This is to certify that the

thesis entitled

Psychophysical Testing in Cartography:

An Evaluation of Methodology

presented by

Daniel Gerard Cole

has been accepted towards fulfillment
of the requirements for

M.A. degree in Geography

Richard Groop
Major professor

Date 5/8/79

O-7639

PSYCHOPHYSICAL TESTING IN CARTOGRAPHY; AN EVALUATION OF METHODOLOGY

By

Daniel Gerard Cole

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

MASTER OF ARTS

Department of Geography

1979

ABSTRACT

PSYCHOPHYSICAL TESTING IN CARTOGRAPHY: AN EVALUATION OF METHODOLOGY

By

Daniel Gerard Cole

Cartographers have employed psychophysical testing to determine the average map reader's perception of the sizes or values of circles, dots, gray-tones and patterns. But numerous variables exist within the testing procedures themselves which have not been examined. Hence, the author uses graduated circle maps in analyzing two variables: short term memory response and task orientation. Ninety-six subjects were tested on their ability to recall or recognize a mapped circle pattern given one of three instructional levels.

Several statistical methods were utilized to evaluate the data. The results indicate that there is indeed a significantly large difference between recall and recognition and a smaller but nonetheless significant difference between very specific and non-specific instructional sets.

DEDICATION

To my parents for their continual support and without whom none of this would have been possible.

ACKNOWLEDGMENTS

I gratefully acknowledge the advisement and assistance received in various forms throughout my program from the following people: Professor Richard Groop who provided the original idea for the topic of this paper, numerous suggestions, much patience and an empty red editorial pen; Professor Richard Smith who provided additional ideas and guidance as well as the impetus for my attendance at Michigan State University; Professor Bruce Pigozzi, without whom, the statistical analysis of my data would have been a disaster; and Professor Dieter Brunnschweiler who helped maintain my interest in Remote Sensing such that my dissertation will fall under that heading. I also wish to thank the Department of Geography for the teaching assistantship awarded during all of my attendance.

On the lighter side, many of you probably realize that graduate students maintain strange hours and eating habits. In that vein, I must thank Jim Johnson for teaching me that while researching and writing into the late hours of the night, one can survive on nothing more than sandwiches of choke and slide.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Figure		Page

Chapter I

Introduction

Over the last 100 years, geographers and others have increasingly relied on thematic maps for the communication of spatial distributions. The thematic map's "...main objective is specifically to communicate geographic concepts such as the distribution of densities, relative magnitudes, gradients, spatial relationships, movements, and all the myriad interrelationships and aspects among the distributional characteristics of the earth's phenomena (Robinson & Sale, 1969, pp. 10-11)." But this objective is not achieved if the map-reader does not understand the data or misinterprets it. Thus, in order to effectively communicate information through maps, cartographers have attempted to standardize symbols, to increase the accuracy of data bases, and most importantly, to understand the process through which spatial information is transmitted from the map to the reader. This latter area of concern has included many tests of map readers' perceptions of cartographic symbology such as circle sizes, gray-tone values, dot densities and pattern correspondence. The present study is directed toward the methodology involved in cartographic perceptual testing.

Problem Statement

Cartographic perceptual testing methodology usually involves the use of psychophysical techniques. In most non-cartographic psychophysical testing, investigators conduct highly controlled experiments whereby only one or a few variables are scrutinized. This allows an evaluation of responses to simple stimuli and leads to more certainty in analyzing results. Cartographic perceptions, however, are usually much more complex because of the presence of a complex stimulus--the map. Experiments are less tightly controlled and results are more tentative. But most importantly, the methodology employed in the experiment is more likely to effect the outcome of the psychophysical test.

One may assume that if the subject's information output does not equal the cartographer's information input, then error is being introduced into the testing procedure by either the map maker or the map reader. In other words, "the act of testing, per se, appears to alter the behavior it sets out to measure (Cooper & Monk, 1976, p. 133)." The variables which induce this error include: (1) the complexity of the test maps; (2) the map reader's prior knowledge (environmental perception) of the mapped area and/or topic; (3) the reader's ability to learn over the course of successive testings; (4) the instructional set (task orientation) assigned to the reader before and during the test; and (5) the reader's ability to either recall or recognize (short term memory) patterns or symbols during the test. To date, little empirical research has been conducted in cartography to evaluate the effects of these variables on the results obtained in cartographic psychophysical testing.

Purpose

The focus of the present study is to examine two of the above variables, task orientation (specificity of instructions) and short term memory. (Note: Technically, short term memory is immediate memory; but since all of the variables pertaining to it are present in this study and since the use of the subjects' long term memory store was avoided, the term short term memory will suffice here.) These variables are scrutinized in the context of pattern recognition and reproduction on thematic maps.

Four critical questions concerning these variables can be raised: What are the factors which influence the map reader's ability to recall or recognize a pattern? How do these factors affect the reader's short term memory? Will the changes of specificity in the instructions create significantly different map reader responses? And will tests involving recall and recognition produce significantly different results? The first two questions are addressed in this chapter while the latter two are dealt with in the following chapters.

Importance

Prior to considering the above questions, the importance of task orientation and short term memory of map patterns should be addressed by examining the psychophysical studies that have investigated and utilized these procedures. Within cartographic perceptual testing, apparently task orientation greatly influences the amount and type of information accessed by a map reader. Most cartographers assign specific tasks to their test subjects, but few have empirically studied the effects of the level of instructional set on map reader responses. Generated responses from various types of instructions may simulate the information transfer in a "normal" map reading situation. If a map reader merely conducts incidental viewing of a map in an article, the amount of information accessed by him is less than it would be if the author directed the reader to the map or to particular portions of the map. In the same way, given non-specific instructions, the amount of information accessed by a test subject is less than it would be if the testor very specifically directed the subject to the pattern or symbols on the map or indicated a later task to be performed by the subject.

The importance of recognition and recall can be seen in most cartographic psychophysical tests since the comparison and discrimination between individual symbols or patterns invariably involves the use of one of these techniques (Flannery, 1956; Williams, 1956; Muehrcke, 1969; Olson, 1970). While recognition has been widely employed in cartographic testing, recall has been limited to test subjects producing cognitive maps of their environment and a single test in

cartography in which the subjects attempted to reproduce a stimulus pattern (Downs & Stea, 1973; 1977; Steinke, 1975). But recall is conceptually more important then recognition: In a map reading situation, the reader seldom refers back to a previously viewed map; instead, he probably relies upon his recollection of the map when viewing a secondary stimulus map. Thus, unless the maps are adjacent to one another, recall may assume the dominant position in common map reading tasks.

Next, one may well ask: Why examine pattern recognition or recall and not that of symbols? The dichotomy between the study of map patterns and symbols is expressed in the following passage:

> Maps are communicative devices designed to display spatial information in a two dimensional format. To some map makers and map users these displays are considered to be aerial data banks or storehouses of a myriad of separate and isolated facts. Other map makers and map users turn to the map as a communicative device because the two dimensional format allows them to display, and see, the new information which derives from the juxtaposition of sets of symbols (Jenks, 1975, p. 311).

