based on a single-process memory model.

For any theoretical account the important problem concerning the feeling of knowing process is understanding how the S can know his capacity for a correct response and how he is later able to prove this on a recognition test. These are the questions to which Study I of the present investigation is directed.

The mechanism used here as an explanation of the "feeling of knowing" phenomenon is that of partial recall. According to this view, Ss in some cases recall less than the whole item, and it is on this basis that they can recognize items which they cannot recall.

To test partial recall as a factor in recognition

McNulty (1965) used recognition distractor items which shared many parts of the target word. For example, when Ss were forced to choose between word pairs such as boldness and coldness, the difficulty of recognition approached that of recall. However, the fact that recognition was still somewhat superior to recall may also have been due to part-whole processes. The Ss in this case may not recall because they only remember the first letter; yet this partial information is enough to make recognition possible.

The present theoretical approach is similarly based on a multicomponent view of memory. It is proposed that an item is represented in memory as a list of discrete features. For example, "lion" might be encoded as animal, ferocious, name begins with L, ... Such a scheme is similar to other multicomponent models (e.g., Bower, 1967; Brown and McNeill, 1966; Norman and Rumelhart, 1970; Underwood, 1969) and no great elaboration is intended here. Rather, the data of Brown and McNeill are taken as evidence that the partial recall of these features is possible and it will be argued that such a mechanism provides a possible explanation of the "feeling of knowing" phenomenon.

Support for this approach is found in a study by DaPolito, Guttenplan, and Steinitz (1968) who modified Hart's procedure by adding an "elimination stage" to the RJR technique. After judgement but prior to recognition Ss were asked to indicate recognition alternatives which they felt were definitely incorrect. The Ss eliminated more incorrect alternatives following a "yes" judgement than following a "no" judgement.

Partial information retained, according to the DePolito et al. and the present view, can explain the result in terms of a hypothetical scanning mechanism which allows Ss to reject incorrect alternatives on the basis of a "mismatch" against this partial information. If this is the case, then one can expect the recall of partial information to account for feelings of knowing. Study I is an attempt to support this hypothesis by comparing partial recall measures and feeling of knowing judgements on the basis of the capacity to predict recognition.

In the second phase of the present research the partial recall model is related to multiple-choice recognition memory. Additionally, the second guessing paradigm is emphasized, since one of the questions to be investigated concerns the conditions under which Ss can exhibit accurate editing responses (certainty judgements) for second guesses.

First choice errors in the multiple-choice task, like recall errors, do not mean that the S is without information as to the correct alternative. Second guessing data provide the primary support for this statement. For example, in a five-choice recognition test the S is informed of his errors and is asked to guess again. The probability correct for second choices, conditional on error, is not one-fourth as one might expect from an "all-or-none" view, but has been shown to be significantly above chance (e.g., Bregman, 1966; Brown, 1965a). The second guess paradigm has also been effectively used in the study of perceptual thresholds and tachistoscopic recognition (e.g., Swets, Tanner, and Birdsall, 1961).

Brown (1965b) has developed an index to estimate "the average probability of rejecting an incorrect alternative" and has shown its relevance in comparing performances on tests which vary in the number of response alternatives. It is one of the purposes of Study II to approach second guessing behavior in another fashion and attempt its explanation through the previously described concept of partial recall.

One criterion for simulation models is "Can it be shown how such a mechanism could be built?" Here a similar question is asked of the multicomponent model, "Can it be shown how such memory representations could lead to the observed phenomenon?" To affirm this, the following hypothetical situation is described; it is felt that such speculation is warranted by the testable propositions which can be derived from the multicomponent model.

Consider again the earlier example and that the S is asked to associate the word lion to some stimulus. Presume further that the orthographic attribute "name begins with L" was forgotten while "animal" and "ferocious" were retained. Next, assume the following multiple-choice recognition terms are given as a test of memory for the associate of the stimulus; book, caterpiller, tiger, lion. It will be seen that the S can eliminate the first two alternatives on the basis of his partial recall, yet must guess between tiger and lion. If he chooses tiger and is told he is wrong, it follows that his second choice will very likely be correct.

The reader should recognize that the number of features encoded may be more than those specified in the present example (cf. Underwood, 1969); the proposed approach is not an attempt to specify these features, it is merely meant to be a sketch of how accurate second guesses might occur. However, several theoretical hypotheses do follow from the multicomponent view.

First, multiple-choice errors should in most cases depend on the correct response. Since, hypothetically, several items may share the remembered attribute or attributes, e.g., "ferocious," the S may be led to an incorrect response. This is especially true if a "liberal naming strategy" (Norman and Rumelhart, 1970) is adopted, as would be the case under forced response instructions. Following such an error the S would be asked to make a second guess. Since he knows the first answer is wrong he can choose another item which shares the quality "ferocious" and this time have a very good chance of being correct. Thus it is proposed that, given the S chooses an incorrect alternative, Y, the response most likely correct is that alternative most similar to Y. In other words, it is felt that errors in multiple-choice responding bear information about the correct alternative.

Next, it is proposed that if and only if such information does present itself in first choice errors, then second choices will be above chance accuracy. The amount of information about the correct alternative which is conveyed in errors should correlate highly with performance on second choices.

Another requirement of the multicomponent mechanism is that second choices be similar to first choices (i.e., first and second choices should show common attributes). In information theory terms the information in first choices should be transmitted to second choices.

Next, the number of attributes remembered should influence the S's confidence rating projections of second test performance. The theoretical question here is quite similar to the relationships described for Study I. The more attributes recalled, the more a response should appear to be on the "tip of the tongue" and as a result, the S should become more confident that the correct response could be identified. Also, a condition implied by this is that confidence ratings should relate to second choice performance. That is, high ratings should indicate a higher probability correct than low certainty ratings.

If response alternatives could be constructed which would make all alternatives highly similar, then information in errors would be minimal and also the performance on second choices would be impaired. However, it is difficult to find such simple elements. What is required is a class of figures which share no common attributes, i.e., each figure is unique to itself within the set. It is possible that no such set exists. For example, the set 1, 2, 3, 4 is differentiable into "odd" and "even"; the last three items in the set a, b, c, d all contain the basic "ee" sound. The alternative to trying to fix a boundary condition for response similarity is to make investigations of more than one set, establishing the hypothetical null condition described above as a reference point from which to judge the model at points more and less removed from the boundary state. According to this view, response similarity should (1) decrease information in errors, (2) decrease second choice accuracy, (3) decrease information communicated from first choices to second responses, (4) decrease the relationship between first responses and later confidence ratings, and (5) decrease the discriminability of correctness for second choice responses.

STUDY I

The present investigation was an attempt to provide evidence for partial recall as a mechanism underlying the ability of *Ss* to predict accurately their recognition of items which they could not recall. A "study-test" paired-associate task was employed in which the recall responses were such that partial recall scores could be easily obtained. It was predicted that partial recall scores provide a substitute measure for "feeling of knowing" judgements made by the *Ss*.

