HALL, DONALD MADISON. The Learning and Recognition of Random Shapes: Effects of Number of Response Categories and Retrieval Cues. (1974)
Directed by: Dr. H. Herbert Wells

Two types of learning task, as well as the presence or absence of retrieval cues, were examined for their effects on long-term recognition memory for random shapes. Twenty college student subjects learned to associate the number "1" with five of ten 24-point random shapes, as used by Vanderplas & Garvin (1959), and a "2" with the other five shapes (Group 2). Twenty other subjects learned to associate a unique number, from one to ten, with each of the same ten shapes (Group 10). One week after the subjects had learned the shapes (to a criterion of three successive correct trials) a recognition test for the shapes was given. The recognition test was given with two cards, each containing fifteen 24-point random shapes (five "old" shapes and ten distractors). On one set of cards all "old" shapes on each card had the same response number, while on another set of cards the shapes were randomly assigned to each card. Half of the subjects in Group 2 were given the former type of cards (categorized) while the other half of the subjects in Group 2 used the latter type of cards (uncategorized). Subjects receiving the categorized cards were informed of the categorization. Half of the subjects in Group 10 received a set of cards on which the "old" shapes were arranged in serial order

according to the numbers associated with the shapes (categorized) while the other subjects in Group 10 received cards containing "old" shapes which had been ordered randomly (uncategorized).

The recognition scores were very high for all groups. Group 2 and Group 10 subjects who received categorized tests differed significantly in recognition performance (the former with 88% and the latter with 98% recognition performance). Analysis of learning performance indicated that Group 2 and Group 10 did not differ in learning speed, though chance performance should have enabled Group 2 subjects to learn faster. A processing limit of 2±1 shapes per trial appeared to have prevented Group 2 subjects from being able to use their advantage.

The random shapes used were scaled for difficulty, but the various difficulty indices did not correlate highly with the measures of association value used by Vanderplas & Garvin (1959). The experimental paradigm was found to be less than optimal for the study of recognition memory, primarily because of ceiling effects that could not be eliminated.

THE LEARNING AND RECOGNITION OF RANDOM SHAPES: EFFECTS OF NUMBER OF RESPONSE CATEGORIES AND RETRIEVAL CUES

by

Donald Madison Hall

A Thesis Submitted to
the Faculty of the Graduate School at
The University of North Carolina at Greensboro
in Partial Fulfillment
of the Requirements for the Degree
Master of Arts

Greensboro
1974

Dec. 19, 19[illegible]
Date of Acceptance

Approved by

Herbert Wells
Thesis Adviser

APPROVAL PAGE

This thesis has been approved by the following committee of the Faculty of the Graduate School at The University of North Carolina at Greensboro.

Thesis Adviser Herbert Wells

Committee Members Ernest A. Lyonsden

Robert W. Pratt

Dec. 19, 1974
Date of Acceptance by Committee

ACKNOWLEDGMENTS

The author is indebted to Dr. H. Herbert Wells, adviser, for his guidance of this research and the use of his laboratory. Sincere thanks are also extended to Drs. Ernest Lumsden and Robin Pratt, members of the thesis committee, for their advice and assistance.

Special thanks are due to Dr. Charles L. Richman, Wake Forest University, for his initial suggestions and continued interest in this research, and to Miriam Freel, for her assistance with experimental equipment and manuscript preparation.

TABLE OF CONTENTS

	Page
ACKNOWLEDGMENTS	iii
LIST OF TABLES	v
LIST OF FIGURES	vi
CHAPTER	
I. INTRODUCTION	1
II. METHOD	21
Subjects	21
Apparatus	21
Procedure	22
III. RESULTS	27
IV. DISCUSSION	54
BIBLIOGRAPHY	59

LIST OF TABLES

Page

TABLE I 2x2 Analysis of Variance for Recognition Scores 45

LIST OF FIGURES

		Page
FIGURE 1	Sample Recognition Test Form	23
FIGURE 2	Learning Curves for Each Shape for a Typical Subject in Group 2	29
FIGURE 3	Adjusted Learning Curves for the 20 Subjects in Group 2	30
FIGURE 4	Adjusted Learning Curves for the 20 Subjects in Group 10	31
FIGURE 5	Illustration of Derivation of the "spread index"	34
FIGURE 6	Scatter Plot for Each Group Showing Correlation between Trials to Criterion and the Spread Index	36
FIGURE 7	Illustration of the "quasi Vincentizing" Procedure	38
FIGURE 8	"Vincentized" Learning Curves for Each Group, Calculated from the Adjusted Learning Curves	39
FIGURE 9	Conditional Probabilities for the Two Groups	42
FIGURE 10	Recognition Performance for the Four Recognition Groups	44
FIGURE 11	Rank Order Correlation Matrix for Shape Difficulty Measures and "Association" Scales	52

CHAPTER 1

INTRODUCTION

Much attention has been focused recently on the effects of organization on learning and memory. The impetus for this renewed interest in grouping processes was given, for the most part, by Miller's (1956) article concerning chunking. He reviewed evidence that suggested a limit on processing capacity which allowed input of only 7±2 chunks of information at a given time. Memory was increased, he argued, only by increasing the number of items in each chunk.

A somewhat earlier treatment of grouping effects (Katona, 1940) gave a more detailed analysis of the process. "Artificial grouping" and "grouping by arrangement" were the two types of organization that occurred, with each type having its own rules. For artificial grouping, in which the subject arbitrarily groups the material to be learned, Katona listed the following rules: "(1) Between members of the same group there is a stronger association than between members of different groups, (2) part of a group has a tendency to reproduce the entire group, (3) groups have their own associations, which may be different from the associations of their members, and (4) grouping facilitates learning" (pp. 168-169). Furthermore, once groups were established, any

later situation which entailed re-grouping tended to subtract from the positive influence of prior use of the same materials. Recent evidence on this point appeared in a study by Postman (1971) in which organization developed during a free recall task interfered with later learning of a paired-associates task that paired the same words arbitrarily.

Grouping according to an arrangement referred to situations where the grouping was obvious due to the nature of the material or was urged by the experimenter. This type of grouping tended to proceed faster than artificial grouping, and the boundaries between groups were stronger. These groups from an arrangement were more difficult for the subject to alter or discard than groups formed artificially.

Mandler built upon the ideas of Katona and Miller in his theories of chunking. Mandler (1962) argued that functional response units were formed during overlearning, after errors had dropped out, from what were previously discrete responses. "The whole sequence is elicited as a unit and behaves as a single component response has in the past; any part of it elicits the whole sequence" (p. 417).

Mandler (1967) further argued that organization is necessary for learning to occur, that memory for verbal

materials is hierarchical (words are organized into successively higher-order categories), that within a category storage capacity is limited, and that the limit on memory is a function of how many hierarchical levels the subject can retain. Mandler (1968), however, set the processing limit at 5±2 chunks per hierarchical level, lower than Miller's (1956) 7±2.

A comprehensive treatment of chunking and recall by Johnson (1970) essentially agreed with Mandler's analysis, except for the substitution of 4±1 chunks as the processing capacity. He referred to memory codes, the "memorial representation of information" (p. 173), which were distinct from the information they represented. A chunk was defined as any response sequence which was represented by a single code. The code contained all the information in the chunk, and recall of the information depended on recovery of the code. These codes could be organized into higher-order units, which also contain all the information subsumed one level below them. Johnson (1970) referred to codes as "opaque containers," stressing that "a coded chunk cannot be considered as a collection of items in memory, but rather as a unitary entity within which item identity is lost" (p. 213). In several experiments, Johnson demonstrated that the probability of a complete omission of a chunk increased when more

chunks were used for a response sequence of fixed length. Adding more items per chunk failed to produce similar complete omissions, indicating that the chunks were recalled in an "all-or-none" fashion.

Tulving (1968a) also conceded organization an important role in learning, and even implied that items must be organized before they can enter secondary memory. He claimed that organization has its effect on retrieval, in the form of a "retrieval plan" which makes items more accessible in memory. Johnson (1970) mentioned that his "opaque code" and Tulving's "retrieval plan" probably referred to the same thing. The code content could be a type of locating system (a "set of ordered addresses") for recalling the items in the chunk.

Postman (1972) basically agreed with the reasoning of Mandler and Tulving, but considered their statements that organization is necessary for learning as too extreme. He referred to such statements as being representative of the "strong principle of limited capacity." A "weak" principle holds that there is, indeed, a limit to the amount of information that can be processed during a specified period of time, but repetition may have effects apart from contributions to organization. Postman (1972) concluded that

> the weak principle of limited capacity constitutes an important argument for the functional utility

> of organizational processes. There is good evidence that unitization can help to overcome the inherent limitations of the memory system, whatever they may be. No compelling evidence has been adduced, however, that the limits are such as to make unitization a necessary, and the only sufficient, condition for the retention of materials exceeding in length the span of immediate memory. Adoption of the strong principle of limited capacity places an unnecessary burden of proof on the organization theorist (p. 10).

Both Tulving's retrieval plans and Johnson's opaque code theories required that the retrieval or code information be established during learning. If so, it would follow that any boost to recall of the code would enhance overall recall. Experimental manipulation of retrieval cues usually indicated better recall when cues were given (Tulving & Pearlstone, 1966; Tulving & Osler, 1968). The cues were either category names or weak associates of the to-be-remembered (TBR) words. Since overall recall of organized material depended on the number of categories recalled, with the number of items remembered per category rather constant regardless of the size of the category (Tulving & Pearlstone, 1966), the category cues made it likely that few categories would be entirely omitted. There appeared to be "two independent retrieval processes, one concerned with the accessibility of higher-order memory units, the other with accessibility of items within higher-order memory units" (Tulving & Pearlstone, 1966). Retrieval of the

higher order units evidently represented the prime chore for memory. Johnson (1970) even argued that the probability of retrieving a memory code was a function of the level of that code in the hierarchy, rather than being a function of the number of items the code represented. He gave some empirical evidence that higher order chunking led to more omissions of material in recall than lower order chunking of the same material (1970).

Lewis (1971) found facilitation with cueing by exemplars of categories when block presentation of items was employed, but not with random presentation. It seemed that blocked presentation of items led subjects to organize the material such that the cues given fit in with their organization scheme. The Tulving & Osler (1968) data were consistent with such an explanation, since it was found that cues had to be present at input and retrieval to be effective, while cues only at retrieval failed to be facilitative. In the Lewis (1971) study, random presentation of items produced equal recall for cued and noncued groups, though cued groups recalled more categories, but fewer items per category, and noncued groups recalled fewer categories, but more items per category. Thus, cues did not aid and even disrupted groups given random presentation, since the initial storage was not amenable to the cues given.