The latter type of map information transfer, i.e., the map reader's ability to recall or recognize the pattern as a whole is the most important in thematic mapping. In fact, the purpose of the thematic map has been defined as that which communicates concepts, not data--the map conveys a pattern, not the components of the pattern (Gerlach, 1971, p. 194). And the map reader can perceive the pattern "as

a simultaneous whole on the basis of the interrelation of all of its parts held together in one immediate representation (Blumenthal, 1977, p. 71)."

For one to examine patterns, and not the individual symbols within the patterns can be seen in that the map reader seldom searches for value-size relationships between symbols. One might also reason that the map reader tends to group the data into simpler patterns (or rejects it entirely if it is too complex) so that it may be more easily remembered for future use. This mental generalization could be due to the reader's indifference to the task of enumerating symbols or to the lack of time spent by the reader on the map.

Recent eye movement studies have tended to confirm the idea that map readers do not ordinarily evaluate individual symbols. For instance, Dobson (1977) noted that the reader spent little time looking back and forth between the legend and the various circles on his test map to check the values. This is not to say that cartographers do not have to worry about the perceptual accuracy of individual symbols; on the contrary, one must still present perceptually accurate symbols because only they will result in perceptually accurate patterns. Fortunately, numerous studies have been conducted in which map readers were tested on their ability to discriminate and assign values to individual circles (Flannery, 1956; Meiheofer, 1979) or to individual gray-tones (Williams, 1956; Kimmerling, 1975). Even so, one must remember that the reader will mentally generalize; that is, forget or coalesce part of the pattern or combine some of the symbols.

Hence, a number of cartographers have considered pattern analysis an important part of map communication. Castner (1964), Jenks (1975) and Steinke (1975, 1979) have attempted to evaluate the visual comparison and reproduction of graduated circle patterns. Still other cartographers have examined the characteristics of patterns on choropleth maps. Several of these people (Olson, 1970, 1972; Monmonier, 1975; Lloyd and Steinke, 1977) set out to measure the effects of class interval systems on the visual correlation (pattern recognition and comparison) of choropleth maps. Other studies (Muehrcke, 1969; Monmonier, 1974; Olson, 1975; Muller, 1976) have been solely concerned with the visual analysis of choroplethic pattern complexity. All of the above studies examined either the subjects' ability to discriminate, compare or reproduce patterns. As a unit, these studies illustrate the manner in which map patterns are processed in the human perceptual system.

Other Variables

Pattern analysis aside, a number of uncontrollable factors are present within the test subjects which may influence their recall or recognition capabilities as illustrated in the following passage:

> We have asked the students to produce external representations as sketch maps. Can we judge similarity by comparing data taken from the

> sketch maps with similar data from a cartographic map? If this is an attempt to solve the accuracy question, the answer is no. People vary widely in simple graphic abilities. Age affects basic manual skills involving eye-hand coordination. Both the young and the old differ from our college students. Even discounting this age factor, we have the problem of differential training in both the rationale for and the manual reproduction of sketch maps. Some people, notably artists, architects and geographers are trained in technical graphic skills. One would expect that the mechanical production and accuracy of the sketch maps would reflect this training...Styles of training and thinking affect representations, even if both people are the same age, experience, skill, training, the question of the similarity of cognitive maps cannot be answered definitively. The nearest that we can come to such a goal is as follows: Parts of our cognitive maps are common to all or most members of a large group of people, parts are common to a subgroup of people, while still other parts are unique to each person (Downs & Stea, 1977, pp. 100-103).

Equally troublesome is the problem of attitude on the subject's part. Since volunteer subjects may bias the results, "non-volunteers" are usually rendered from classroom settings, however, no one is ever forced to participate. The use of non-volunteers, however, probably involves the use of some students who care little about the experiment. Since this study is only concerned with the responses of the "average" map reader, the effects of the above variables upon the individual reader's expression of short term memory may be offset by the reactions of other subjects in the same test group. Hence, while being aware of the aforementioned uncontrollable variables, they are relegated to a position of minor importance within this study.

Of greater importance, however, are the variables that are intimately related to task orientation and short term memory and their effects on the visual comparison or reproduction of patterns on maps. This relationship may be transformed into the following cause and effect perceptual processes: (1) the test results are dependent upon short term memory and its transmittors (eyesight and/or fine motor skills); (2) short term memory is dependent upon the amount of information received; (3) the amount of information received is dependent upon eye movements; and (4) the eye movements are in part dependent upon the task orientation (instructions) as initially given to the subject, and in part dependent upon the pattern's complexity, the amount of time spent viewing the map, the length of time delay after viewing the map before starting the recognition or reproduction task, and the subject's prior knowledge of the mapped topic and/or area.

These processes may be addressed individually. First, a number of studies have established that the type of instructions given in a psychophysical test do indeed affect the subject's performance. Tversky (1973) noted that if the subject was previously informed of the type of task to be performed, be it recognition or recall, he would perform better than one who was not informed or incorrectly informed. In all cases, however, the recognition tasks had a higher response probability than the recall tasks (Figure 1). Even though Loftus and Loftus (1976) agreed with this statement, they found that the difference between incidental (nonspecific) and intentional (specific) types of directions is negligible in recognition tests while the difference is significant in recall tasks (Figure 2).

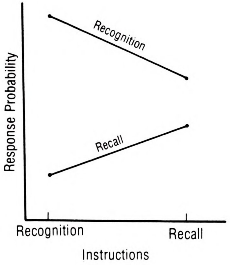


Figure 1 Results of Tversky's Tests

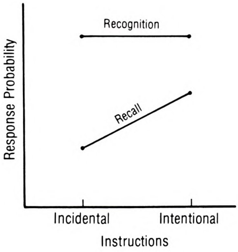


Figure 2 Results of Loftus and Loftus' Tests

Warrington and Ackroyd (1975) tested three tasks: (1) no orientation, (2) relevant orientation and (3) non-relevant orientation. They found no significant difference between the no orientation and the non-relevant orientation tasks but the relevant orientation task produced a significantly better performance on the subjects' part (Table 1).

Table 1

Warrington and Ackroyd's Test of Orientation Tasks

(Mean error)

	Non-relevant	None	Relevant
Words	9.90	12.60	4.75
Faces	10.00	11.10	7.00

DeLucia also tested three different tasks, but he evaluated the directions in terms of eye movement patterns across a map. His tasks included: (a) no orientation, (b) general orientation, and (c) specific orientation. In the no orientation task, the subjects primarily conducted free scans of the map. In the general orientation task, the subject was "only told what to look for but not where and therefore could not narrow down his search area on the basis of the assigned task prior to commencing his scan," i.e., the subject conducted a free scan until he found what he thought he was looking for. The specific orientation task was one of comparison: the subject was told what to look for and where to look on the map and therefore the areas not fixated upon "are avoided by the viewer because they do not contain information relevant to his problem at the moment (1974, pp. 239 and 241)." Unfortunately, DeLucia did not conduct any quantitative analysis of his data.

Steinke conducted a test whereby the subjects knew about the eye movement recording part of his experiment before it began because of an earlier class presentation and a brief introduction when they first arrived for the experiment, but they were not aware that they would have to reconstruct the target map body later. That is, Steinke purposely designed his test as non-specific because he wanted to define what people do under a free look situation even though relatively little real map reading occurs in this way. Likewise, telling the subjects before or during the experiment that they would have to reproduce the map body later would no doubt have increased motivation but at the same time would have resulted in very different map reading activity since few people read maps with the idea of reproducing them later (correspondence with Steinke, March 1979).