Method

Subjects

The 24 *Ss* were students from Iowa State University introductory psychology courses; extra credit was given for participation. An equal number of *Ss* were assigned to each of the two conditions (1 versus 2 training trials) on the basis of order of appearance.

Materials

The response terms learned were highly codifiable "concepts" presented as figures. These consisted of four dimensions of two attributes per dimension. Dimensions and attributes were: Number- 1, 2; size- large, small; color- white, black; shape- square, triangle. The 16 response terms were the unique combinations of these attributes and were paired with 16 high meaningfulness nouns (stimulus terms) from the Pavio, Yuille, and Madigan (1968) list. Two orders of these pairs were prepared for presentation on 3 x 5 cards.

Procedure

Using the "study-test" method of paired-associate learning, *S*s viewed the lists under instructions to learn as many pairs as possible. Following the learning stage *S*s in both conditions were given a recall test. The recall test consisted of a sheet containing the 16 nouns, each followed by a grid which listed all possible attributes (see Appendix A). The recognition test was also printed on this sheet and was folded under and kept out of sight until the end of recall.

The *S*s indicated their recall responses by circling one attribute for each dimension, where this mode of response had been fully described in the instructions prior to the learning stage. Next, the *S*s placed confidence ratings (6 point scale) beside each noun to indicate their "feelings that they could recognize the correct figure on a multiple-choice recognition test." It was stressed that these ratings should reflect the degree to which they felt they knew the answer. (For full instructions, see Appendix B).

After the recall task was completed Ss were given the recognition test. After turning their answer sheets over to the recognition test, the Ss made a check mark beside one of the two figures paired with each noun. Completion of the recognition test terminated the experimental session.

Results

Analysis of recall and recognition scores

For the group given one presentation of the 16 pairs the average recall and recognition scores were 4.45 and 12.91, whereas for the Ss given two presentations the corresponding statistics were 7.75 and 14.58. Memory scores were significantly higher for the two-list group for both recall ($t=3.00$) and recognition ($t=2.06$); $p<.01$ for each.

Signal detection analyses

The discrimination of recognition correctness was investigated both as a function of feeling of knowing judgements and as a function of partial recall scores. However, since the ability to discriminate the correctness and incorrectness of recognition is indeterminate when recognition is perfect, Ss who did not make recognition errors were omitted from analysis. There was one such S in the single presentation condition and four in the group who received two presentations. One additional S in the latter group was not used because of omissions on both recall and recognition items. This occurred in spite of the general instructions to guess on every item not recalled or recognized. The discriminability in feeling of knowing judgements is also only of interest in the absence of full recall. Therefore, items for which all four attributes were recalled were also omitted from analysis.

In general there were few recognition failures and it was not possible to obtain stable operating characteristic (ROC) curves for the present data. Instead, a single d' (discriminability) index was calculated for both part scores and confidence judgements within each of the two groups of Ss. Part scores were the number of attributes (less than four) recalled for each item and for the detection analysis were treated the same as the confidence ratings of items.

Within a group part scores and confidence ratings were partitioned into "yes" and "no" responses at a point determined for each S as his point of neutral bias. This was done so the event matrices could be collapsed over Ss without fear that bias differences would lower the discriminability index.

All tests of d' values were computed using the formulae of Gourevitch and Galanter (1967). For the group who received one list presentation the part score and confidence judgement d''s were .29 and .37; these values did not differ significantly (G=.30). For the group who received two lists the corresponding values were .99 and .75, and these also were not different (G=.66).

Each *d*' was, however, significantly greater than zero ($p < .05$ for each single list *d*'; and $p < .01$ for each under the two list condition). Additionally, the part score *d*' was significantly higher for the two presentation group (G=2.23, $p < .05$), but confidence rating *d*''s did not differ as a function of number of lists presented (G=1.22).

To summarize these results, both part scores and confidence ratings indicate accuracy in the prediction of future recognitions. However, the feeling of knowing judgements did not surpass part scores in this regard. Further, there was evidence that the discrimination of recognition correctness was better for the condition of greater learning.

Discussion

It appears that the concept of partial recall as an explanation of the feeling of knowing phenomenon has received support. However, there are some limitations on this conclusion. First, the high recognition scores detracted from the stable estimation of recognition predictability; the two-choice recognition items were undoubtedly responsible for this. It is possible that more false recognitions (the analogous counterpart of noise trials in the signal detection experiment) would have produced a more powerful test of d' differences, and perhaps this manipulation would have shown greater predictability for the feeling of knowing judgements. The reversal of this trend found in the two-list group, however, tends to minimize the lack of power argument.

Next, the relationship between d' and recognition performance is in essential agreement with a recent report by the present author in which factors determining the discrimination of correctness in multiple-choice tasks were analyzed (Hochhaus, 1970). Guessing was found to be an important factor and it was argued that d' should be expected to correlate with probability correct. Evidence for this was also found in the present study; the d''s were higher in the group receiving two list presentations.

Although the present data indicate that the discrimination of future recognition correctness could be based on the recall of parts of an answer, it is not strictly necessary that this be the case. Feelings of knowing could be based on sources of information as yet unspecified, whereas the partial recalls might relate to recognition performance in an independent fashion. At this point it should be worthwhile to point out alternative explanations of the "feeling of knowing" phenomenon.

First, response competition might produce situations where the S considers more than one response alternative as an answer. In cases where an incorrect associate is given as an intrusion the S may still feel the correct response is ultimately available. The test of this possibility would require the provision for or the encouragement of multiple responses. Following this, an analysis of feeling of knowing judgements for items to which the correct answer was never given might show little or lessened recognition correctness discrimination.

Another possibility is that Ss can remember having learned something without actually knowing what it is they have learned. For example, in scanning reference books for forgotten information, people often report whether the material sought was on a right or left hand page; it seems likely that even the quadrant of the page could be remembered. Thus, it could be that "knowing that you know" is based on a direct process of "knowing that you once knew." Experimental evidence is lacking here, however, and it would appear that investigation of the phenomenon of memory for learning locations might be fruitful in leading to understanding the feeling of knowing experiences.

Still another means of predicting recognition could be based on the S's assessment of the difficulty of the question being asked. In the original experiments (Hart, 1965, 1966) general information questions were used. Cues within the question such as the familiarity of terms used might provide a basis for discriminating recognition capacity. This argument breaks down, however, for the paired-associate tasks, although Ss could conceivably discriminate the recency of having seen the stimulus on the previous learning trial and use this information to judge recognizability. As such, the process would require knowledge and application of learning curve principles, and does not appear to be a strong explanatory concept for Hart's later studies and the present investigation.