The evidence for facilitation of recall by giving cues indicated that items could be stored in memory even though they were not "recallable." Tulving & Pearlstone (1966) referred to such items as being "available." Recallable items were called "accessible" items. Such a distinction for recall tests seems quite reasonable and helpful. Does recognition memory operate the same way? Is there retrieval in recognition memory? If so, does organization affect recognition memory? Convincing answers to these questions are not yet available (much less, accessible), though some evidence has begun to give some cues.

Until recently, the prevailing idea was that the recognition process required no retrieval phase, and was consequently not affected by organizational factors (Bartlett, 1932; Kintsch, 1968; McCormack, 1972). Support for such a notion came from both experimental evidence and intuitive appeal. The superiority of recognition over recall scores, which was almost always found (Postman & Rau, 1957) suggested the existence of some process necessary for recall, but not for recognition. This process was assumed to be retrieval. Kintsch (1970) believed that

> in recognition, the problem of retrieval is simple: the item is sensorily present and it is a simple matter to retrieve its corresponding representation

> in memory (although how this is done is by no means obvious); the subject then has some means of judging the newness of the trace (response strength, familiarity); if the newness satisfies some criterion, the subject says he recognizes the item; otherwise, he calls it new; irrelevant alternatives are not considered in this judgment (p. 337).

McNulty (1966) and Murdock (1963) believed that in some types of recognition test alternatives were not only relevant, but the primary foundation of the subject's strategy. In a recognition test the subject eliminates incorrect alternatives and then selects from the remaining items in a somewhat random fashion. McNulty (1966) also mentioned other possible explanations for the generally found superiority of recognition over recall: (1) There are usually fewer alternatives from which to choose in the recognition test, (2) In a recognition test a decision can be made on the basis of "partial" learning of stimulus items, while partial learning is less effective for recall.

Bahrick & Bahrick (1964), however, found that they could obtain scores for recognition tests which were either higher or lower than scores for recall tests. The subjects' task was to remember the position of a dot in a complex matrix. The experimenters contended that previous studies of various measures of retention (Postman and Rau, 1957; Luh, 1922) had used tests which were too easy, allowing simple differentiation of correct

and incorrect alternatives. Thus, recognition and recall may be more similar than they were previously believed to be.

The strongest case for effects of organization on recognition was presented by Mandler (1972). Putting the problem in historical perspective, he said that

> five years ago the problem of recognition and organization seemed to be relatively simple and we entered on what we thought would be a series of minor experiments to demonstrate what was then conventional wisdom and what is still accepted as a plausible position (e.g., Kintsch, 1970; Murdock, 1968). Specifically it is assumed that while recall depends on the retrieval of items by means of organization of storage, recognition does not. We believed then, as apparently everybody else did too, that subjects would not use retrieval systems during recognition, whether the task presented them with organized or with unstructured lists of items. What was presumably the case was that we were dealing with two distinct processes in memory -- a retrieval process and a decision process. Recognition depended -- by that argument -- solely on a simple decision process unaffected by organizational variables (1972, pp. 140-141).

Mandler's search for a simple way to distinguish recall and recognition led to the realization that a simple distinction was not possible (Mandler, Pearlstone, & Koopmans, 1969). Recognition was greater for more organized material (with organization defined as the number of categories used in a sorting task). It was suggested that studies not finding effects of organization often used insufficient degrees of organization for the effects to occur.

After reviewing the evidence and conducting several experiments on recognition and organization, Mandler concluded that

> the present data suggest that subjects may first check occurrence tags and then use organizational tests on those items that have weak tags. The more such uncertain items are generated, the better the correlation should be between the organization and recognition. This particular analysis of the situation also explains a previously puzzling finding, namely that the correlation between recognition and organization was higher after a two-week delay than during an immediate recognition test. List tags presumably decay markedly over two weeks and more items are thrown into the category of weak or uncertain items (1972, pp. 152-153).

Mandler (1972) mentioned several studies which supported the notion that organization affects recognition. Lachman & Tuttle (1965), using syntax as the basis for a high degree of organization, found better recognition for organized than for unorganized items. Using random and hierarchically organized lists, Bower, Clark, Lesgold, & Winzenz (1969) found recognition scores higher for groups using organized lists. Bruce & Fagan (1970) varied the number of categories in a 42 word list. Recognition was better for groups using 6 categories than for groups using 42 categories.

Other evidence concerning organization and recognition has come from studies employing random and blocked presentation of items prior to tests of memory. Kintsch (1968) found no facilitation of recognition performance

after blocked presentation (relative to random presentation) but his experiments have been criticized by Postman (1972) and by D'Agostino (1969). D'Agostino replicated Kintsch's (1968) experiment, using better controls, and found a slightly significant effect of blocked presentation.

Jacoby (1972) determined that the effect of type of presentation depended somewhat on the order of items on the recognition test. In his first experiment, there was a significant interaction between the type of training (blocked vs. random) and the type of test (categorized-ordered vs. categorized unordered vs. uncategorized). Thus, the facilitation due to the ordering of items on the recognition test in a manner similar to that during initial learning was greater for more highly organized lists (blocked presentation). In the second experiment, blocked presentation was superior to random presentation and a categorized-ordered test was superior to an uncategorized test, but the interaction failed to reach significance. Jacoby (1972) concluded that

> organization variables do have an effect on recognition memory. The conditions under which an effect should be expected, however, have not been completely determined . . . Conflicting results from investigations of organization effects on recognition may be partially due to differences in degree and ease of organization. Degree of organization probably depends on category size, the particular categories employed, the normative frequency of study words as category instances, S's strategy and a number of other unspecified variables (pp. 330-331).

In addition, Jacoby (1972) mentioned that the distance (lag) between related items in a randomized list would affect the ease and, hence, the degree of organization. Jacoby & Hendricks (1973) spaced category instances 0, 1, 3, or 11 items apart and found significantly greater recognition memory for the first two groups over the last two groups. Apparently, the separation of category instances made organization more difficult. Categorized tests yielded significantly higher recognition scores than did uncategorized tests (though with greater spacing, this superiority was reduced or lost completely). The false alarm rate increased with spacing, indicating that some category information was encoded even with greater spacing, though this information was not sufficient for discrimination of "old" and "new" items. The experimenters contended that their results were in accord with Mandler's (1972) suggestion that retrieval is involved in recognition memory only when items are highly organized during study, with low degrees of organization sometimes disrupting recognition performance.

There is some evidence that the context of a test item (either items that accompany it at presentation or the general structure of the list) determines how that item will be stored in memory. Context effects have been shown to be important for recognition memory.

Tulving & Thomson (1971) changed context by adding, deleting, or changing a "context word" (a word that accompanied the test item) during the recognition test. Any such change in the context was found to be detrimental to performance. Evidently, only location of the specific encoding of an item, which was affected (determined) by its context, allowed correct recognition. Thus, differences in context between first presentation of an item and later presentation of the item tended to produce different and independent functional stimuli, while the nominal stimulus (the test item itself) remained unchanged. Tulving & Thompson (1971) believed that information had to be available <u>and</u> accessible for recognition as well as recall.

Light & Carter-Sobell (1970) found higher recognition performance when the meaning of homographs (adjective-noun pairs in which the adjective tends to determine which meaning of the noun is used) was the same during study and test, indicating that context information was stored at input which aided retrieval when it was preserved at output. They argued that proponents of theories of recognition memory who contend that recognition requires only a decision concerning the familiarity of an item should add some mechanism which can specify the particular memory representation that would be examined for recency.

"It is clearly not sufficient to simply state that presentation of a test item obviates the need for retrieval operations by directing S to the memory representation of a test item for purposes of making recency judgments" (Light & Carter-Sobell, 1970, p. 9).

Winograd, Karchmer, & Russell (1971) found a "cueing effect, defined as the facilitation of recognition of a to-be-remembered (TBR) word by the reinstatement of an item which accompanied it at encoding" (p. 199). This cueing effect was found only when subjects were instructed to form a compound image (a "bizarre mental picture combining the words" (p. 201)), and not when they were given associative instructions. Winograd et al. (1971) argued that a "compound image was stored as a functional unit, with each part undoubtedly altered by its incorporation into that whole . . . With associative instructions, we must assume that what is stored is something like separate elements along with information that they go together" (p. 204). Although a cueing effect was found with associative instructions for recall, while not for recognition, it is likely that the difference was due to the amount of "cueing" in the "uncued recognition test." In other words, presentation of the TBR word on the recognition test was more of a cue than none at all, which was the case on the recall test.

In the same vein, Tulving (1968b) used compound words (such as tooth-ache) during learning and tested for recognition only one word of the pair. The decrement in scores found was presumably due to breaking up the encoded unit, an explanation which Winograd et al. (1971) would find agreeable. DaPolito, Barker, & Wiant (1971) studied retrieval processes in recognition memory similarly by altering the context of the item. Semantically related word triplets were presented, then one of the words was presented for recognition either in the original context, with no context (words deleted), with new, related context words which suggested a different meaning for the target word, or with unrelated context words. Using hit rate as a measure of recognition performance, the old context condition produced the highest scores, and the new unrelated context condition the worst. The deleted and new related conditions were intermediate and did not differ significantly from each other. Using a signal detection theory (SDT) application, however, the best groups were the old context and new related context groups, while the worst were the deleted and new unrelated groups. DaPolito et al. (1971) argued that the SDT method was probably more accurate, since it controlled for decision criterion changes. That the new related items did not hinder recognition performance was at odds with the

Light & Carter-Sobell (1970) results. Instead, the results in the DaPolito et al. study suggested that

> during recognition tests, a specific semantic context permits S to retrieve multiple semantic representations for a single word, especially if the representations are strongly coupled in associative memory. Although all internal representations may not be searched, the multiple retrieval hypothesis seems more than a logical possibility, since associative responses to equivocal words depend upon comparisons of different meanings (1970, p. 181).

The deleted and new unrelated groups were, however, hindered on the recognition test, which indicated interference (or lack of facilitation) due to changed context.

Light & Schurr (1973) incorporated the techniques of Tulving & Thomson (1971) and Winograd et al. (1971) in a recent study of context effects in recognition memory. Groups learned a list of unrelated words either by (1) inventing a story which preserved input order for blocks of eight words, (2) studying the list grouped into blocks of eight, or (3) studying the items singly. The "story" group performed better on the recognition test than the other two groups, which did not differ. Performance on a recognition test which preserved input order was better than on a test which had different order. However, the prediction that this superiority would be greatest for the "story" group (since it was presumably the most highly unitized) was not confirmed, thus arguing against the unitization hypothesis (Winograd et al., 1971) which

would predict a greater effect of changed context for more highly unitized materials. Subjects performed better for tests with order similar to input order in both the "story" and the "single" groups, indicating that the "hypothesis that context affects recognition performance may thus be extended from the 2-word context situation investigated by Tulving and Thomson (1971) to the more general case in which the entire list in which a word is embedded is viewed as its context" (Light & Schurr, 1973, p. 138).