On the other hand, one might reason that for a specific orientation task, the test subject will probably scan a test map in much the same way that the stimulus map was scanned. In other words, if the search pattern for familiarization is repeated for recognition, then the subject will have an easier and faster job of recognizing the correct pattern (Norton and Stark, 1971; Whiteside, 1978). Hence, in preparing a test, one should assign one of three tasks for the subjects: non-specific, somewhat specific and very specific. The specificity of the instructions will at least partly determine the test results.

Aside from the instructions given, a number of other variables affect the subject's short term memory. First, the complexity of the pattern must be considered. Phillips (1974) discovered that as the complexity of the pattern increases, the error in short term memory increases. The error may be due to the fact that as pattern complexity increases, the short term memory store becomes overloaded whereby parts or all of the pattern may be forgotten or may not even be initially absorbed (Herriott, 1974). In fact, Kaufman, et al.,(1949) provided evidence which indicates that a subject cannot perceive the pattern as a whole beyond approximately eight elements in the pattern. But French (1954) stated that target recognition improved with an increase in target complexity (possibly due to the "uniqueness" of the more complex patterns) and became worse with an increase in visual noise. French's data suggested, however, that this function is negatively accelerated.

The recognition of pattern types seems to involve some additional influences of complexity. For example, Fitts et al. studied the effects of redundancy, i.e., an inverse measure of complexity, and their results indicated that "there is no simple relationship between the redundancy of figures and pattern recognition. The introduction of redundancy may either facilitate or hinder pattern recognition, depending on the way in which it is introduced (1956, p. 10)."

Fitts, et al. also found that "random figures were recognized more rapidly than were constrained figures. Symmetrical and vertically oriented figures were recognized more rapidly than were single or double assymmetrical figures or horizontally oriented figures of the same complexity (1956, p. 10)." And during the time that the subject is viewing the stimulus map and during the delay time between the initial exposure and the recall or recognition task, it appears that the pattern is organized into a coding scheme along horizontal and vertical axes within the human visual system (Dodwell, 1970, pp. 112-13). Given the above, perhaps the short term memory coding scheme is stronger along the vertical axis than it is along the horizontal axis.

Obviously, the construction and arrangement of circle patterns influences one's ability to recognize that pattern. Consequently, in constructing a test map, one should try to create a pattern of "medium" complexity, i.e., a pattern that has a clearly recognizable shape but has no symmetry and does not contain too few or too many circles, while at the same time keeping that pattern "realistic" in appearance.

The amount of time that the subject spends looking at the map may also affect the amount of information that is being stored. Dobson observed that "the speed at which information was accessed varied from subject to subject (1977, p. 53)." And Steinke found that there was "little relationship between how much time a person looks at a map and his ability to reproduce it (1975, p. 220)." (Note: the amount of time that the subjects were exposed to the stimulus map in each test varied from 12.99 to 32.65 seconds and 12.6 to 40.3 seconds, respectively.)

The passage of time between the end of the presentation of the stimulus map and the start of the recall or recognition test may also be a factor influencing what is retained in the short term memory store. "In general, short term memory is observed as a temporal constraint on recall capacity and not as a constraint on recognition capacity, recognition being a distinct and powerful long term memory ability. An object briefly seen can be totally unavailable to recall, yet days later its recurrence may be recognized immediately (Blumenthal, 1977, p. 72)." Apparently, the specific impressions of a circle pattern disintegrates with the passage of time and one should therefore minimize the delay time in the testing process.

Taken one step further, "the last few items presented (or looked at) tend to be recalled first (Baddeley, 1976, p. 103)." But "elapsed time, per se, does not affect retention at all, and that the retrievability of memory is solely a function of how much interference has occurred at the time of recall (Blumenthal, 1977, p. 72)." Like time, interference is less likely to affect recognition than it will recall (Wicklegren & Norman, 1966, p. 346). But neither recognition nor recall are seriously affected by interference in the short term memory store (Herriot, 1974, p. 6).

Mandler (1972) noted that the proportion of information (pattern) which was presented but not recalled had not been organized by some people; and he stated that some of the information recalled was not originally presented. Had his tests included maps, this "new" information was probably derived from past (long term) memory or environmental perception of the mapped area or topic and organized into the stimulus map pattern. Finally, Tulving and Pearlstone (1966) discovered that there was much more memorized material available at the time of recall than can actually be retrieved. This discrepancy may be due to and widened by the use of fine motor skills and eyesight (each of which may or may not be well developed), in the reproduction and recognition of circle patterns, respectively.

Summary

Given that this thesis is focused upon two important test situation variables, task orientation and short term memory, one finds that not only do these two variables affect test results but a number of other variables also influence the results either directly or indirectly through the above two variables. Many of these variables can be controlled or at least held constant while others cannot and one must therefore be aware of their presence when evaluating the test data. Overall, cartographic psychophysical test results are not merely influenced by the instructional set or the type of memory response being elicited; instead, a myriad of complex variables are present which effect both each other and final test results.

Chapter II

Hypotheses and Methodology

Introduction

This section specifically addresses the manner in which three levels of task orientation (non-specific, NST; somewhat specific, SST; and very specific task, VST) should hypothetically affect the test results with either the goal of pattern recognition or pattern reproduction of a graduated circle map. A detailed explanation of the testing methodology will follow since, after all, one of the primary foci of this thesis is to "test the test."

Hypotheses

Based upon the considerations noted in Chapter I, one may make the following general hypotheses:

I (a) Error will be less for recognition tasks than for recall tasks (Figure 3).

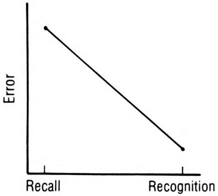


Figure 3 Hypothesized Difference of Error Between Recall and Recognition.

(b) Error will decrease from a non-specific to a somewhat specific to a very specific instructional set (Figure 4).

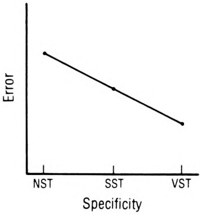


Figure 4 Hypothesized Difference of Error Between Instructional Sets.

(c) Therefore:

(1) The greatest amount of error by the test subjects will come from those who are given a non-specific task requiring a reproduction skill.

(2) The least amount of error by the test subjects will come from those who are given a very specific task requiring a recognition skill (Figure 5).

	NST	SST	VST
Recall	Greatest Error (6)	(5)	(4)
Recognition	(3)	(2)	Least Error (1)

Figure 5 Hypothesized Progression of Error Terms.

II (a) The stimulus pattern will always be more consistently recognized than it will be reproduced (Figure 6).

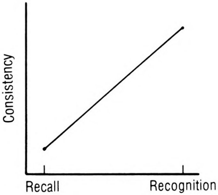


Figure 6 Hypothesized Differences in Consistency Between Recall and Recognition.

(b) The consistency with which the stimulus pattern is recognized or reproduced will increase as the task becomes more specific (Figure 7).

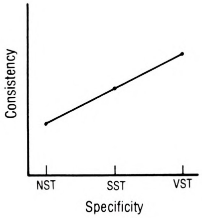


Figure 7 Hypothesized Differences in Consistency Between Instructional Sets.