As a final alternative, the two-process model of memory perhaps deserves the most consideration. Although similarities and differences between recognition and recall have been investigated for many years now (e.g., MacDougall, 1904; Luh, 1922) it is surprising how little has been learned. Arguments for a single-process account of recognition and recall (e.g., Postman, Jenkins and Postman, 1948) seem as frequent as two-process views (e.g., Kintsch, 1970). Like Adams (1967), Kintsch argues for a two-process explanation and perhaps offers the most convincing argument to date, although he does not relate his model to the "feeling of knowing" phenomenon. The following is a review of what Kintsch feels are "important qualitative differences" between recognition and recall.

First, the meaningfulness relationship is reversed for recognition. Low-frequency Thorndike-Lorge words are better recognized than high-frequency words (e.g., Shepard, 1967). Also, in unpublished research the present author has found this reversal with words scaled for meaningfulness (words were rated for "ideational frequency and intensity" by independent Ss). In contrast, however, it should be noted that Martin (1967) found no such reversal of the meaningfulness relationship with CVC syllables; high meaningfulness syllables were recognized better than low meaningfulness ones.

Next, the intention or determination to learn has been found to aid recall but not recognition (Achilles, 1920; Hollingworth, 1913). Kintsch used similar findings (Postman, Adams and Phillips, 1955; Estes and DaPolito, 1967) to argue that recall involves an additional search process while recognition does not.

The next basis for Kintsch's two-process conclusion is the lack of interference effects in recognition. Although recognition studies have shown both proactive inhibition (Peixotto, 1947) and retroactive inhibition (McKinney, 1935; Postman, 1952), the magnitude of these effects has been less than that in recall. In some cases interference effects have been almost entirely absent (Bower and Bostrum, 1968; Postman and Stark, 1969).

Last, the organization (i.e., inter-relatedness) of items to be learned does not appear to be a factor in recognition (Kintsch, 1968). The strong facilitative effects of organization on free recall are, however, well known (e.g., Cofer, Bruce, and Reicher, 1966), and Kintsch used this to add to his argument that search processes are not a factor in recognition memory.

Because the factors of meaningfulness, intent to learn, interference, and organization all seem to show highly different effects on performance depending on the method of measurement, Kintsch has concluded that "the basic difference between recall and recognition appears to be that recall involves a search process and recognition does not," (p. 337). However, this account and the two-trace system proposed by Adams (1967) do not seem operationally distinct.

In terms of the multicomponent view, it is as if items in memory consist of a set or cluster of attributes containing two basically distinct types of attributes or "tags," those used in retrieval and those used in recognition. For the purpose of exposition let's call these attributes "S-features," or "R-features," depending on whether they are used in search or recognitive processes. S-features would be used in locating an item in memory, presumably by some sort of "content addressable" system. R-features, on the other hand, could not serve this purpose and could only be used as a basis for identifying the appropriateness or correctness of the item. This identification might be based on novel or ideosyncratic association leading <u>from</u> the item to other memory locations.

While such a conceptualization is not alien to the multi-component view, it is contrary to the spirit of the partial recall account proposed here as the basis of the "feeling of knowing" phenomenon. In effect, the present data do not require the postulation of separate R-feature processes; it is felt that small groups of S-features (partial recalls) are sufficient for the prediction of recognition by individuals.

In spite of this adequacy, however, several important questions are left unanswered concerning the role of recall in the recognition process. Yet, on the other hand, certain paths of inquiry are indicated. First, to what extent is recognition based on information not available through recall? Is the information retrieved through search sufficient to explain recognition performance? What should false recalls say about the capacity to recognize; for example, if in the memory of nonsense syllables the correct response is CHJ and the S recalls "CHX" does this mean he is guessing about the last letter or does it mean that his stored information has dictated his response? What if he had instead been asked to make a recognition choice between CHJ and CHX; which would he have chosen?

Unfortunately, certain properties of logical analysis may preclude the answer to this last question. We would like to know what recognition choice would be made if the S were tested with the right answer in addition to the answer most representative of his recall storage, yet we cannot know the latter unless we have actually tested recall.

If recall is tested and later the S is given a contingent recognition test, part of the problem is that it is possible that because we have introduced this activity the intrusion could be learned and upon subsequent recognition testing the S may be in a memory state different from that just prior to recall. Furthermore, if the recognition test is based on the S's recall performance, experimental control is lost; cause-effect relationships cannot be inferred.

Recognition distractors based on errors common to a large population of Ss might help answer the question experimentally, and such a study has been done. Dale and Baddeley (1962) found that the use of common recall errors as recognition decoys reduced recognition; yet performance was still accurate. The difficulty here is the lack of control over ideosyncratic errors, and we return to the problem of sampling or correlational analysis.

A meaningful alternative would be to observe recall performance and infer what recognition performance would have been on the basis of partial recalls. An assumption is made that the recalls best represent what has been stored in memory, and the question then is whether or not this information is sufficient to account for recognition performance---without the necessity of postulating additional information accessible only in recognition testing. It is possible that the data of the present experiment might shed light on this question.

Because of the contrast of approach between the present view and the dual process notion of recall and recognition memory, an additional analysis was carried out on the data of Study I. It is as if a simulation of recognition were carried out on the basis of recall performance. Instead of using separate Ss for recall and recognition, or the same Ss tested at different times with similar materials, identical materials were used with recognition following recall. This method takes advantage of intra-individual covariation in performance, and, fortunately, makes error learning in the recall interval of diminished importance to the conclusion of equality between recognition and recall. Recall testing prior to recognition does reduce recognition (Postman, Jenkins and Postman, 1948), but such an effect would serve to make the simulation easier. Thus, a simulation failure could be considered as evidence for the existence of process differences between recall and recognition, with the latter having access to information beyond that of recall.

The following is a report of a simulation of recognition based on partial recall scores. As indicated, the data of Study I were used as the basis of the simulation.

For each S-R pair for each S, the parts of a response were analyzed to predict each later recognition choice that would be made for that pair. The (simulated) recognition choice was always that term which overlapped on the greatest number of dimensions with the recall response. When ties occurred a guess was recorded and it was assumed that in the two-choice situation that prevailed half of these guesses would be correct. Next, the 16 simulated recognition responses for each S were combined to provide a simulated error total; these are shown in Figure 1, as are the actual recognition error scores achieved.

It can be seen that simulated error scores are higher than actual error scores, and a Sign test (Walker and Lev, 1953) showed the difference as significant ($p=.01$). It appears that at the time of recall, Ss did not provide all of the information about the stimuli which was ultimately available to them at the time of recognition testing. Several factors, however, must be considered before the data are taken as evidence for the dual process theory.