The present study examined the effects of organization and cueing on recognition memory for random shapes. Random shapes were used for several reasons. First, they were considered less likely than words or letters to be affected by preexperimental associations, although some of the shapes were expected, upon exposure to the subjects, to evoke consistent associations across subjects. Secondly, interference effects during the retention interval were expected to be minimal, since few subjects were likely to encounter similar random shapes elsewhere during the week. Third, it was hoped that complex random shapes would be difficult to remember for a week and ceiling effects on the recognition test would be avoided. Fourth, the recent technique used by Charles Richman and his associates (Richman & Trinder, 1968; Trinder, Richman,

& Gulkin, 1969; Trinder, Metzger, Sherman, and Richman, 1972; Richman, Bidwell, & Henson, 1973; Richman, 1974) for inducing chunking with random shapes could be used to study the effects of experimentally acquired-organization on recognition memory. Their method of evoking organization with these shapes was used rather uniformly in all of their studies, and will be briefly described. In general, subjects learned to give one response (in some cases a verbal response; in others, a lever press) to four random shapes and another response to four others. After the subjects had been either undertrained or overtrained on the discrimination task, they were given either a reversal or a nonreversal shift. In the reversal shift, all Response "A" shapes became Response "B" shapes, and Response "B" shapes became "A" shapes. Thus all four shapes in each response group were still united by having one response in common, but the responses were now switched between groups. In the nonreversal shift, responses for only two shapes in each response group were changed (two A's and two B's). Thus, the original groups (A and B) were now broken up. It was hypothesized that prior to reaching criterion, subjects responded in an S-R fashion to each individual shape, and had not yet firmly organized the shapes into groups according to common responses. After overtraining, however, two

chunks should have formed, with the shapes having the same label forming the chunks. Evidence for chunking was obtained by comparing the two types of shift. As expected, the undertrained group performed better on the nonreversal shift, since subjects responding on an S-R basis for each individual shape would benefit from the smaller number of response changes in the nonreversal shift (4 as compared to 8 in the reversal). With overtraining, however, subjects performed better on the reversal shift, indicating that the chunks were well integrated and that all the subjects had to do was to switch chunk labels. On the nonreversal shift, the chunks were disrupted and relearning was not facilitated by having the items highly organized.

Given that a common response can lead to grouping of random shapes, the effects of such grouping on memory can be assessed. Predictions concerning the effects of chunking on memory for random shapes should parallel those for verbal materials. Specifically, if the number of categories (chunks of random shapes, as defined by having a common response) is varied for a constant number of shapes, memory should be better when an optimal number of categories and items per category is used. In the present study, it is predicted that two categories with five items per category (Group 2) should be better

than ten categories with one item per category (Group 10).

As is probably obvious, random shapes do not readily lend themselves to the recall paradigm. Therefore, a recognition test is used. If the predicted superiority of Group 2 over Group 10 is found, it constitutes evidence for the effect of organizational processes, and thus retrieval processes, in recognition memory. In addition, cueing effects are examined by constructing different types of recognition test. Tests with retrieval cues are expected to yield higher recognition performance than those with no cueing.

CHAPTER 2

METHOD

Subjects

Forty students taking introductory psychology courses at the University of North Carolina at Greensboro during the spring of 1974 were used as subjects.

Apparatus

A total of thirty 24-point random shapes (Vanderplas & Garvin, 1959) were used in the experiment. Ten of the shapes were randomly selected from this pool and used as test items. The remainder were used as the distractor shapes on the recognition test. The test shapes were rear-projected by a Kodak Carousel 800 projector, which was controlled by solid state equipment, onto a translucent screen. The projected size of each shape was approximately six inches high and eight inches wide. The subject sat 3 feet on the far side of the screen. A small metal box containing red feedback lights was fastened to the table on which the screen rested.

For the recognition test, 5 test and 10 distractor shapes were printed on each of two cards, with the shapes on each card in a single column. The shapes on each card were lettered sequentially from top to bottom with the letters A-O appearing to the left of the shapes. The cards were 11 inches long and 4 inches wide. Responses by subjects were written on an answer sheet like the one

shown in Figure 1.

Procedure

There were two conditions of original learning and two types of recognition test in the 2x2 factorial design. In the learning task, 20 subjects used the numbers "1" and "2" to label the 10 shapes (Group 2) and 20 subjects used the numbers "1-10" (Group 10). Group 2 subjects thus responded to 5 shapes with "1" and the other 5 shapes with "2". Shapes having the same response will be referred to as being in the same "category". Subjects were read the following instructions:

> I am going to project 10 shapes on the screen in front of you, one at a time. I have assigned a number to each shape and you will try to learn, initially through guessing, what number goes with each shape. Use the numbers 1-10 for the 10 shapes (for Group 2 subjects: Use the numbers 1 and 2 to label the 10 shapes. Thus there will be 5 shapes you call 1 and 5 you call 2). A shape will be shown for 7 seconds. While it is being shown, please call out the number you think goes with it. Please speak loudly and distinctly, so I can be sure what number you are saying. Call out only one number for each shape presentation. If you are correct, the two red lights on the box in front of you will light up. We will continue until you get all 10 shapes correct 3 times in a row. Do you have any questions?

There was an interval of 3 seconds between shapes and a 13 second interval between the blocks of 10 shapes. Six different random orders of shape presentation were used with each subject. All subjects learned to a criterion of three successive correct trials.

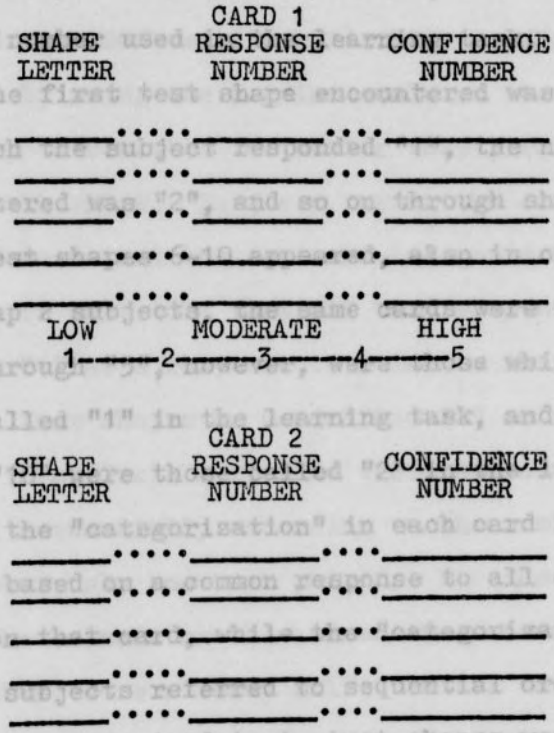


Fig. 1. Sample recognition test form.

One week after completion of the learning task, the appropriate recognition test was administered. The two types of test will be referred to as "categorized" and "uncategorized". The categorized test for Group 10 had the test shapes in sequential order, according to the response number used in the learning task. Thus, on Card 1, the first test shape encountered was the shape to which the subject responded "1", the next test shape encountered was "2", and so on through shape "5". On Card 2, test shapes 6-10 appeared, also in order.

For Group 2 subjects, the same cards were used. Shapes "1" through "5", however, were those which Group 2 subjects called "1" in the learning task, and shapes "6" through "10" were those called "2" in the learning task. Thus, the "categorization" in each card for Group 2 subjects was based on a common response to all of the test shapes on that card, while the "categorization" for Group 10 subjects referred to sequential ordering.

For the uncategorized test, test shapes were assigned to the two cards randomly, thus mixing the categories for Group 2 subjects and scrambling the sequence for Group 10 subjects. The position of test shapes within the columns was determined randomly, and was the same for both forms of the recognition test. Each distractor shape was in exactly the same location on both test forms.

Thus, the assignment of different test shapes to fixed test shape positions within the columns constituted the only difference between recognition tests.

When the recognition test was given, subjects were told that they were to pick out 5 shapes from each card that they had seen before, indicate their confidence in their choices by using the 5-point scale on the answer sheet, and try to remember the number they previously paired with each shape. Subjects taking the categorized test were informed that the test was categorized. Thus, Group 2 subjects were told that all test shapes appearing on Card 1 had "1" as a response term in the learning task, and all test shapes on Card 2 had "2" as a response term. Group 10 subjects were told that the test shapes were arranged sequentially according to the number with which they previously had been paired. For these cued conditions, the columns on the answer sheet headed by "Response Number" were already filled in (by the experimenter) with the appropriate numbers. Subjects taking the uncategorized recognition test were not given any such retrieval cues.

To fill out the answer sheet, subjects were told to write the letter that appeared next to each shape they chose (as one they had seen before) in the appropriate blanks on the answer sheet -- under the heading "Shape".

Then, they were to look at the confidence scale, select a number from the scale which indicated their degree of confidence that they had actually seen each chosen shape before, and place that number under the heading "Confidence Number" opposite the corresponding shape letter (the letter used to designate the shape). Then, subjects were to write the numbers they thought were previously paired with each shape in the column headed "Response Number." After completing this process for Card 1, subjects repeated the process for Card 2.

After the recognition test had been completed, the experimenter discussed the design of the experiment with the subject, and asked what types of strategies the subject had used. Subjects in cued conditions were asked if they thought the cueing helped, and subjects in non-cued conditions were asked if cueing probably would have helped them. Some of the subjects were asked for the specific associations they had used to remember each shape, and whether they had memorized only a specific portion of the shape instead of the entire shape.

CHAPTER 3

RESULTS

Learning data, recognition data, and shape difficulty information were analyzed separately, except where an aspect of one type of data was examined for its effect on some aspect of another type of data. The analysis of the learning began with a comparison of Group 2 and Group 10 on the standard measure "trials to criterion." Group 2 subjects took an average of 14.9 trials to reach criterion (S.D.=4.74), while Group 10 subjects took an average of 16.5 trials (S.D.=4.18). A *t* test indicated that the two groups did not differ significantly on this measure ($t=1.1$, $df=38$), assuring that later comparison of recognition scores of the two groups was not biased by differential frequency of exposure to the random shapes. The lack of a difference between these groups on the average trials required to reach criterion was simultaneously encouraging (since the recognition scores could confidently be compared) and puzzling, since there seemed to be reasons to expect the groups to differ on this measure. The primary difference between the two groups concerned the role of chance performance. Since chance performance for Group 2 subjects was 50%, while chance performance for Group 10 subjects was, at first, 10% and declined to 1 of 9, 1 of 8, and so on as shapes were learned, it seemed that Group 2 subjects would learn

faster because they got more shapes correct by chance and thus had more opportunity to learn the shapes in the early trials. Furthermore, Group 2 subjects could learn a shape even when they were wrong on a particular trial, since there were only two response possibilities.