(c) Therefore:

(1) The test subjects will exhibit the least consistency for a non-specific task requiring a reproduction skill.

(2) The test subjects will exhibit the greatest consistency for a very specific task requiring a recognition skill (Figure 8).

	NST	SST	VST
Recall	Least Consistency (1)	(2)	(3)
Recognition	(4)	(5)	Greatest Consistency (6)

Figure 8 Hypothesized Progression of Consistency.

These hypotheses may be addressed individually. Based upon the results of Tversky (1973) and Loftus and Loftus (1976), one would expect a higher response probability, or a lower error, for all recognition tests than for any recall test. This difference in error terms between recall and recognition is probably due to the sheer difficulty of recalling and reproducing a mapped circle pattern. And if the test map is to be in any way realistic in appearance, the number of circles will probably exceed the eight element limit that Kaufman, et al. (1949) defined as that which can be efficiently recalled or recognized.

Concerning the instructional set, the conclusions of Loftus and Loftus (1976) and Warrington and Ackroyd (1975) indicate that the subject should perform significantly better given a specific task than if they were given a non-specific task. And Loftus and Loftus found that this difference was much more pronounced in recall than in recognition tests. The difference between either the NST or the VST and the SST, however, may not be significant. Even though this difference is not pronounced, the results, nonetheless, should show a progression in the error terms as hypothesized.

The reader may note that the slope of the lines illustrating the differences in mean error between recall and recognition (Figure 4) is steeper than that between the different levels of task orientation (Figure 5). Given the above prior research, and due to the fact that reproduction of a circle pattern is more difficult than the recognition of that pattern, the different instructional levels will not produce as much error as the two memory tasks.

In regard to the subjects' consistency of response within each test group, one would expect an inverse relationship to that of the error terms (Figures 7 & 8). As with error, a significant difference in consistency between recall and recognition should be evident in the data analysis. A somewhat different reason may account for the discrepancy in consistency, however. In a recall test, given an infinite number of subjects, an infinite variety of patterns could be reproduced; but in a recognition test, the subjects are constrained to whatever choices the researcher gives them.

Neither the memory task nor the instructional set appear to have been quantitatively examined in terms of the subjects' consistency. But a descriptive measure of consistency was provided by DeLucia in his tests on task orientation: He recorded the subjects' eye movements after reciting a particular level of instruction; the narrowness of the area scanned decreased as the specificity increased (1974, p. 239; 241). In other words, one may assume that the greater consistency in eye movements over the stimulus map, the greater the consistency of the test subjects in recognizing or reproducing the target map. Finally, in following this line of reasoning, like error, the slope of the lines illustrating the differences of the subjects' consistency is steeper between the levels of memory task as opposed to that between the levels of instructional set.

Testing Methodology

Six different map oriented tasks were administered to six different groups of sixteen people yielding a total of 96 subjects. The subjects consisted of a variety of undergraduate and graduate geography students (majors and non-majors) at Michigan State University. None of the subjects were tested more than once. Given the range of students tested, the sample population appeared to approximate the "average" map reader, thereby reducing the effects of the problems addressed by Downs and Stea (1977, pp. 100-03).

The stimulus map was designed for simplicity, employing standard cartographic principles. The circle pattern on the map consisted of a selected assortment of 20 different-sized, non-overlapping circles. The map as such presents a pattern that is not easy for the subjects to recall while the task required is not impossible. Overlapping circles were not used because the element of overlap introduces additional unwanted variables into the test (Groop & Cole, 1978). Nor were county borders, a north arrow or a scale used on the map. Whereas county borders may provide a locational impetus, their presence creates additional background noise--the influence of which is hard to measure; on the other hand, the presence of the latter two items, being merely incidental to the map's message, adds only clutter to the map. However, a source and a legend (adjusted to Flannery's constant) were present on the map because not only are they commonplace but they also provide necessary information to the reader.

The stimulus map used in the tests, which appears in Figures 9a and 9b, covered an area and topic which presumably most students from the state of Michigan would be unfamiliar with, i.e., "State Park Attendance in Arkansas." Supposedly, a map of this sort reduces the effects of previous learning or experience on the test results. Hence, only in those areas of the map that the subject did not look at (terra incognita), would one expect great variations in the components of the pattern.

In each test, the subjects were given two envelopes: A and B. First, they were told to remove the map from envelope A. Then, one of the levels of instruction was

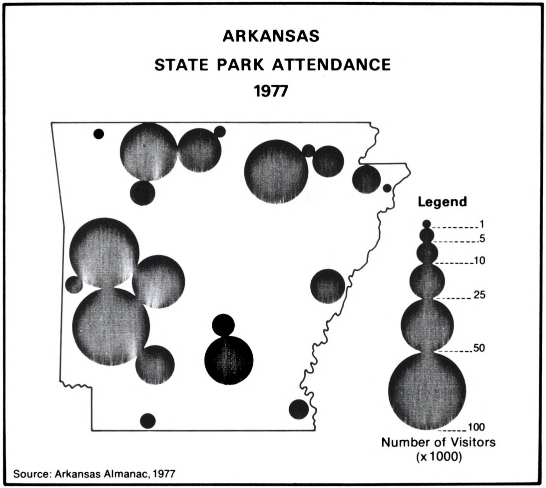


Figure 9a Stimulus Map Used in the Recall Tests (Reduced 10%).

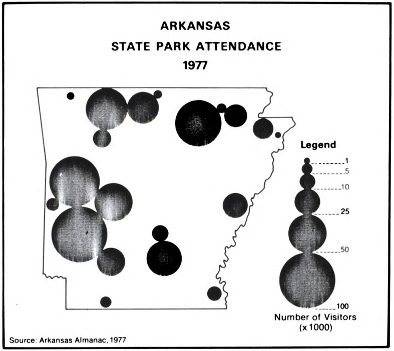


Figure 9b Stimulus Map Used in the Recognition Tests (Reduced 10%).

delivered to the subjects:

(1) NST -- "Look at the map."

(2) SST -- "Look at the pattern of circles on the map."

(3) VST -- "Study the pattern of circles on the map; a) later, you will be asked to rank a set of maps from most similar to least similar in relation to the one that you are looking at now (recognition) or (b) later, you will be asked to reproduce the circle pattern that you are looking at now (recall)."

After the initial instructions were read, the subjects were given 30 seconds to view the stimulus map. At the end of that time, they were told to "put the map back in envelope A and place it off to the side." The final tasks were then read to the subjects. Clarity <u>and</u> expediency were required of the testor at this point because, as Phillips noted, short term memory decays rapidly after 10-20 seconds (1974, p. 284). Granted, the reading of the final task may be deemed interference by Brown (1958) or Winklegren and Norman (1966), but the instructions will be held constant for each of the recall tests and each of the recognition tests and thus the verbal instruction's effects should be uniform for each group. Regardless, as Herriot noted, interference does not seriously affect recall or recognition in the short term memory store.

The final tasks themselves can be broken down into two types:

Recall and Replicate

The subjects were instructed that "Envelope B contains a blank map (Figure 10), a slip of paper with a set of stick-on circles on it, and a knife. Please remove them carefully. From what you remember of the map that you saw earlier, very carefully remove all of the circles from the piece of paper and arrange them on the map until the circle pattern duplicates the pattern of the first map. You may rearrange the circles at your discretion."