First, the memory items were presented as visual figures in the recognition test and symbolically in recall testing (see Appendix A). It is possible that other combinations of testing modes would have changed the result. Second, because of differences in difficulty between recall and recognition in terms of the information and number of

	1 presentation		2 presentations	
	simulated	actual	simulated	actual
	3.5	3	3	3
	5	4	2.5	3
	6.5	6	2	1
	5	3	4.5	6
	1	1	1.5	2
	2.5	1	2.5	1
	1	0	3	1
	5.5	8	.5	0
	5	3	1	0
	4	0	1	0
	3	3	1.5	0
	4	6	0	0
	9	5		
$\bar{X}$	4.23	3.31	1.92	1.41

Figure 1. Recognition error scores; simulated and actual performance

alternatives with which Ss dealt, motivational factors could have made recall responses less informationally rich; in future studies of this design it would be advisable to pay Ss on the basis of recall and recognition performance to insure a concerted approach to both tasks. Last, it was assumed that in cases of conflicting partial recalls a guess would be made and that half of these guesses would be correct. On the other hand, it is possible that Ss didn't treat each recalled dimension as equally valid in making their recognition choices. If Ss could weigh these validities appropriately, it is even possible that all cases of ties would have led to correct recognition.

In spite of these criticisms, the outcome of the simulation is in line with the dual process interpretation; it remains to be seen whether controlled experiments specifically designed to test this issue will replicate the present effect. Thus, while the data are somewhat equivocal, the present approach has provided a research design which shows some promise for distinguishing the two theoretical viewpoints.

STUDY II

Study II was an attempt to learn about the correctness discrimination of second guesses in the multiple-choice recognition paradigm. Additionally, several predictions drawn from the multicomponent theory of memory were tested. At present, theoretical accounts have only emphasized an elimination process as the explanation of second guessing behavior (e.g., Brown, 1965a, Murdock, 1963). It was hoped that the multicomponent model would probe more deeply into the dynamics of the elimination process itself.

As indicated earlier, when response terms are specifiable in terms of the organization and inter-relatedness of component attributes, it may be possible to discover factors both in correctness discriminability and second guessing behaviors. In general, where response terms are more structured, more information will be found both in first choice errors and in second choices. Additionally, it was predicted that factors which lead to response structure also lead to increased correctness discrimination.

Subjects

The 28 *S*s were students from the source described earlier. Assignment of the *S*s to the response similarity conditions, dimensional stimuli (DS) and nondimensional stimuli (NDS), was counterbalanced on the basis of order of appearnce.

Instrumentation

Figure 2 shows the face plate of one of six electroluminescent panels used. The panels measured 3" x 4" and were displayed in a vertical column with a 2" separation between panels. A wide versatility of displays was possible with the equipment; individual elements of each panel could be operated independently and combinations of elements could also be used. The programming of patterns as well as duration intervals was achieved through the use of a Wang Block Tape Reader; paper tapes used to control the reader were punched by a Friden Flexowriter.

Task

The general design was a modification of the "probe" paired-associate technique (e.g., Murdock, 1966) where the stimulus terms were represented by the six panels. The probe in this case was the dot (r) in the lower right hand corner of each panel. A probe lasted 18 sec. and was the S's cue to recall the pattern (single element) which had been displayed on that trial. A trial consisted of lighting one of four elements (b, f, j, or n for the DS condition and c, e, k, or m for NDS) on one panel for 2 sec., and doing the same for each other panel in a random sequence. In a sense, each panel could be in one of four possible "states" and these states were shown to the S one at a time. It was the

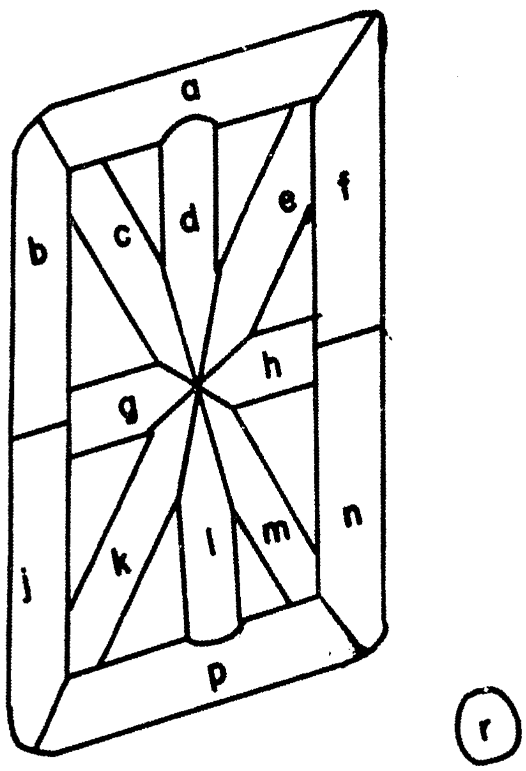


Figure 2. Face plate of an electro-luminescent panel

S's task to remember these states as accurately as possible. Following presentation, one panel was probed. The dot appeared for 25 sec. on the panel whose state was to be reported.

Lists

For the present task it is possible to define serial position in two ways, namely, in terms of the temporal order of inputs or in terms of the spatial arrangement of stimuli. To balance the inputs a modification of the latin square design was used. A 4 x 4 latin square with two rows and two columns added was used in two replications. This allowed a desirable balance wherein each of the four response terms was represented as the tested item three times in each temporal order and three times in each spatial position.

Having thus fixed the temporal and spatial position of each probe as well as which response would be correct for each of these, the remaining five inputs of each trial were determined. The five unused temporal and spatial positions were randomly permuted to exhaust each category. Also, response alternatives were assigned with the following restrictions: No response term could represent more than two input states within a single trial, and each response term appeared an equal number of times in each temporal position. In addition to the 72 experimental trial lists described above, 21 other trials were drawn up. Of these, 15 were practice trials used to acquaint the Ss to the task. There were also three warm up trials which preceded each experimental block of 36 trials. These additional trial lists were representative of the actual experimental trials used.

Paper punch tape was prepared coding the above information onto four separate blocks of paper. Tape 1 contained three trials which were presented with no testing during the reading of instructions. Tape 2 contained the 18 practice trials, whereas tapes 3 and 4 contained three warm up and 36 experimental trials each.

Procedure

The Ss were tested individually in sessions lasting approximately an hour. A 10 minute rest interval was provided between the two major trial blocks.

The two experimental conditions of stimulus similarity, DS and NDS, were determined by the four elements chosen as response terms. Elements b, f, j, and n correspond to the DS condition, whereas elements c, e, k, and m represent the NDS condition (see Figure 2). The shift in similarity condition was made by rewiring the equipment rather than by making additional paper tapes.

To begin, the S was shown a card depicting the four elements appropriate to his experimental condition; these were labeled a, b, c, and d, respective of the above orders. Also, the dot was shown on the card and was labeled as a "query or test of memory." Next, Ss were told to merely watch the three practice presentations "to get an idea of what it is you'll be asked to remember." At this point a trial was defined for the S; also, he was informed of the nature of the probe test of memory. The S was then instructed in the second guess and confidence procedure, and when this was understood the practice trials began. Following the first recall attempts the S was given feedback concerning the accuracy of his responses. If the first response was correct he was asked to merely wait for the next trial. However, if the first response was an error he was asked to make a second choice from among the remaining three alternatives. Also the S was asked to rate each second choice as to how certain he was that he was correct. Confidence ratings were indicated by a verbal declaration of "right" or "wrong" by the S. No feedback was given concerning second choices.