The learning patterns for the two groups were found to be different, due largely to the different chance performance levels. The learning data for each subject consisted of ten separate learning curves for each of the ten shapes (see Figure 2). These ten curves were aligned at the trial of last error for each subject, to allow a better representation of actual performance prior to learning. When aligned in this manner, curves for some shapes were longer (covered more trials) than curves for other shapes, which would make overall curves for each subject misleading if they were drawn using "number of shapes correct" as the ordinate. Therefore, the proportion correct on each trial prior to the aligned trial of last error was used to draw adjusted learning curves for each subject. Adjusted curves for Group 2 are shown in Figure 3 and adjusted curves for Group 10 are shown in Figure 4. Curves characteristically began between 40% and 50% for Group 2 subjects and then progressed erratically to criterion. Group 10 subjects started near the chance level of 10% and advanced to

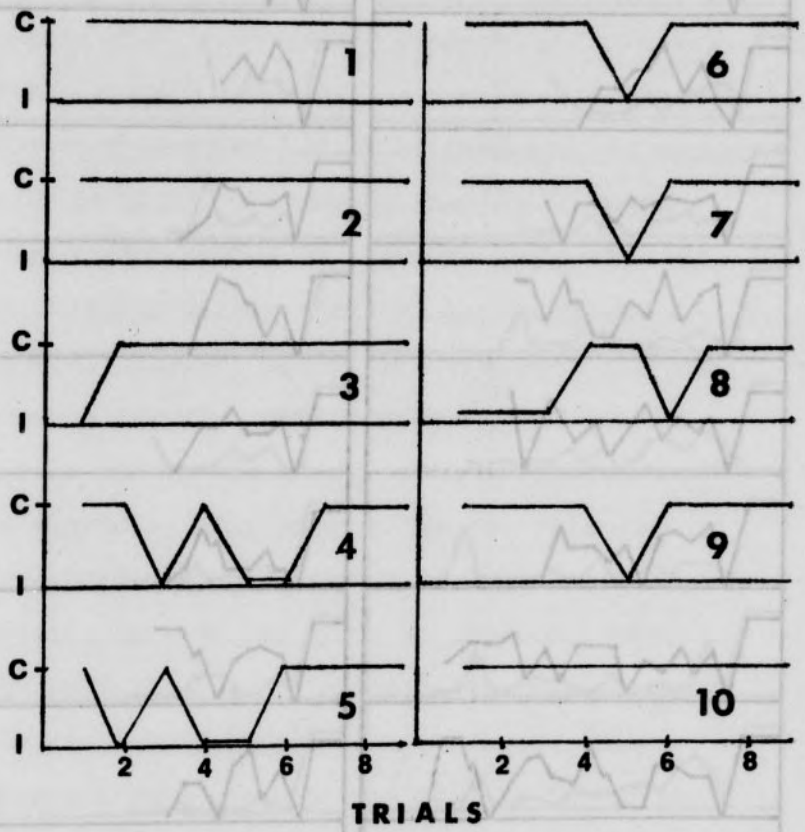


Fig. 2. Learning curves for each shape for a typical subject in Group 2 (C=correct, I=incorrect). Numbers refer to specific shapes.

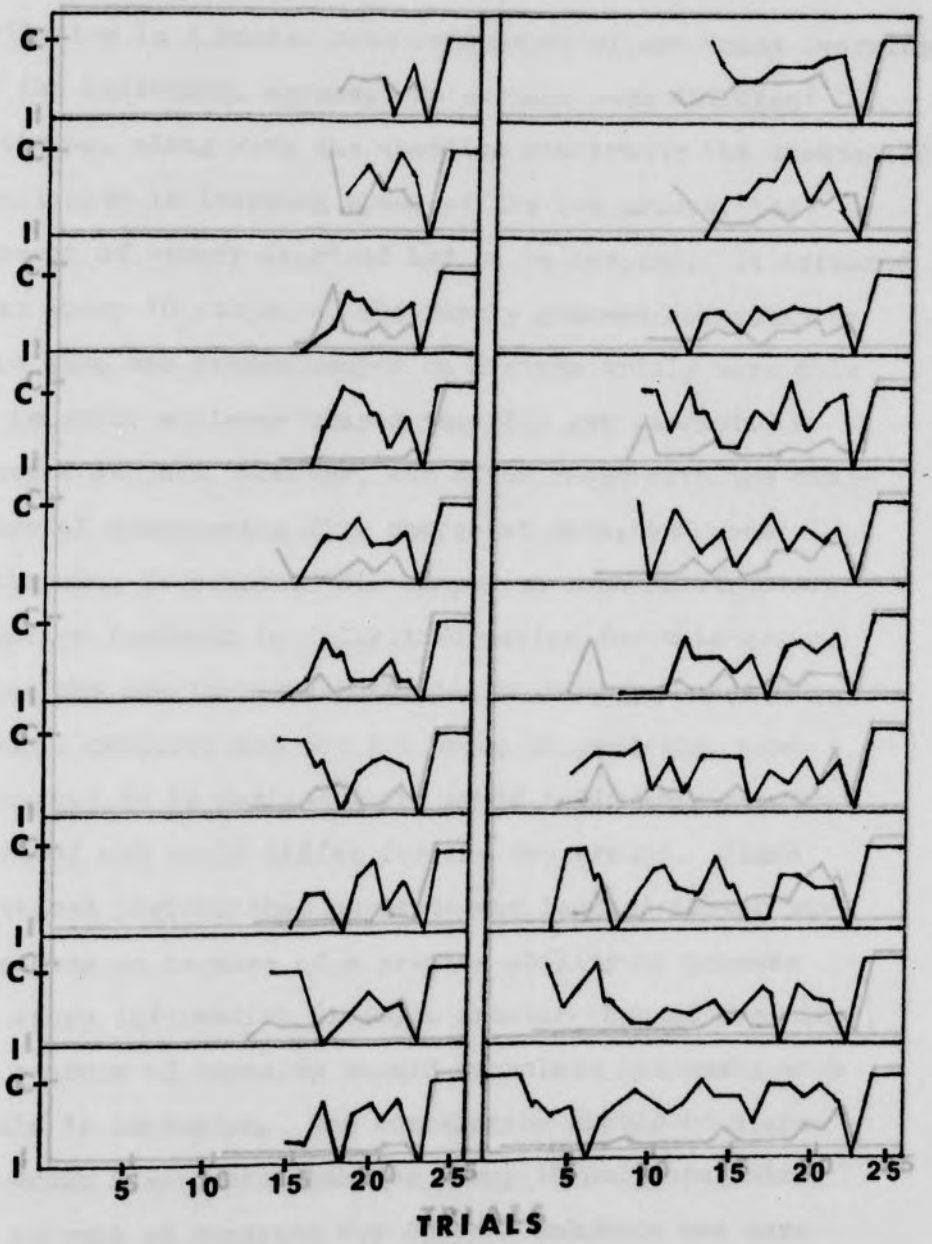


Fig. 3. Adjusted learning curves for the 20 subjects in Group 2. C=correct and I=incorrect.

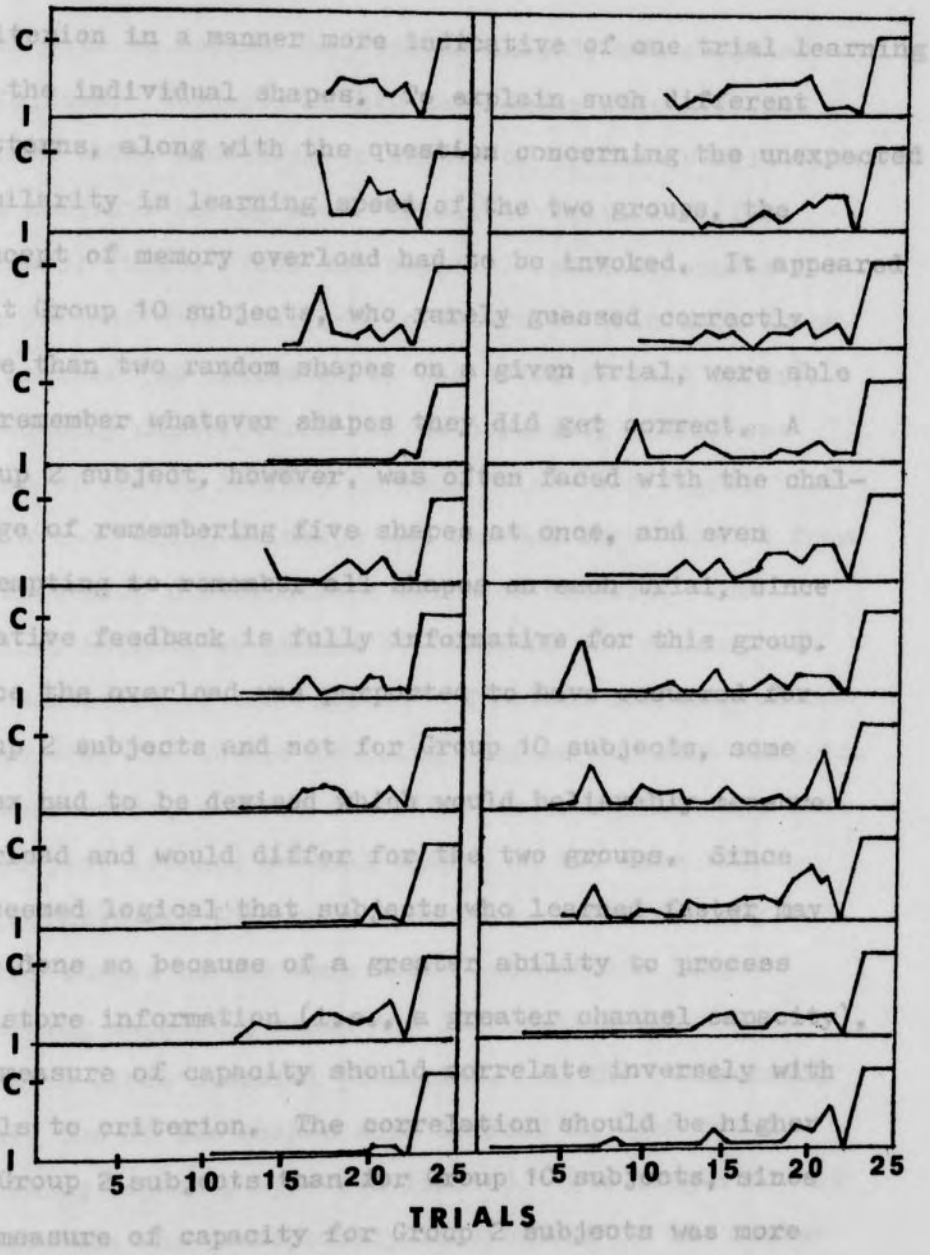


Fig. 4. Adjusted learning curves for the 20 subjects in Group 10. C=correct and I=incorrect.