The above instructions follow Brown's (1976) definition of recall in that the testor asked his subjects to generate a target(s) which matched the initial stimulus to the best of his memory. The response accuracy depended upon how complete the subject's information (determined in part by the instructional set) was of the target stimulus. Once the map is complete, the subject then decides whether or not he recognizes the pattern; if not, he may then rearrange some or all of the circles. Figure 11 illustrates a reconstructed pattern produced by one of the subjects in VST.

The use of stick-on circles in this study is an outgrowth of the work conducted by Steinke (1975). In his test, Steinke presented his subjects with a stimulus map containing 39 circles divided into seven classes. Later, the subjects

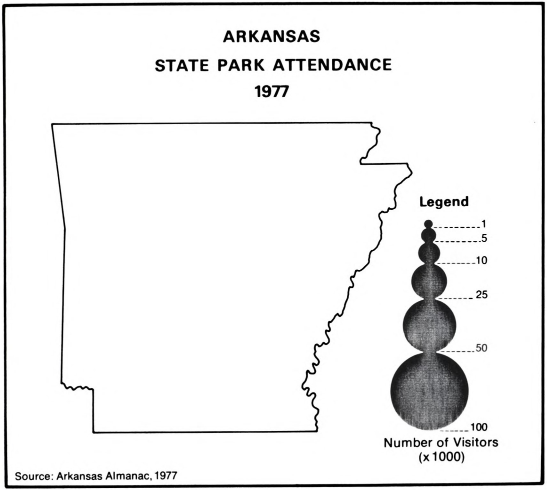


Figure 10 Blank Map Used in the Recall Tests (Reduced 10%).

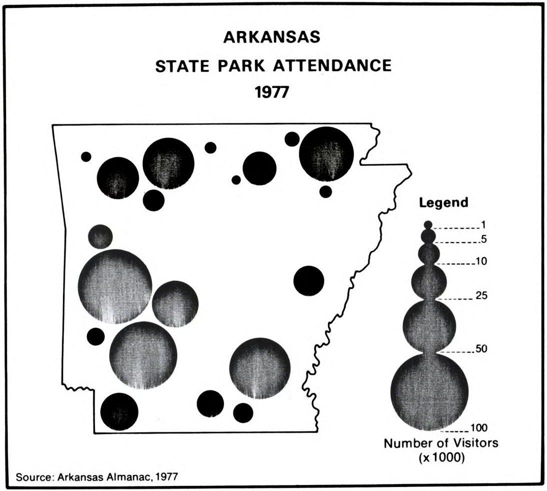
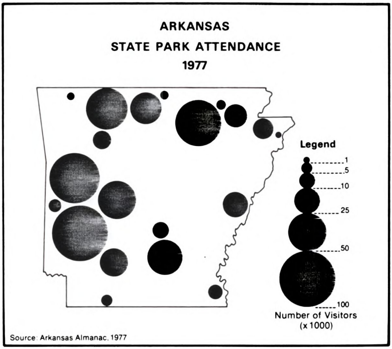


Figure 11 An Example of a Reproduced Map: Subject #4 in the VST (Reduced 10%).

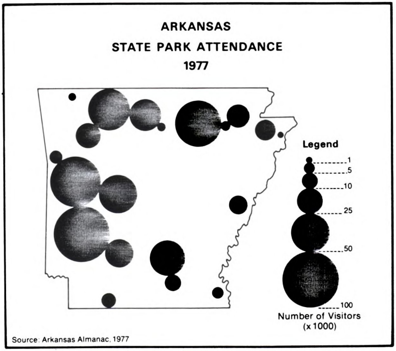
were asked to reproduce the pattern to the best of their memory using a set of adhesive-backed circles. Unfortunately, he encountered a problem in evaluating the reproduced patterns because he had let the subjects use more circles than were in the stimulus pattern. This procedure may have inadvertently allowed his subjects to inject circles into the pattern that were not originally present. Because different numbers and sizes of circles may have been used on the target maps as opposed to those circles on the stimulus map, a quantitative comparison between the two was difficult to make. A constraint was therefore imposed on the subjects in the present test, i.e., the same number and sizes of circles were placed on the slip of paper in envelope B as were on the stimulus map. In this way, the "movement" of each circle and the pattern as a whole could be quantitatively measured.

Recognition

The subjects were instructed that "envelope B contains a set of eight maps similar to the one that you just saw, plus a slip of paper numbered 1-8 (Figures 12a through 12h and 13, respectively). Please remove them. From what you remember of the map you saw earlier, arrange the maps from most similar to least similar. Record the symbols located in the lower right hand corner of the maps in the appropriate blanks on the slip of paper."

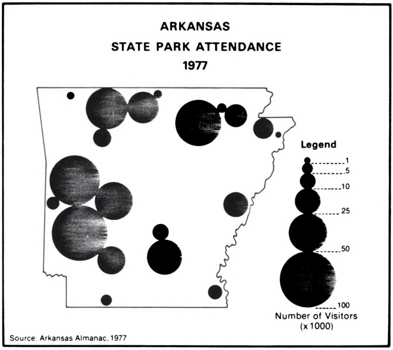


X a

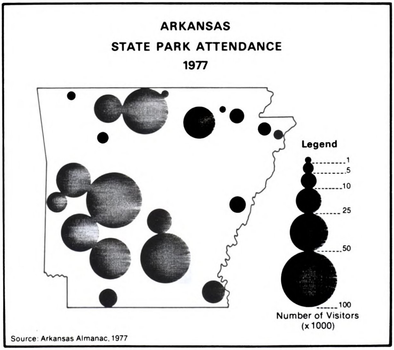


! b

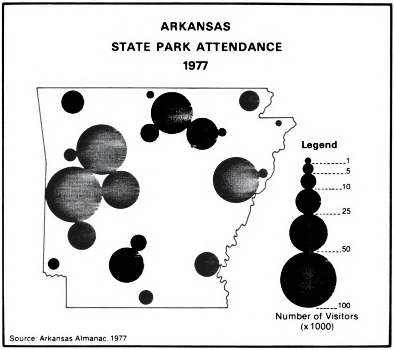
Figure 12 Distractor Stimuli Used in the Recognition Tests (Reduced 10%).