The E recorded all responses on an answer sheet (IBM document 505). Rather than blackening spaces on the sheet, E instead wrote a 1 for the first answer and a 2 for the second. Beside each second answer the E recorded R (right) or W (wrong) corresponding to S's two-point confidence rating for that response.

Results and Discussion

Memory curves

Figure 3 shows memory performance both as a function of the temporal order of inputs and in terms of the spatial arrangement of stimuli. The temporal serial position curve is characterized by a marked recency effect and only a slight primacy effect, whereas the analysis of spatial positions showed slight superiority for the top and bottom positions. Overall probability correct was .60.

Tests of the multicomponent model

To begin, only trials on which an error occurred were analyzed. Additionally, this is the case for each of the statistical tests to follow.

First, it was hypothesized that errors depend on correct responses, and that this dependency should be greater for the DS than for the NDS condition. To test this, errors were recorded in a "confusion matrix" (separately for each condition). These matrices are shown in Figure 4. Here, *Ss'* responses were entered as a function of the correct response. Under null conditions there is no relationship between correct responses and the errors which are given. That is, the entries of each matrix are entirely reproducible from the marginal proportions. In effect, we have a 3 x 4 Chi-square test

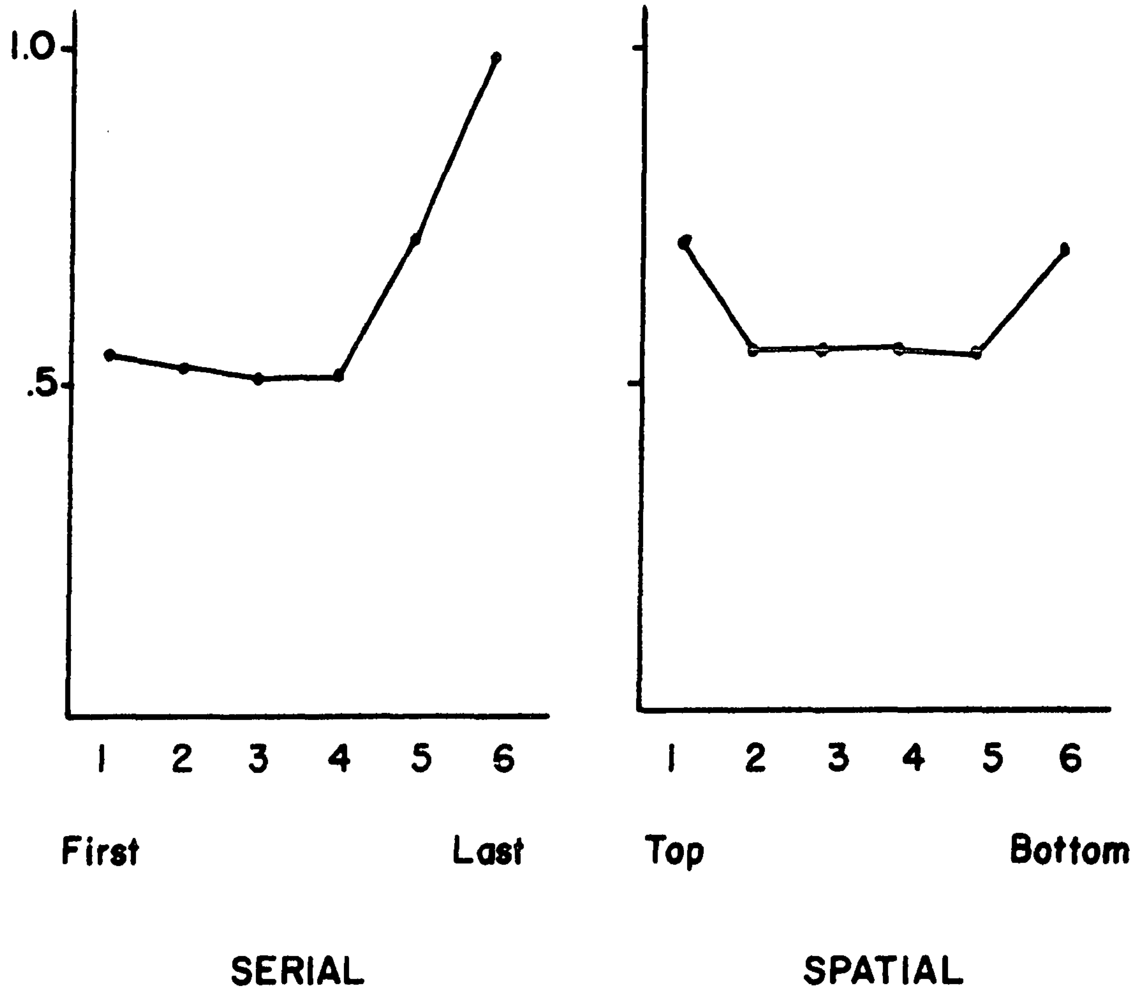


Figure 3. Probability correct for serial and spatial positions

(6 df.) where a significant value of this statistic indicates a rejection of the above stated null condition. The Chi-square values for the DS and NDS conditions were 28.24 ($p < .01$) and 15.56 ($p < .05$). There is evidence for the predicted interaction with similarity conditions ($p < .01$, cf. Knepp and Entwisle, 1969).

A second method of investigating possible memory representations in terms of dimensions is to note the common dimension, horizontal (H), vertical (V), or diagonal (D), between correct responses and errors. The proportion of each of these types of errors was as follows: DS condition; H=26%, V=44%, D=30%; NDS condition; H=26%, V=47%, D=27%. Thus in both similarity conditions there was a strong tendency to make errors which bore a common horizontal attribute (right or left) with the correct response. In other words, errors tended to be above or below the target item.

Next, second guesses were predicted to be above chance accuracy and again that the effect interacts with similarity conditions. Chance performance on second choices in the present task is one-third correct. The proportions of second guesses correct for the DS and NDS conditions were .54 ($Z=6.68$, $p < .01$) and .45 ($Z=3.99$, $p < .01$). As expected, the difference in these proportions was also significant ($Z=2.64$, $p < .01$).