criterion in a manner more indicative of one trial learning of the individual shapes. To explain such different patterns, along with the question concerning the unexpected similarity in learning speed of the two groups, the concept of memory overload had to be invoked. It appeared that Group 10 subjects, who rarely guessed correctly more than two random shapes on a given trial, were able to remember whatever shapes they did get correct. A Group 2 subject, however, was often faced with the challenge of remembering five shapes at once, and even attempting to remember all shapes on each trial, since negative feedback is fully informative for this group. Since the overload was purported to have occurred for Group 2 subjects and not for Group 10 subjects, some index had to be devised which would believably measure overload and would differ for the two groups. Since it seemed logical that subjects who learned faster may have done so because of a greater ability to process and store information (i.e., a greater channel capacity), the measure of capacity should correlate inversely with trials to criterion. The correlation should be higher for Group 2 subjects than for Group 10 subjects, since the measure of capacity for Group 2 subjects was more valid (since they were presumably more likely to have been overloaded with information). The devised index,

which will be called the "spread index", was computed as follows. Subjects within each group were arranged in rank order according to number of trials to criterion, with subjects having fewer trials to criterion (fast learners) first. Then the original learning curves for each subject were analysed to determine when each shape was learned (as indicated by the trial of last error). A frequency distribution was constructed, giving the number of shapes learned on each trial (see Figure 5). Trials upon which no shape was learned were deleted from the frequency distributions for each subject. The physical height, or number of items in the column, was representative of the "spread" of actual learning of the shapes over the entire learning process for each subject. In other words, a short distribution would indicate that the subject learned several shapes on the same trial, and probably did this more than once. The learning of several shapes on the same trial indicated a large processing capacity, since the information was all stored in close temporal proximity. On the other hand, if only one or two shapes were learned on each trial, processing capacity appeared limited. The length of the distribution was conceptually independent of the number of trials to criterion, since several shapes could be learned on the same trial anywhere in the learning process.

GROUP 2
Trials to Criterion

7	9	11	11	12	12	12	12	13	13	14	14	15	17	17	18	21	22	22	26
5-2	7-2	9-1	9-1	10-1	10-3	10-1	10-2	11-1	11-1	12-1	12-1	13-2	15-1	15-1	16-1	19-2	20-1	20-2	24-1
3-4	6-4	8-2	8-1	9-1	9-1	9-2	9-1	8-1	10-2	10-2	11-1	11-4	14-1	14-1	13-1	18-1	17-2	18-1	21-2
2-2	2-1	7-1	7-2	7-3	8-1	7-2	8-1	7-1	9-1	9-1	9-1	8-1	11-1	13-2	12-2	15-2	13-1	17-1	20-1
1-2	1-3	6-2	6-2	6-1	5-1	6-1	7-1	6-1	6-1	8-2	6-4	7-1	9-3	11-1	10-2	14-1	11-1	13-1	19-1
		5-1	3-4	5-1	1-4	5-1	6-1	4-3	5-2	7-2	5-2	6-1	8-1	10-1	7-1	12-1	10-2	10-1	14-1
		4-2		4-2		4-1	4-2	3-1	2-2	2-2	4-1	3-1	7-1	9-1	6-1	8-1	9-1	9-1	13-1
		3-1		3-1		2-1	3-1	2-2	1-1				5-1	7-1	5-1	3-1	3-1	7-1	12-1
						1-1	1-1						2-1	6-2	1-1	1-1	1-1	5-1	10-1
																		3-1	3-1
4	4	7	5	7	5	8	8	7	7	6	6	5	8	8	8	8	8	9	9 SI

GROUP 10
Trials to Criterion

10	10	12	13	13	14	14	14	14	16	16	17	17	19	20	21	21	21	23	25
8-4	8-1	10-2	11-1	11-2	12-1	12-1	12-2	12-3	14-1	14-4	15-2	15-1	17-1	18-2	19-2	19-1	19-1	21-1	23-1
7-2	7-3	9-1	10-1	10-1	11-2	11-1	11-1	11-1	13-1	13-1	14-1	14-2	16-1	17-1	18-1	17-1	17-1	18-1	21-1
6-3	5-2	8-1	9-2	9-1	8-1	10-3	10-1	10-1	12-1	5-1	13-1	13-1	14-1	16-1	15-1	16-1	16-2	17-1	19-2
4-1	4-1	6-1	5-1	8-1	7-2	9-1	9-1	7-1	10-1	4-1	11-1	12-1	12-2	12-1	14-1	14-1	15-1	16-2	17-2
	3-1	5-3	4-1	7-1	6-1	8-1	8-1	5-1	9-2	3-2	10-3	11-1	11-2	9-1	13-1	13-1	14-1	15-1	13-1
	2-2	4-1	3-1	6-2	4-1	7-1	7-1	4-1	8-1	2-1	9-1	9-1	10-1	8-1	11-1	10-1	11-3	14-1	11-1
		1-1	2-1	2-2	2-1	5-1	6-1	3-1	7-1		8-1	5-2	6-2	7-1	10-1	7-2	10-1	6-1	10-1
			1-1		1-1	2-1	5-1	1-1	6-1			2-1		6-1	8-2	4-1		4-1	2-1
							3-1		2-1					4-1		1-1		1-1	
4	6	7	8	7	8	8	9	8	9	6	7	8	7	9	8	9	7	9	8 SI

Fig. 5. Illustration of derivation of the "spread index." Left member of each hyphenated number pair is the trial number; right member is the number of shapes learned on that trial. "SI" refers to the "spread index."

Thus the link between trials to criterion and the spread index is theoretical, and not a mathematical inevitability. The Pearson correlations performed were attempts to measure the relationship between trials to criterion (speed of learning) and the spread index for each group. The Pearson correlation for Group 2 was .779 ($df=18$, $p<.01$) while that for Group 10 was .502 ($df=18$, $p<.05$). Thus, the hypothesis concerning processing and memory capacity was supported. The lack of difference between the two groups in trials to criterion can be understood as due to the inability of Group 2 subjects to take full advantage of the more abundant information afforded them, with the inability presumably due to a limit on the amount of information that they could process at a given time. Scatter plots showing the relationship between the spread index and trials to criterion appear in Figure 6.

Overall learning curves were constructed for each group by a quasi-Vincentizing procedure. The adjusted learning curves for each subject were examined and the number of correct responses per trial was recorded. These numbers were each divided by the number of shapes which actually were represented by learning curves on each trial. The resulting index indicated the proportion correct for each trial. These indices were listed for

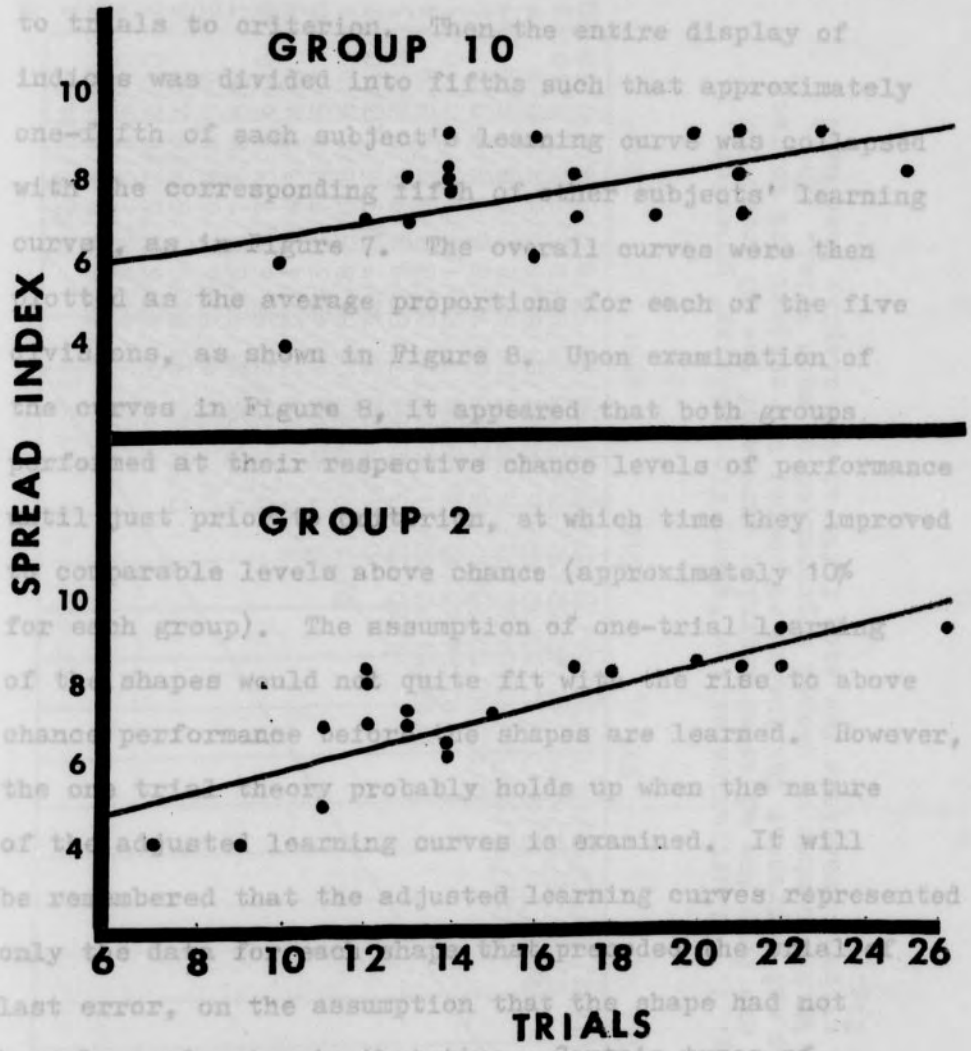
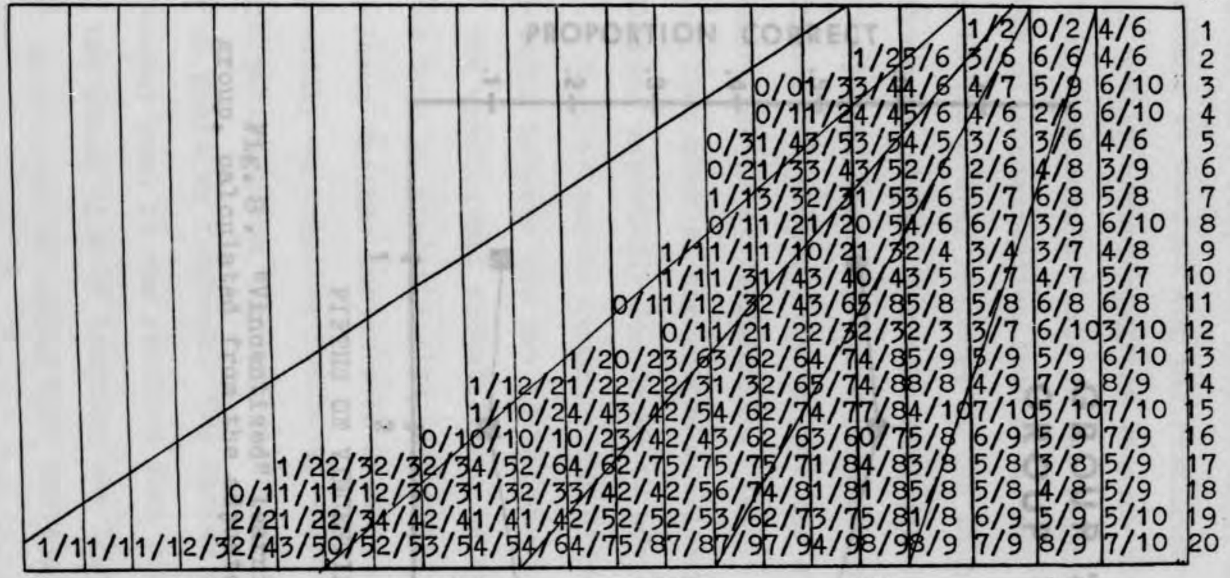