ARKANSAS
STATE PARK ATTENDANCE
1977
Legend
1
5
10
25
50
100
Number of Visitors
(x 1000)
Source: Arkansas Almanac, 1977

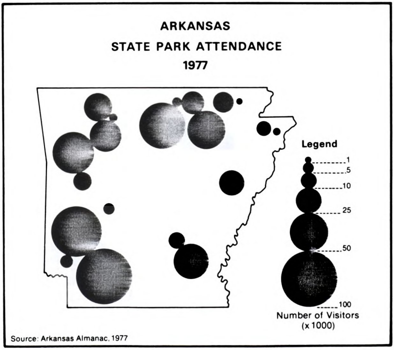
< c


ARKANSAS
STATE PARK ATTENDANCE
1977
Legend
1
5
10
25
50
100
Number of Visitors
(x 1000)
Source: Arkansas Almanac, 1977

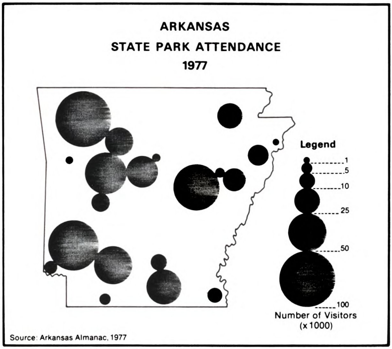
Δ d


ARKANSAS
STATE PARK ATTENDANCE
1977
Legend
1
5
10
25
50
100
Number of Visitors
(x 1000)
Source: Arkansas Almanac, 1977

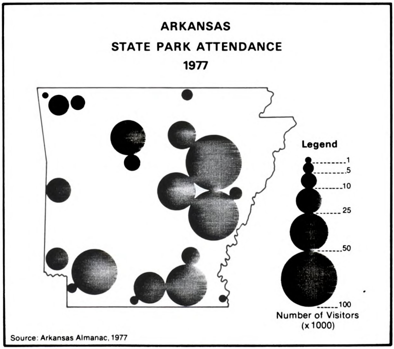
e


ARKANSAS
STATE PARK ATTENDANCE
1977
Legend
1
5
10
25
50
100
Number of Visitors
(x 1000)
Source: Arkansas Almanac, 1977

→ f


ARKANSAS
STATE PARK ATTENDANCE
1977
Legend
1
5
10
25
50
100
Number of Visitors
(x 1000)
Source: Arkansas Almanac, 1977

g


ARKANSAS
STATE PARK ATTENDANCE
1977
Legend
1
5
10
25
50
100
Number of Visitors
(x 1000)
Source: Arkansas Almanac, 1977

h

1.____________ Most similar

2.____________

3.____________

4.____________

5.____________

6.____________

7.____________

8.____________ Least similar

Figure 13 Slip of Paper Used for Ranking in the Recognition Tests.

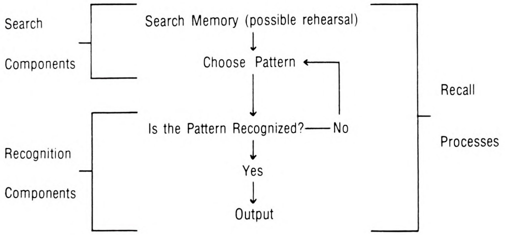


Figure 14 Mental Processes During Recall and Recognition (After Loftus and Loftus, 1976).

These directions were similar to Brown's (1976) and Loftus and Loftus' (1976) definitions of recognition in which one or more potential targets were presented to the subject who searched his memory, selected a pattern, and then decided whether or not that pattern most closely resembled the initial stimulus that he remembered (Figure 14). The response accuracy depended upon whether or not the subjects had "as much information (determined by the instructional set) as is necessary to discriminate the target stimulus from the distractor stimuli (Brown, p. 87)."

This type of test involves two additional subtasks on the subject's part: First, the subject conducts a discrimination task where he "judges whether the stimulus pattern seen is different from some other one or from some other set of patterns." Next, a judgmental task is performed where the subject "is required to assign the presented pattern(s) some value on a scale of judgment, such as size, complexity, attention value" or similarity (Hake, 1966, p. 146).

Since a rank-ordering of the patterns took place, two decisions were made before the test was given: (1) Symbols, and not numbers of letters, were used to prevent a biased rank-ordering of the pattern. (2) Because Neisser (1967) had found that test subjects did a faster and better job of recognition for dissimilar target arrays than for similar target arrays, eight similar target arrays were presented. While the use of eight similar patterns probably increases the distances between mean error terms for recognition and recall tests, one should expect that said presentation is a more realistically difficult assignment than the use of dissimilar patterns.

At this point, one must also note that there are two features in recognition tests which are likely to encourage the use of recall:

(a) if the recognition test requires the subject to search for a single target amongst many choice items (in this case, eight), his best strategy may be to recall the target first (curing the search memory phase of Figure 14) and then to search for it.

(b) in a short term memory experiment, the subject may still be rehearsing the material (also during the search memory phase) at the time the test was administered. Since rehearsal involves repeated recall, recall inevitably mediates recognition in the test (Brown, 1976, pp. 2-3).

But rehearsal is dependent upon the specificity of the instructions, i.e., the more specific the task, the more likely active rehearsal will occur.

At the conclusion of each test, the subjects were asked informally, "Did you remember any circles in particular?" Most of the subjects indicated that they remembered the locations of the larger circles and groups of circles better than smaller and isolated circles. They were also asked, "Did you remember a shape to the pattern?" This question elicited a variety of responses including the shape of a C, an obtuse triangle, a hook, and so forth. The responses to the first question qualitatively verifies the impression that relatively large amounts of blackness attract the eye while the responses to the second question merely indicates the variety of memory patterns (organization) present in different people.

Chapter III

Data Analysis

Introduction

In order to quantitatively evaluate the results, the x,y coordinates of all the circles in each pattern reproduced in the recall test or chosen in the recognition test, plus the stimulus pattern, were plotted on graph paper. Since the maps in the recognition tests were smaller than those in the recall tests, the x,y grids were adjusted accordingly so that relevant comparisons could be made between the groups. From this base, a series of statistical analyses were conducted in order that the data could be meaningfully interpreted.

These analyses included: (1) frequency of pattern choice in the recognition test; (2) the mean squared distance between the point locations of the stimulus circles and the chosen or reproduced circles in both tests; (3) the difference between homologous circles after adjusting for orientation; (4) the mean deviation of the percent black per x,y quadrant; (5) the use of the circular normal distribution to obtain mean distance from the origin and the mean angular deviations; (6) a weighting factor (area of the circles) applied to 5; (7) centroids of the patterns and their associated vectors from the stimulus centroid plotted; and (8) a weighting

factor applied to each of the centroids in 7.

A logical question at this point of the study might be: Why use all of the aforementioned statistical procedures? The answer is simply because no one statistic adequately explains all of the variations in the data. While all are related, each analysis describes a different aspect of the data; thus, each is used to compliment one another so that any conclusions that are made have sufficient quantitative backing. A discussion of the use of these techniques and interpretations of the results follows.

Before proceeding, a brief examination of the hypotheses is in order. If one attaches significance levels to the difference in means for each group using a two-tailed Student's $\underline{t}$ test, then one might realistically formulate the following specific hypotheses: (1) for all test groups, one can predict a significance level of $a = .01$ $(t = 2.75)$ between the mean error terms of recall and recognition. (2) Based upon earlier research and given the instructions that were read to the subjects, there should not be a significant difference (although a progression should exist) between the mean error terms of the non-specific and somewhat specific and between the somewhat specific and very specific instructions. (3) But a significance level of $a = .05$ $(t = 2.042)$ should exist between the non-specific and the very specific instructional sets.

Frequency of Choices--Recognition

First, an examination of the relative rank that the subjects assigned to the stimulus map and the distractor stimuli is now in order. This particular measure was employed to assess the frequency of responses for each of the test maps. Figures 15a, 15b and 15c show that the stimulus map was ranked no "worse" than third by any of the three instructional groups. Clearly evident is that very little consistency exists within that range for the NST group.