Dimensional Stimuli

errors

correct response	a	b	c	d
a	-	14	33	23
b	14	-	18	27
c	34	9	-	12
d	21	29	30	-

Non-dimensional Stimuli

errors

correct response	a	b	c	d
a	-	26	32	23
b	15	-	13	26
c	28	19	-	14
d	22	26	13	-

Figure 4. Confusion matrices showing the relationship between errors and correct responses

In the next analysis the dependency of second choices on first choices was of interest. According to the multicomponent view, the reason second choices relate to first choices is because of common encoding attributes. In the present experiment Ss may have emphasized the encoding of items on the basis of the horizontal dimension, devoting less time to the processing of other aspects of the stimuli. Thus, first and second responses would tend to share attributes (right or left) on this dimension. It follows that the most likely first-second choice combination would be to pick elements on the same side of the display panel. This effect was predicted to be less under the NDS similarity condition where the right-left dimension is less vividly portrayed.

To test this hypothesis a second set of matrices was developed and these are shown in Figure 5. Under null conditions no contingency should be found between first and second responses. Again the Chi-square test was used and as before both DS and NDS conditions showed a significant contingency relationship; Chi-square =27.44 and 26.81, respectively. The patterns are highly similar to those found between correct responses and first choice errors. In this case, however, there was no evidence for the predicted interaction.

Last, it was hypothesized that confidence ratings depend on first choices. In terms of the importance of the horizontal

Dimensional Stimuli

first responses	second responses: a	b	c	d
a	-	16	38	15
b	14	-	10	28
c	36	24	-	21
d	19	22	21	-

Non-dimensional Stimuli

first responses	second responses: a	b	c	d
a	-	20	31	14
b	16	-	24	31
c	28	14	-	16
d	15	33	15	-

Figure 5. Matrices showing the relationship between first and second responses

attribute discussed in the preceding analyses, when a first choice error contains this memory attribute in common with the correct response, the certainty rating of the second response should be greater. To test this hypothesis, errors were classified on the basis of the common dimension (H, V, or D) between the error and the correct response. The confidence ratings of second guesses associated with these classifications were subjected to an analysis of variance. The effects of the similarity treatment were also included in the analysis, as was the interaction of these treatments with the classifications. The result is shown in Figure 6; none of the factors were significant. A specific linear comparison of horizontal versus nonhorizontal memory showed the contrast to be of marginal importance in confidence ratings ($F=3.19$, not significant). It does not appear that the retention of the major memory dimension leads to significantly higher confidence in second guessing ability.

Signal detection analysis

It is already established that *S*s can discriminate the correctness of second guess responses (Brown, 1965a). Here the question is whether this ability is systematically affected by the dimensionality of the stimulus array.

An attempt was made to estimate discrimination parameters for individual *S*s. The attempt was somewhat successful, but

Source	df	Mean Square	F
Treatments (T)	1	144.98	<1.00
Subjects (S)/T	16	666.76	
Classifications (C)	2	258.20	1.64 (n.s.)
T x C	2	11.99	<1.00
S/T x C	32	157.30	

Figure 6. Analysis of variance

one difficulty was encountered. Certain Ss achieved hit or false alarm rates of 1.00 or 0.00, and in such cases d' is not defined. The elimination of these Ss left eight d' values in the DS condition and six in the NDS condition. The d' averages for these groups were, respectively, .32 and .41. Thus, while each of these values is significantly greater than zero (t-tests), there is, of course, no evidence that Ss in the DS group were better in correctness discrimination than those in the NDS condition.

By way of summary, the main effect of stimulus dimensionality in the present investigation has not been a powerful one; in only two of the five analyses did this factor significantly change the test outcome. A clue to this result is given in terms of the pattern of errors observed; nearly half of these represented a generically correct response to the appropriate side of the display, under both conditions of stimulus dimensionality. Furthermore, it seems highly likely that the vertical column of display panels was responsible. Whether the ultimate explanation shall be given in terms of relevant eye movements, interference effects for the coding of vertical and diagonal values, or some other factor, unfortunately, cannot be specified for the present experiment. However, a few cautious conclusions can be drawn. First, Ss appear to have been sufficiently challenged by the task. Probability correct was .60, where .25

represents chance performance. Next, the differential forgetting of the potential encoding dimensions points to a systematic attempt on the part of the Ss to reduce their cognitive strain in the task. If the dimensional representation of items in memory did occur, the effect was the same for both experimental treatments; objectively, the important feature of the stimulus was its right or left position on the instrument panel.

GENERAL DISCUSSION

Several things have been learned from the present investigation. In Study I, a first attempt was made in explaining the "feeling of knowing" phenomenon in terms of part-whole processes. Tulving (1970) has implied that memory's "knowledge of its own contents" reflects some special talent held only by humans; the present data, however, do not demand this interpretation. "Feeling of knowing" judgements contained no information about future recognition performance which was not already contained in partial recalls. Thus, as an alternative to the dual process view, it is felt that Ss based their judgements directly on the parts of a response which were remembered.

However, while both interpretations are adequate to the facts of feeling of knowing experiments, the present study does not distinguish the validity of either position. It is a weak argument to counter the dual process conception as lacking in parsimony; research is needed to resolve the issue. The simulation of recognition on the basis of partial recalls was an attempt to provide this. In this case, the information in partial recalls was not sufficient to predict all of recognition. Dual process theories would perhaps interpret this as support; however, the present account has suggested that the residual information resides in the S's knowledge of validity differences in the parts remembered. Despite the ambiguity, it was suggested that the simulation technique might prove useful in resolving the issue.

In the last study it was hoped that additional tests of the multicomponent model could be made. However, the manipulation of structure among the elements of the six display panels was completely outweighed by dimensional preferences which seemed to arise out of the gross structure of the task. When the latter preferences were analyzed in terms of the multicomponent view of memory, some support was found for the notion that the confidence ratings of second guesses in the multiple-choice task depended on partial recall. When the dominant "part" (right versus left) was remembered on the first choice, the confidence ratings of second guesses were higher than when the primary part was not remembered. As in Study I, it appears that "knowing that you know" may not be an additional process beyond merely "knowing part." If anything, the process of correctness discrimination is a logical one rather than memorial; individuals remember a part, and when asked about certainty, their judgements are based on the number of parts recalled.

LITERATURE CITED

Achilles, E. M. Experimental studies in recall and recognition. Archives of Psychology, 1920, 6, 1-80.

Adams, J. A. Human memory. New York: McGraw-Hill, 1967.

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

Bower, G. H., and Bostrum, A. Absence of within-list PI and RI in short-term recognition memory. Psychonomic Science, 1968, 10, 211-212.

Bregman, A. S. Is recognition memory all-or-none? Journal of Verbal Learning and Verbal Behavior, 1966, 5, 1-6.

Brown, J. A comparison of recognition and recall by a multiple response method. Journal of Verbal Learning and Verbal Behavior, 1965, 4, 401-408. (a).

Brown, J. Multiple response evaluation of discrimination. British Journal of Mathematical and Statistical Psychology, 1965, 18, 125-137. (b).

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

Cofer, C. N., Bruce, D. R., and Reicher, G. M. Clustering in free recall as a function of certain methodological variations. Journal of Experimental Psychology, 1966, 71, 858-866.