Fig. 6. Scatter plot for each group showing correlation between trials to criterion and the spread index.

each subject, with the subjects being ordered according to trials to criterion. Then the entire display of indices was divided into fifths such that approximately one-fifth of each subject's learning curve was collapsed with the corresponding fifth of other subjects' learning curves, as in Figure 7. The overall curves were then plotted as the average proportions for each of the five divisions, as shown in Figure 8. Upon examination of the curves in Figure 8, it appeared that both groups performed at their respective chance levels of performance until just prior to criterion, at which time they improved to comparable levels above chance (approximately 10% for each group). The assumption of one-trial learning of the shapes would not quite fit with the rise to above chance performance before the shapes are learned. However, the one trial theory probably holds up when the nature of the adjusted learning curves is examined. It will be remembered that the adjusted learning curves represented only the data for each shape that preceded the trial of last error, on the assumption that the shape had not been learned prior to that time. Certain types of response sequence, however, can falsely be included in the adjusted learning curves, in that the shape may have already been learned. For example, occasionally a subject would respond correctly to a particular shape for ten



																						S
																			1/2	0/2	4/6	1
																	1/2	5/6	3/6	6/6	4/6	2
															0/0	1/3	3/4	4/6	4/7	5/9	6/10	3
															0/1	1/2	4/4	5/6	4/6	2/6	6/10	4
														0/3	1/4	3/5	3/5	4/5	3/6	3/6	4/6	5
														0/2	1/3	3/4	3/5	2/6	2/6	4/8	3/9	6
														1/1	3/3	2/3	1/5	3/6	5/7	6/8	5/8	7
														0/1	1/2	1/2	0/5	4/6	6/7	3/9	6/10	8
													1/1	1/1	1/1	0/2	1/3	2/4	3/4	3/7	4/8	9
													1/1	1/3	1/4	3/4	0/4	3/5	5/7	4/7	5/7	10
												0/1	1/1	2/3	2/4	3/6	5/8	5/8	5/8	6/8	6/8	11
													0/1	1/2	1/2	2/3	2/3	2/3	3/7	6/10	3/10	12
											1/2	0/2	3/6	3/6	2/6	4/7	4/8	5/9	5/9	5/9	6/10	13
									1/1	2/2	1/2	2/2	2/3	1/3	2/6	5/7	4/8	8/8	4/9	7/9	8/9	14
									1/1	0/2	4/4	3/4	2/5	4/6	2/7	4/7	7/8	4/10	7/10	5/10	7/10	15
								0/1	0/1	0/1	0/2	3/4	2/4	3/6	2/6	3/6	0/7	5/8	6/9	5/9	7/9	16
					1/2	2/3	2/3	2/3	4/5	2/6	4/6	2/7	5/7	5/7	5/7	1/8	4/8	3/8	5/8	3/8	5/9	17
				0/1	1/1	1/1	2/3	0/3	1/3	2/3	3/4	2/4	2/5	6/7	4/8	1/8	1/8	5/8	5/8	4/8	5/9	18
				2/2	1/2	2/3	4/4	2/4	1/4	1/4	2/5	2/5	2/5	3/6	2/7	3/7	5/8	1/8	6/9	5/9	5/10	19
1/1	1/1	1/1	2/3	2/4	3/5	0/5	2/5	3/5	4/5	4/6	4/7	5/8	7/8	7/9	7/9	4/9	8/9	8/9	7/9	8/9	7/10	20

Fig. 7. Illustration of the "quasi Vincentizing" procedure. Numbers are from the adjusted learning curves of the 20 Group 2 subjects and represent the number of shapes the subject got correct divided by the number of shapes represented on each trial. The overall average fractions for each fifth of the display were used to draw Fig. 8.

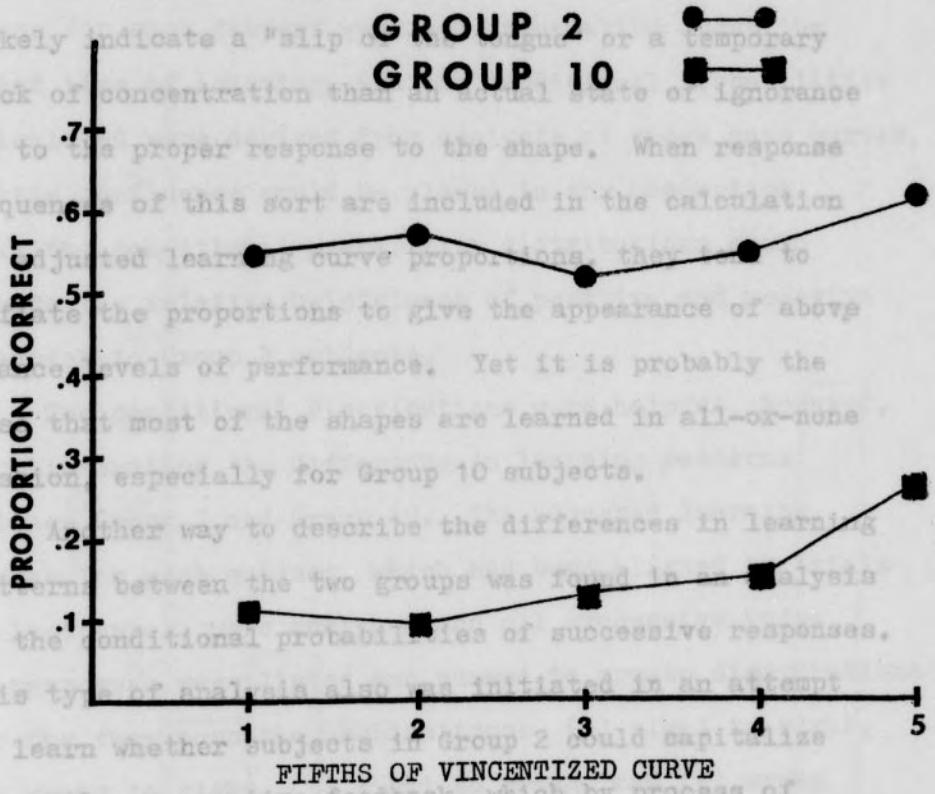


Fig. 8. "Vincentized" learning curves for each group, calculated from the adjusted learning curves.

trials in a row, then miss it once, and continue on getting it correct for the remainder of the experiment. The one incorrect response in such a case would more likely indicate a "slip of the tongue" or a temporary lack of concentration than an actual state of ignorance as to the proper response to the shape. When response sequences of this sort are included in the calculation of adjusted learning curve proportions, they tend to inflate the proportions to give the appearance of above chance levels of performance. Yet it is probably the case that most of the shapes are learned in all-or-none fashion, especially for Group 10 subjects.

Another way to describe the differences in learning patterns between the two groups was found in an analysis of the conditional probabilities of successive responses. This type of analysis also was initiated in an attempt to learn whether subjects in Group 2 could capitalize as easily on negative feedback, which by process of elimination gives the correct answer to the subject, as on positive feedback. The latter quest ran aground somewhat when it was realized that precise specification of the exact time of learning was impossible. It was not sufficient to assume that the shapes were "learned" exactly when they passed from the trial of last error to the next trial. It could have been that the shape

was answered correctly by chance on that trial and was actually learned on the trial or trials following the trial of last error. Since the learning curves for each shape for each subject were not informative about the exact time of learning, and the conditional probabilities calculated were derived from analysis of these same curves, little confidence could be placed in the assumption that the conditional probability distributions could decide the relative helpfulness of positive and negative feedback to Group 2 subjects.

The conditional distributions were helpful, however, in illustrating the difference in learning patterns between Group 2 and Group 10. The adjusted learning curves for each subject, which had been aligned by trials of last error, were analyzed and all successive pairs of responses were listed and summed to create distributions for the four possible combinations: (1) right to right, (2) wrong to right, (3) right to wrong, and (4) wrong to wrong. Thus, for each subject a 2x2 matrix was created which showed the total number of instances which fell into each of the four categories. The totals in each cell were then converted to percentages for each subject, and these cell percentages were used to calculate mean cell percentages for Group 2 and Group 10 independently. The conditional distributions are shown in Figure 9.