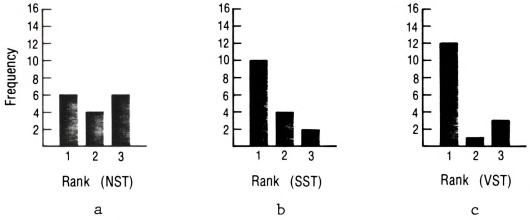


Figure 15 Rankings for the Stimulus Map.

Comparing Figures 15a and 15b, one notes that as the task becomes more specific, the number of subjects assigning a first order ranking to the stimulus map increased at the expense of the third order rankings. Further, a distinctive downward trend now appears for the SST group as opposed to the rather even spread of rankings for the stimulus map within the NST group. These rank frequencies indicate that

not only do somewhat specific instructions generate more consistent results than do non-specific instructions, but they also convey the idea that progressively fewer people assign the stimulus map to the less similar rankings.

Such is not the case when comparing Figures 15b and 15c. Whereas the frequency of first order choices in the VST is greater than that of the SST, the difference between the two is not as great as expected; in fact, it is less than the difference between NST and SST. And the progression through the rankings does not get successively lower for VST.

Before postulating as to the reasons why, one should look at Figures 16a, 16b and 16c which illustrate the first order rankings for all maps. No more than five of the eight maps given were chosen in all of the test groups (Figures 12a-12e). One may assume that the patterns of the three maps not chosen must be, in the visual sense, sufficiently different from the stimulus map. Conversely, one might also assume that the patterns of the maps that were assigned to the first order are visually "close" to the stimulus map pattern. The term "visual" qualifies the last two statements because, as discussed later, visual rankings are not necessarily equivalent to mathematical rankings.

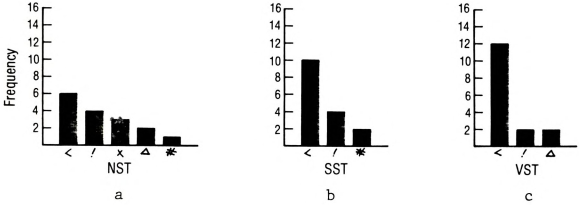


Figure 16 First Order Rankings for all Maps.

Figure 16 illustrates a variety of responses (non-consistent) to NST. On the other hand, in response to SST and VST, not only do a greater number of subjects choose the stimulus map over all others combined, but fewer of the distractor stimuli are chosen for these two tasks as well. Like Figure 15, however, Figure 16 also shows a greater difference in responses between NST and SST than between SST and VST. One might expect, though, that the error terms calculated for VST are lower than those for SST. But the patterns chosen by the subjects in VST dictate otherwise. Because two of the subjects in VST chose the map symbolized with a triangle (Figure 12d), most statistical measures will calculate the error in that group to be greater than the error in SST. In other words, the progression from high to low mean error terms should be NST-VST-SST, respectively. Had the test sample been larger, the expected progression of error terms might have materialized. These

are important points to remember later on when the output of the other statistical procedures are analysed.

Mean Deviation--Percent Black Per Quadrant

In this particular statistical test, an x,y grid was laid down on top of each of the reproduced and chosen maps and the percent black per quadrant was calculated for all maps. The absolute deviation of each quadrant of the maps was derived from the difference between the percent black per quadrant of the stimulus map and all other maps (e.g., mean deviation of quadrant 1: $Q_1 = \frac{/E(\bar{X}s_1 - \bar{X}i_1/}{/ \quad 16 \quad /}$). The mean absolute deviation was then calculated for each test group ($\frac{E(Qi)}{4}$). In effect, what this statistic measures is only the percent black per quadrant; it tells us little about the positions of the individual circles or the patterns. Since a number of small circles in one quadrant can equal one large circle in another quadrant, the interchange of the two would probably result in little or no change in the respective deviations. While a reproduced pattern may look very different in comparison to the stimulus map, the percent black per quadrant could conceivably be equal.

The data for each quadrant indicates where the average map reader in a particular test group remembers clustering or blackness to occur and the statistic itself reflects that memory, i.e., the larger the mean deviation, the less the subject remembers about the overall pattern. First, Table 2

illustrates that while both the absolute total and mean deviations are less for recognition than for recall, only relatively small variations from the stimulus map occur in any one quadrant for any one test. Overall, Table 2 indicates that the subjects approximately remember the relative amounts of black in each quadrant, regardless of the instructional set. The absolute mean deviation (in parentheses) points out the fluctuations among the different task levels for each quadrant. These fluctuations are rather hard to interpret; however if one examines Figure 17, the overall error for each instructional set is clearly outlined. As expected, the differences between NST and SST for both recall and recognition are less than the differences between SST and VST. But the difference of means test revealed that NST and VST for both recall and recognition are not significant at the a = .05 level, although the difference between recall and recognition are significant at the .01 level.

Table 2 Absolute Total and (Mean) Deviations

Quadrant		1	2	3	4
Stimulus		17.7	7.1	33.9	31.0
Recall	NST	22.2 (5.74)	9.6 (5.14)	26.9 (7.86)	31.7 (3.24)
	SST	20.9 (5.33)	8.4 (3.76)	30.2 (5.46)	30.4 (6.97)
	VST	19.7 (3.98)	8.9 (4.47)	34.9 (2.61)	27.0 (5.90)
Recog.	NST	16.8 (1.80)	8.7 (1.70)	33.1 (1.80)	32.9 (2.15)
	SST	18.5 (1.06)	7.3 (0.90)	32.1 (1.80)	33.2 (2.20)
	VST	16.4 (1.26)	8.2 (1.16)	34.2 (0.60)	31.6 (0.70)

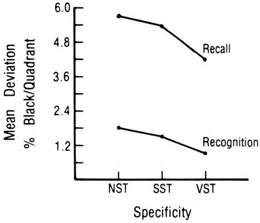


Figure 17 Mean Absolute Deviations of the Percent Black Per Quadrant.

Mean Squared Distance

Two related measures of error were on techniques suggested by Sneath (1967). Those measures were used to produce similarity indices, based upon the movement of the individual circles, between the resultant patterns and the stimulus. More specifically, the first of these techniques measured the sum of the squared distances between the circles in the stimulus pattern and the respective circles in the recognized or recalled patterns. In essence, this program performed a comparison between those patterns and the stimulus pattern and yielded the sum of the squared distances which are interpreted as dissimilarity indices. The mean index for each test group was then calculated thereby producing a measure of similarity between the test groups.

The large differences in error terms between the recall and recognition groups can be easily seen in Figure 18. As hypothesized, the difference between NST and SST is less than that between SST and VST for recall. With recognition, on the other hand, VST registered a higher degree of error than SST for reasons already explained. A significant difference is present between the means of recall and recognition at the .01 level. NST and VST exhibited a significant difference for recall at the .1 level while no significant difference could be detected between any of the instructional sets for recognition.

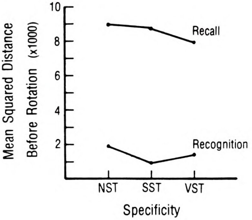


Figure 18 Test Group Values of the Mean Squared Distance Before Rotation.