Dale, H. C. A., and Baddeley, A. D. Alternatives in testing recognition memory. Nature, 1962, 196, 93-94.

DaPolito, F., Guttenplan, H., and Steinitz, H. Accuracy of subjective judgements of information in long-term memory. Psychonomic Science, 1968, 13, 227-228.

Estes, W. K., and DaPolito, F. Independent variation of information storage and retrieval processes in paired-associate learning. Journal of Experimental Psychology, 1967, 75, 18-26.

Freud, S. Psychopathology of everyday life. New York: Mentor, 1954.

Gourevitch, V., and Galanter, E. A significance test for one parameter isosensitivity functions. Psychometrika, 1967, 32, 25-33.

Hart, J. T. Memory and the feeling-of-knowing experience. Journal of Educational Psychology, 1965, 56, 208-216.

Hart, J. T. A methodological note on the feeling-of-knowing experiments. Journal of Educational Psychology, 1966, 57, 347-349.

Hart, J. T. Memory and the memory-monitoring process. Journal of Verbal Learning and Verbal Behavior, 1967, 6, 685-691.

Hochhaus, L. Correct response discrimination as a function of multiple recognition choices: The effect of correct guessing on Type II d'. Journal of Experimental Psychology, 1970, in press.

Hollingworth, H. C. Characteristic differences between recall and recognition. American Journal of Psychology, 1913, 24, 532-544.

James, W. The principles of psychology. New York: Dover, 1950.

Kintsch, W. Recognition and free recall of organized lists. Journal of Experimental Psychology, 1968, 78, 481-487.

Kintsch, W. Models for free recall and recognition. In D. A. Normal (Ed.), Models of human memory. New York: Academic Press, 1970.

Knepp, D. L., and Entwisle, D. R. Testing significance of differences between two Chi-squares. Psychometrika, 1969, 34, 331-333.

Luh, C. W. The conditions of retention. Psychological Monographs, 1922, 31 (Whole No. 142).

MacDougall, R. Recognition and recall. Journal of Philosophical Psychology and Scientific Methods, 1904, 1, 229-233.

McKinney, F. Retroactive inhibition and recognition memory. Journal of Experimental Psychology, 1935, 18, 585-598.

McKinney, F., and McGeoch, J. A. The character and content of transfer in retroactive inhibition: Disparate serial lists. American Journal of Psychology, 1935, 47, 409-423.

McNulty, J. A. An analysis of recall and recognition processes in verbal learning. Journal of Verbal Learning and Verbal Behavior, 1965, 4, 430-436.

Martin, E. Stimulus recognition in aural paired-associate learning. Journal of Verbal Learning and Verbal Behavior, 1967, 6, 272-276.

Melton, A. W. Decision processes in retrieval from memory: Comment on the contributions of Lloyd R. Peterson and Charles N. Cofer. In B. Kleinmuntz (Ed.), Concepts and the structure of memory. New York: Wiley, 1967.

Mowrer, O. H. Learning theory and the symbolic processes. New York: Wiley, 1960.

Murdock, B. B., Jr. An analysis of the recognition process. In C. N. Cofer and B. S. Musgrave (Eds.), Verbal behavior and learning. New York: McGraw-Hill, 1963.

Murdock, B. B., Jr. The criterion problem in short-term memory. Journal of Experimental Psychology, 1966, 72, 317-324.

Murdock, B. B., Jr. Short-term memory for associations. In D. A. Norman (Ed.), Models of human memory. New York: Academic Press, 1970.

Norman, D. A., and Rumelhart, D. E. A system for perception and memory. In D. A. Norman (Ed.), Models of human memory. New York: Academic Press, 1970.

Pavio, A., Yuille, J. C., and Madigan, S. Concretness, imagery, and meaningfulness values for 925 nouns. Journal of Experimental Psychology, 1968, 76 (1, Pt. 2).

Peixotto, H. F. Proactive inhibition in the recognition of nonsense syllables. Journal of Experimental Psychology, 1947, 37, 81-91.

Postman, L. Retroactive inhibition in recall and recognition. Journal of Experimental Psychology, 1952, 44, 165-169.

Postman, L., Adams, P. A., and Phillips, L. W. Studies in incidental learning: II. The effects of association val value and of the method of testing. Journal of Experimental Psychology, 1955, 49, 1-10.

Postman, L., Jenkins, W. O., and Postman, D. L. An experimental comparison of recall and recognition. American Journal of Psychology, 1948, 61, 511-519.

Postman, L., and Stark, K. The role of response availability in transfer and interference. Journal of Experimental Psychology, 1969, 79, 168-177.

Shepard, R. N. Recognition memory for words, sentences, and pictures. Journal of Verbal Learning and Verbal Behavior, 1967, 6, 156-163.

Swets, J. A., Tanner, W. P., and Birdsall, T. G. Decision processes in perception. Psychological Review, 1961, 68, 301-340.

Tulving, E., and Madigan, S. A. Memory and verbal leaning. Annual Review of Psychology, 1970, 21, 437-484.

Underwood, B. J. Attributes of memory. Psychological Review, 1969, 76, 559-573.

Underwood, B. J., and Schulz, R. W. Meaningfulness and verbal learning. Philadelphia: Lippencott, 1960.

Walker, H. M., and Lev, J. Statistical inference. New York: Holt Rinehart and Winston, 1953.

Yavuz, H. S., and Bousfield, W. A. Recall of connotative meaning. Psychological Reports, 1959, 5, 319-320.

ACKNOWLEDGMENTS

It is a pleasure to acknowledge my debt to Dr. Wayne H. Bartz, Chairman of my Graduate Committee; my opinion is that his enthusiasm and encouragement in research pursuits can be matched by very few men. I also wish to thank Drs. Donald C. Charles, Terry L. Dickinson, David C. Edwards, and Richard C. Warren of my Graduate Committee for the valuable time and service they have directed to this work.