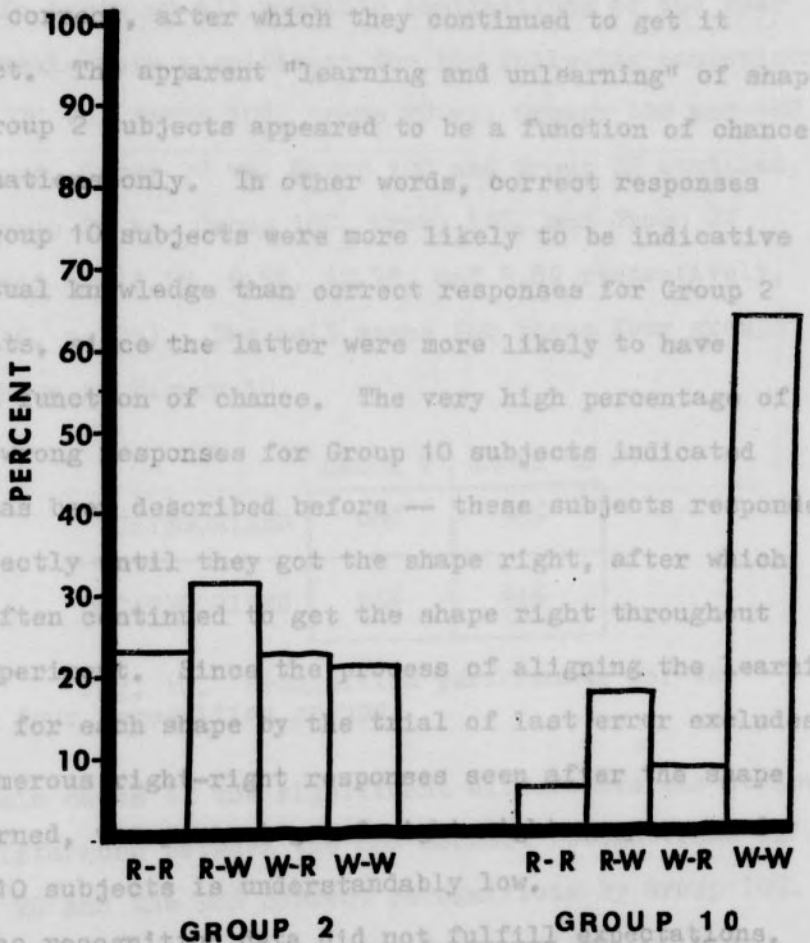


Fig. 9. Conditional probabilities for the two groups. R-R=right to right, R-W=right to wrong, W-R=wrong to right, W-W=wrong to wrong.

The response sequences wrong-to-right and right-to-wrong comprised a larger proportion of all responses for Group 2 than for Group 10 subjects. Group 10 subjects tended to respond incorrectly until they finally got a shape correct, after which they continued to get it correct. The apparent "learning and unlearning" of shapes for Group 2 subjects appeared to be a function of chance fluctuations only. In other words, correct responses for Group 10 subjects were more likely to be indicative of actual knowledge than correct responses for Group 2 subjects, since the latter were more likely to have been a function of chance. The very high percentage of wrong-wrong responses for Group 10 subjects indicated what has been described before -- these subjects responded incorrectly until they got the shape right, after which they often continued to get the shape right throughout the experiment. Since the process of aligning the learning curves for each shape by the trial of last error excludes the numerous right-right responses seen after the shape is learned, the percentage of right-right responses of Group 10 subjects is understandably low.

The recognition data did not fulfill expectations, primarily because of unwanted ceiling effects that could not be eliminated. A 2x2 analysis of variance on the recognition scores of the four groups -- categorized and

uncatcgorized by Group 2 and Group 10 -- revealed significant main effects for number of responses ($F=4.52$, $df=1,36$, $p<.05$) and a significant interaction ($F=6.78$, $df=1,36$, $p<.05$), as summarized in Table 1. Scheffe tests on all possible combinations of the four cell means were significant for the following comparisons: Group 2C vs. Group 10C, Group 2C vs. Groups 10C and 10U combined, Group 2C vs. Group 10C and Group 2U combined, and Group 2C vs. Group 10C, Group 10U, and Group 2U combined ($F=11.19$, 9.55, 10.78, and 9.86 respectively, $df=3,36$, $p<.05$). The cell means for these four groups are shown in Figure 10.

	GROUP 2	GROUP 10
CATEGORIZED	88%	98%
UNCATEGORIZED	95%	94%

Fig. 10. Recognition performance for the four recognition groups.

The main cause of the significant differences was obviously the difference between the 88% correct recognitions of Group 2C and the 98% correct recognitions by Group 10C. It must be remembered that the "categorization" was different for the two groups, with response number constituting the organization for Group 2 and serial position (1 through 10) for Group 10. It appeared that

TABLE 1

2x2 Analysis of Variance for Recognition Scores

Source of Variance	Calculations				
	SS	df	MS	F	p
No. of response categ.	2.02	1	2.02	4.52	<.05
Categor-ization	.22	1	.22	.49	NS
Inter-action	3.03	1	3.03	6.78	<.05
Error	16.10	36	.45		

the serial position cue for Group 10 subjects was a more effective cue than the category membership cue for Group 2 subjects. Some Group 10 subjects mentioned in the post-experimental interview that they were helped in situations where they could, for example, remember shapes 1 and 3 and then know where to search more specifically for shape 2. Since Group 2C was different from other groups only when Group 10C was included in the analysis, it would be inappropriate to state that Group 2C subjects were actually hindered by the category cueing. Instead, the results indicated that Group 10C was different from Group 2C and not much else. The serial position cueing was helpful, but the response number category cueing was not.

The confidence scores, which were indicative of the subjects' confidence that the shapes they listed on the answer sheet were the correct shapes, were not systematically related to the actual recognition scores of the subjects, as indicated by a nonsignificant Pearson correlation between the confidence and recognition scores ($r=.263$, $df=38$). The lack of significance of this correlation should not be extremely surprising, since there was much opportunity for response bias to be much more different for subjects than the very homogeneous recognition scores. Some subjects reported extreme

confidence to the experimenter after the experiment, along with indicating the maximum degree of confidence on the answer sheet. Other subjects, however, verbally expressed an uncertainty that they had selected the correct shapes. These subjects tended to diverge widely in their interpretation of the confidence scale, some using the lowest point on the scale to indicate a degree of uncertainty which corresponded, according to verbal description, with degrees of confidence of other subjects who used higher points on the scale. Often the subjects who recorded uncertainty scored perfectly on the recognition test. Evidently the measure of confidence as used in the present experiment was not a particularly valid measure.

The random shapes used in the present study were all 24-point shapes, yet they varied in difficulty of learning. Attempts were made to determine the difficulty of the various shapes, and the possible relationship of scaled psychological responses to the shapes with their difficulty. Three different measures of difficulty were used. First, the trial of last error was used, with the assumption that shapes learned first were easier than those whose trial of last error came later. This method probably yielded a more accurate measure of shape difficulty for Group 2 than for Group 10, since the

bias from which shapes were first guessed correctly was greater for the latter group. The second and third measures of shape difficulty were devised to overcome the biasing effect of chance to more adequately guage the actual relative difficulty of the shapes. One of the measures devised consisted entirely of trials between the first correct response for a shape and the trial following the trial of last error. It was reasoned that longer intervals would signify more difficulty in learning the shape than would shorter intervals. The shapes were put in rank order according to the size of these intervals. The interval sizes for each shape were averaged across all subjects in each of the two main groups (Group 2 and Group 10). This measure was, however, likely to be insensitive to certain subtle types of response sequences which falsely indicated that a shape was difficult. There were two types of response sequence which were deceptive. One type, which has already been discussed in the consideration of the adjusted learning curves, was characterized by a string of correct responses followed by a single (or in some cases, two) incorrect response, and then a succession of correct responses that continued until the criterion was reached. Such a pattern would have too large an interval index for the actual degree of difficulty suggested, since the single incorrect

response more probably indicated a slip of the tongue or a momentary lapse of concentration on the subject's part. At the other extreme, subjects would sometimes get a shape correct initially, and then miss it for one, two, or three trials in a row before getting it correct for the remainder of the learning task. Adjustments were sought to keep such patterns from being added to the index of difficulty, since it was likely that the first correct response in this case was a chance occurrence and did not indicate that the subject had knowledge of the shape. When the number of trials containing errors following a correct response exceeded three, however, however, shape difficulty was assumed to be indicated (although this argument tends to be more valid for Group 2 than for Group 10, since in the latter case it was not as likely that the subject would again get the shape correct by chance in a small number of trials). Thus, a third index of difficulty was devised to adjust for these two specific types of response sequence. The proportion of the responses which were correct within the interval used for the second index was calculated for each shape and expressed in decimal form. A value of "1" was assigned to all instances where this proportion was greater than or equal to .19 and less than or equal to .81. When the numerator of a proportion was zero and

the denominator three or less, a value of "0" was assigned. When the numerator of the proportion was zero and the denominator exceeded three, a value of "1" was assigned. The "1" values were indicative of shape difficulty, and were summed for each shape across all subjects in Group 2 and Group 10 independently. The ten shapes were then rank ordered according to the totals obtained.

In addition to the measures of shape difficulty, the shapes were rank ordered according to three other measures, which were obtained by Vanderplas & Garvin (1959). The shapes were scaled on association value, proportion of content responses, and heterogeneity. The association value represented the percentage of all subjects questioned in their study who either gave an association to a shape or who indicated that the shape reminded them of something, though they did not verbalize it. The proportion of cases in which actual responses were given, instead of just saying "I had an association", was the "proportion content response" measure. The final measure indicated the degree to which responses varied for each shape across subjects. These three indices were used in the present study to determine whether they could help explain the relative difficulty of the shapes.

Rank order correlations were calculated for all meaningful combinations of shape orderings, and the

correlations are listed in Figure 11. The correlations did not indicate strong relationships between shape difficulty and the association-type measures. Since only five out of twenty-seven correlations were significant and only two of these concerned the association measures, it appeared that prediction of shape difficulty would have to encompass other factors before becoming accurate. It did appear that Scheme A, the method in which certain misleading types of response sequence were eliminated, allowed the best determination of shape difficulty, especially for Group 10. That trial of last error was a confounded measure for Group 10 is evident when the correlation of Scheme A and Proportion Content Responses (.676) is compared to the correlation of Trial of Last Error and Proportion Content Responses (.043). A similar comparison for Group 2 shows a greater agreement between the two measures, with the correlation for Scheme A and Proportion Content Responses being .530, and the correlation for Trial of Last Error and Proportion Content Responses being .285. The trial of last error for Group 10 subjects was, as expected, affected too much by the order in which the shapes were guessed.