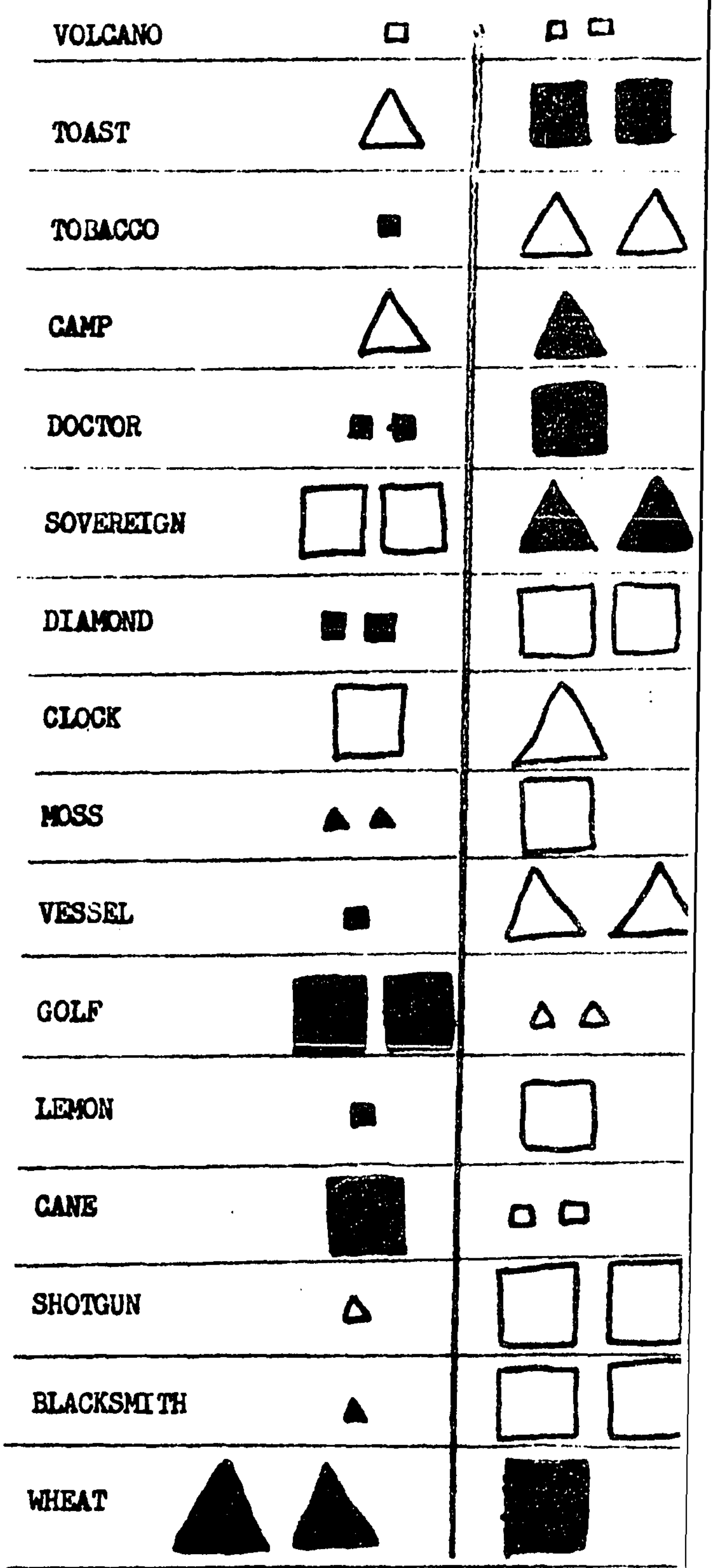
Others, though they held no designated status regarding my research efforts, are highly deserving of mention. Thus, a special thanks to Harold Gelfand and Glen Raser who spent many long hours with me in analytical discussion of design problems and critical concepts concerning the investigation of the ecphoric processes covered by this thesis. Also, other members of the Bartz Research Team have served in this capacity as well; they are, Richard Anderson, Marion and Robert Lewis, Mary Salehi, Gary Schumacher, William Stock, and Gene Swinton.

Above all, I wish to express my gratitude and appreciation to my wife, Jean, for her understanding, encouragement, and patience during the writing of this thesis.

APPENDIX

Name________________________ 51b

LEMON	1 2	L Sm	B W	Sq T
MOSS	1 2	L Sm	B W	Sq T
DIAMOND	1 2	L Sm	B W	Sq T
GOLF	1 2	L Sm	B W	Sq T
CLOCK	1 2	L Sm	B W	Sq T
SOVEREIGN	1 2	L Sm	B W	Sq T
CAMP	1 2	L Sm	B W	Sq T
TOAST	1 2	L Sm	B W	Sq T
VOLCANO	1 2	L Sm	B W	Sq T
TOBACCO	1 2	L Sm	B W	Sq T
VESSEL	1 2	L Sm	B W	Sq T
DOCTOR	1 2	L Sm	B W	Sq T
WHEAT	1 2	L Sm	B W	Sq T
BLACKSMITH	1 2	L Sm	B W	Sq T
CANE	1 2	L Sm	B W	Sq T
SHOTGUN	1 2	L Sm	B W	Sq T



APPENDIX B

The following instructions were read to each S:

"This is a learning experiment in which you will learn to associate words and geometric figures. It is very important that you follow the instructions carefully. Should you fail to follow any instruction, be sure to tell me since this would be important to the interpretation of the results."

"The list will consist of 16 pairs of items like the pair on this card. (The E shows an example card.) These pairs will be shown to you one at a time above this partition; you will see each pair for about 2 sec. You are to study each pair such that later when you are shown the words alone you can recall the figure that was paired with each word earlier."

"Now, let me show you some more example figures.... The figures we will be dealing with were constructed by a few simple rules: There are only two possible colors, white or black; only two possible sizes, small or large; and only two possible shapes, square or triangle. Last, the number of figures paired with the stimulus word may vary. In some cases a single figure will be paired with the word, while sometimes the figure will be presented doubly, that is, merely repeated on the same card." (During the discussion of figures E shows example response terms for each dimension as these are described.)

"After you have studied the full list of 16 word-figure pairs you will be given a sheet of paper containing the 16 words. As you look at each word try to recall the figure that was paired with it earlier. Next, you are to make your answer by marking a set of symbols which you'll find to the side of each word. These symbols are arranged in 4 pairs corresponding to the dimensions I have just described. Your first choice will depend on whether the figure was single or double, and here you are to circle either a 1 or a 2. Next is size, and you circle either large or small; next color, black or white; and finally, shape, either square or triangle. For example, if your answer was 2 large black squares you would circle the 2, then large, then black, and then square. Do you understand? If you are unsure about any of these dimensions do not leave any answer out. Go ahead and guess since this will not be counted against you and it is important to have a complete record of your learning."

After the study trials the instructions were extended as follows:

"Here is the list of words; work from the top of the page down and do not omit any answers. However, before you begin I have a further instruction. In the left hand margin you are to put down a number from 1 to 6 to tell me how sure or confident you are that you could recognize the correct answer on a multiple-choice recognition test. Use the number 6 if you are very certain that you know the answer and use the number 1 if you are not at all sure. Additionally, you may use any whole number in between to indicate greater or less confidence. Remember, these numbers should reflect your feeling that you could recognize the correct answer... regardless of whether you recalled the answer or not. When you have finished do not turn you paper over; just put your pencil down and look up. Do you understand? Go ahead."

Upon completion of the recall task the recognition test was administered with the following instructions:

"Now I want you to take a multiple-choice recognition test. When I say 'ready' I want you to turn your paper over and again work through the list of words from the top to the bottom of the sheet. Beside each word you'll find two response terms separated by a thickened line. Make a small check mark by the figure which you feel is correct. Do you understand? Ready? Go."