The three measures devised to ascertain shape difficulty agreed highly for Group 2, as indicated by the high correlations among these measures for that

	SCHEME A	TRIAL OF LAST ERROR	LENGTH OF INTERVAL	ASSOC-IATION VALUE	PROP. CONTENT R's	HETERO-GENEITY
SCHEME A		.791 -.027	.882 .524	-.209 .355	.530 .676	.021 .421
TRIAL OF LAST ERROR			.915 .358	.079 .200	.285 .043	.213 .576
LENGTH OF INTERVAL				.067 .648	.515 .333	.225 .625
ASSOCIATION VALUE					.382	.527
PROPORTION CONTENT RESPONSES						.564
HETERO-GENEITY						

Fig. 11. Rank order correlation matrix for shape difficulty measures and "association" scales. When two numbers are present, top is for Group 2 and other for Group 10. Critical values: p .05=.648, p .01=.794.

group, but did not seem to adequately measure the same thing for Group 10 subjects. The lack of agreement for Group 10 was most likely attributable to the inadequacy of the Trial of Last Error and the Length of Interval measures for this group. The Length of Interval measure was not adequate because, for many shapes in Group 10, it was often zero. In other words, many shapes were learned in one trial and thus were not easily distinguished in terms of ease of learning. The Scheme A approach also suffered from the presence of such one-trial learning, but was more useful because it eliminated the two types of misleading response sequence which appeared mainly for Group 10.

CHAPTER 4

DISCUSSION

The recognition test data were disappointing, in that unwanted, but seemingly uncorrectable, ceiling effects appeared to preclude the appearance of a meaningful pattern of results. The ceiling effects could have been avoided, of course, if delays of four weeks instead of one week were used. However, such delays would presumably nullify the possibility for cueing effects, since the numbers associated with the shapes, which served as the basis for organization, would most likely be forgotten during longer intervals. Using fewer than ten shapes to speed the learning process, in order to decrease frequency of exposure to the shapes, would be inadvisable for at least two reasons. First, the possible benefit for grouping would be lessened as the total number of items came within the processing capacity of subjects. Secondly, the meaningfulness of the recognition scores would be lessened, since the percentages correct would be based on a smaller number of items (e.g., 8 out of 10 correct may be a better indication of performance than 4 out of 5). Using more difficult shapes would presumably make recognition more difficult, but would also increase learning time and, ipso facto, shape exposure frequency. Using simpler shapes to speed learning and reduce exposure frequency would work, except that the

simpler shapes would probably be more easily recognized. It appears that the method of inducing organization used here, with the stringent criterion of three successive correct trials, does not lend itself to the study of recognition memory because the overlearning produces, inevitably it seems, ceiling effects on the recognition test. Other methods of organization (spatial, chromatic, categories, etc.) could possibly be used more profitably, although Mandler's (1972) requirement of a high degree of organization should be kept in mind.

The learning data proved much more interesting, and were analyzed in much detail partly because of the absence of exciting recognition data. The discovery of the role of processing capacity in equalizing the learning times of the two groups and in explaining the main reason for differences among subjects in learning speed was the most significant aspect of the study. It appeared that the limit on learning rate for the random shapes as used in the present experiment was 2±1 shapes per trial. Comparison of this measure of processing capacity is not easily compared to limits established using other paradigms. Perhaps it is best to state simply that the maximum learning rate for random shapes seemed to indicate that they are harder to learn than verbal materials.

Some of the subjects in Group 2 reported after the experiment that they had visualized two "chunks", with "1" shapes in one chunk and "2" shapes in the other. Other subjects reported no such distinct imagery, yet used associative schemes in which all shapes with the same response number were were associated with the same mnemonic. A frequently used strategy was to find in the shape (in part by self-persuasion) the number that was to be associated with it. Some subjects could find "ones" in the "1" shapes, but were unable to find "two's" in the "2" shapes. These subjects would sometimes adopt a strategy which seemed most efficient to the experimenter (perhaps because he used such a strategy when he was in a similar experiment as a subject): Find something similar about one chunk's items and remember that any shape not possessing the required characteristic was a member of the other chunk. The subject who learned in the fewest number of trials (seven trials) reported using just such a strategy. Group 10 subjects usually learned a separate association for each shape, and were able to remember almost every association they had used for the one week interval between the learning task and the recognition test.

The development of an adequate way to measure shape difficulty was considered worthwhile, though the attempt

to determine why some shapes were more difficult than others fell short of being conclusive. The lack of similarity of shape difficulty rank orderings between Groups 2 and 10 probably indicated that which shapes were confused by the subject depended on the mnemonic devices used to remember them, which, in turn, depended on the nature of the task specific to each group. For example, Group 2 subjects who tried to find a "one" in the "1" shapes were bothered by the existence of "2" shapes that contained portions looking like a "1". Such confusions were subtly different for different subjects, largely because they tended to focus in on different parts of the shapes (as their functional stimulus), and even the best attempts of the experimenter to anticipate which shapes a particular subject would find difficult to discriminate were thwarted. Group 10 subjects tended to have the most difficulty with shapes that did not readily look like anything in particular. The significant correlation between Proportion Content Responses and the Scheme A difficulty index for Group 10 supports such a contention. Some subjects formed associations to all the shapes they could, and then memorized the number that went with the "shapes that didn't look like anything."

Since the recognition data were to have been the most important data in the experiment, and ceiling effects

substantially clouded the results, the paradigm of Richman and his associates must be considered limited in its applicability to the study of recognition memory. The paradigm is, however, interesting from a learning, instead of a memory, standpoint. Further work of a parametric nature could be done with the paradigm, examining the effects of various rates of stimulus presentation, difficulty of shapes, and number of response categories.

BIBLIOGRAPHY

Bahrick, H. P., and Bahrick, P. O. A re-examination of the interrelations among measures of retention. Quarterly Journal of Experimental Psychology, 1964, 16, 318-324.

Bartlett, F. C. Remembering: A study in Experimental and Social Psychology. Cambridge at the University Press, 1932.

Bower, G. H., Clark, M. C., Lesgold, A. M., and Winzenz, D. Hierarchical retrieval schemes in recall of categorized word lists. Journal of Verbal Learning and Verbal Behavior, 1969, 8, 323-343.

Bruce, D., and Fagan, R. L. More on the recognition and free recall of organized lists. Journal of Experimental Psychology, 1970, 85, 153-154.

D'Agostino, P. R. The blocked-random effect in recall and recognition. Journal of Verbal Learning and Verbal Behavior, 1969, 8, 815-820.

DaPolito, F., Barker, D., and Wiant, J. Context in semantic information retrieval. Psychonomic Science, 1971, 24, 180-181.

Jacoby, L. L. Effects of organization on recognition memory. Journal of Experimental Psychology, 1972, 92, 325-331.

Jacoby, L. L., and Hendricks, R. L. Recognition effects of study organization and test context. Journal of Experimental Psychology, 1973, 100, 73-82.

Johnson, N. F. The role of chunking and organization in the process of recall. In Gordon H. Bower (Ed.) The Psychology of Learning and Motivation, Vol. 4, 1970.

Katona, G. Organizing and Memorizing. New York: Columbia University Press, 1940.

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

Kintsch, W. Models of free recall and recognition. In D. A. Norman (Ed.) Models of Human Memory. New York: Academic Press, 1970.

Lachman, R., and Tuttle, A. V. Approximation to English and short-term memory: Construction or storage? Journal of Experimental Psychology, 1965, 70, 386-393.

Lewis, M. Q. Categorized lists and cued recall. Journal of Experimental Psychology, 1971, 87, 129-131.

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

Light, L. L., and Schurr, S. C. Context effects in recognition memory: Item order and unitization. Journal of Experimental Psychology, 1973, 100, 135-140.

Luh, C. W. The conditions of retention. Psychology Monographs, 1922, 31, No. 142.

Mandler, G. From association to structure. Psychological Review, 1962, 69, 415-427.

Mandler, G. Organization and memory. In K. W. Spence and J. T. Spence (Eds.) The Psychology of Learning and Motivation, 1967, Vol. 1.

Mandler, G. Association and organization: Facts, fancies, and theories. In T. R. Dixon and D. L. Horton (Eds.) Verbal Behavior and General Behavior Theory. Englewood Cliffs, N. J.: Prentice-Hall, 1968.

Mandler, G. Organization and recognition. In E. Tulving and W. Donaldson (Eds.) Organization of Memory. New York: Academic Press, 1972.

McCormack, P. D. Recognition memory: How complex a retrieval system? Canadian Journal of Psychology, 1972, 26, 19-41.

McNulty, J. A. A partial learning model of recognition memory. Canadian Journal of Psychology, 1966, 20, 302-315.

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

Murdock, B. B. An analysis of the recognition process. In C. N. Cofer and B. Musgrave (Eds.) Verbal Behavior and Learning: Problems and Processes. New York: McGraw Hill, 1963.

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

Postman, L., and Rau, L. Retention as a function of the method of measurement. University of California Publications in Psychology, 1957, 8, 217-270.

Richman, C. L. Chunk formation. In preparation, 1974.

Richman, C. L., Bidwell, L., and Henson, K. Stimulus organization: Type of overtraining criterion and temporal factors. Unpublished manuscript, 1973.

Richman, C. L., and Trinder, J. Effect of a novel stimulus dimension on discrimination learning. Journal of Experimental Psychology, 1968, 77, 163-165.

Trinder, J., Metzger, K., Sherman, A., and Richman, C. L. Development of stimulus organization. Developmental Psychology, 1972, 6, 26-31.

Trinder, J., Richman, C. L., and Gulkin, R. Effects of a Slamecka shift on reversal and half-reversal shift tasks. Proceedings of the 77th Annual Convention, APA, 1969, 63-64.

Tulving, E. Theoretical issues in free recall. In T. R. Dixon and D. L. Horton (Eds.) Verbal Behavior and General Behavior Theory. Englewood Cliffs, N. J.: Prentice-Hall, 1968(a).

Tulving, E. When is recall higher than recognition? Psychonomic Science, 1968(b), 10, 53-54.

Tulving, E., and Osler, S. Effectiveness of retrieval cues in memory for words. Journal of Experimental Psychology, 1968, 77, 593-601.

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

Tulving, E., and Thomson, D. W. Retrieval processes in recognition memory: Effects of associative context. Journal of Experimental Psychology, 1971, 87, 116-125.

Winograd, E., Karchmer, M. A., and Russell, I. S. Role of encoding unitization in cued recognition memory. Journal of Verbal Learning and Verbal Behavior, 1971, 10, 199-206.