The second of these techniques was attempted to bypass several difficulties associated with the above measure and other analyses conducted by Jenks (1975) and Steinke (1975). In both of these studies, subjects were asked to visually compare mapped circle patterns and to assign a value of similarity to them (1-7 and 1-5, very dissimilar to very similar, respectively). While their results may have been overtly subjective, they provided a basis for comparison between perceptual and mathematical pattern similarity. Jenks also tried to evaluate the similarities between circle patterns by means of a correlation grid. Unfortunately, the correlation between the stimulus pattern and the resultant pattern was either very low or zero. This problem may also exist to a lesser extent in the analysis conducted below.

In an attempt to solve this problem, while at the same time maintaining an objective measure of pattern comparison, the following steps were performed: the x,y locations of all the circles, stimulus and resultant, were transformed into standard deviation units. Each resultant pattern was then rotated over the stimulus pattern until the "best fit" between the patterns was achieved as measured by the least squared distances between homologous circles. The mean deviation of each group was also calculated to compare test groups.

While compensating for minor rotational errors on the subjects' reconstructed patterns, this technique introduced a different problem: it rotated the patterns and examined them out of the context of the map. For example, if a subject had chosen the map symbolized by a star (Figure 12h) as the most similar pattern, the index between it and the stimulus pattern was similar to the stimulus pattern (with minor variations when rotated 180 degrees). In other words, this technique overcompensates for the subjects' mental errors in orientation and would subsequently assign a similarity index between patterns that may be significantly different from an index created by a visual comparison of the patterns. That problem aside, when the mean squared distances after rotation were plotted for each test group (Figure 19), surprisingly, the shape of the plots was basically the same as those on Figure 18. The similarity between Figures 18 and 19 may indicate that the average test

subject performed very little mental rotation of the stimulus pattern.

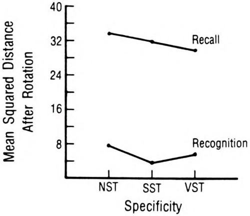


Figure 19 Test Group Values of the Mean Squared Distance After Rotation.

A significant difference at the .01 level was found between recall and recognition. The relative significant difference between NST and VST for recall was strengthened, in comparison to the mean squared distance before rotation, in that their difference was significant at the .05 level. No significant differences were noted between any of the recognition tasks.

Circular Normal Distribution

Based on research conducted by Reyment (1971) and Mardia (1972), several measures of positional error of the reproduced or recognized circles were made. The goal of both of the previous studies was to determine whether or not

the data contained in a circular distribution was normal. Of primary interest here, however, are the statistics used to arrive at a determination of normality: the distance from the origin which reveals how close to the stimulus the circles in the resultant patterns come, and the angular deviation which indicates a test group's angular consistency in the positioning and locating of the circles. The distance (r) from the origin (the point of origin is arbitrary, but it must be the same for all circles) is found by the following formula:

$$r = (\bar{X}^2 + \bar{Y}^2) \text{ where } \bar{X} = \frac{E \text{ (cos ai)}}{N} \text{ and } \bar{Y} = \frac{E \text{ (sin ai)}}{N}$$

(Reyment, 1971, p. 23). This statistic establishes the length of the mean vector for a circle. The length is a unit measurement, i.e., as $\underline{r}$ approaches 1.0, the mean circle location approaches that of the stimulus. The mean vectors for the patterns were also calculated and plotted in Figure 20, illustrating the differences between recognition and recall and between the instructional sets for $\underline{r}$. This graph also points out the inverse relationship between $\underline{r}$ and error.

The mean angular deviation(s) is essentially the standard deviation of a test group measured along the circumference of a circle instead of a straight line. It is defined as: $s = 2\ (1-r)$ (Reyment, 1971, p. 27). This computation yields the value of $\underline{s}$ in radians. By converting radians to degrees, one arrives at the mean angular deviation,

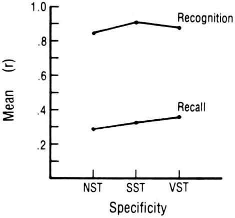


Figure 20 Test Group Values of the Mean Distance From the Origin (r).

i.e., the angle which defines one standard deviation unit to either side of the mean vector. As s approaches zero, the mean circle location approaches that of the stimulus. The mean angular deviation for all circles in each test group was also computed and plotted in Figure 21 which illustrates the consistency of each group's ability to recall or recognize the stimulus pattern.

As anticipated, the mean differences between the recall and recognition test groups for both r and s were significant at the .01 level. But no significant difference at the .05 level could be found between any of the instructional sets in either memory group for both r and s. The relationships between the mean vectors and the spread of the observations within one standard deviation unit

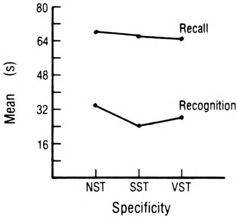


Figure 21 Test Group Values of the Mean Angular Deviation (s).

underlying those vectors are shown in Figures 22a through 22f. The mean directions of the vectors were determined from the mean x,y positions. Again, the differences between the instructional levels is not as noticeable as that between recall and recognition.

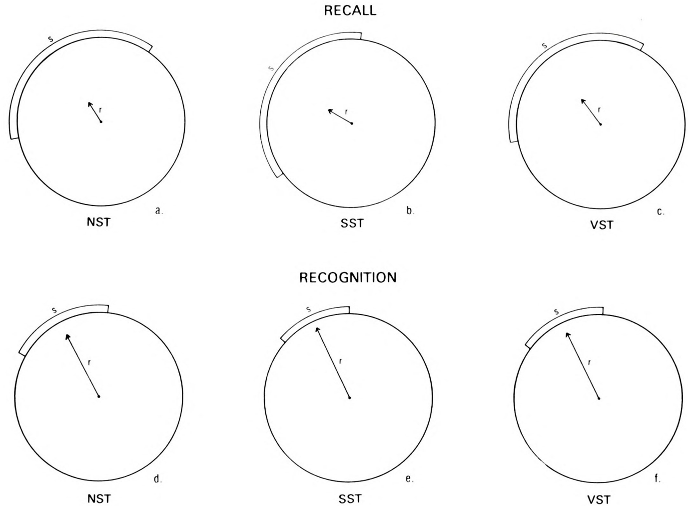


Figure 22 Graphic Representations of the Mean Distance from the Origin and the Mean Angular Deviation for all Test Groups.

A weighting factor was also applied to r and s; the formulas were the same as the above except now:

$$X = \frac{E(\cos a \cdot w)}{Ew} \quad \text{and} \quad Y = \frac{E(\sin a \cdot w)}{Ew}$$

where w equals the area of a circle. The rationale for using a weighting system of this sort has been set forth in eye movement research (Steinke, 1975; Dobson, 1977) which showed the greater visual "importance" of large black figures (circles) on a map versus smaller figures. And the subjects in the present study verbally indicated a better memory of the larger circles as opposed to the smaller ones. This concept is supported by Figures 23 and 24. These two figures graphically represent the mean r and s values, respectively, for each circle in the recall test before the weighting factor was added. Note that error increases dramatically at first, and then levels off as circle size decreases.

The use of the weighting factor produced unexpected results (Figures 25 and 26) where the weighted r and s values for the test groups, while not being mirror images of, are nonetheless the reverse of those seen in Figures 20 and 21. But the reasons for this reversal are not very apparent and do not appear to be forthcoming.

Centroids

Another measure of positional error was calculated for the response patterns in the tests. Unlike most of the

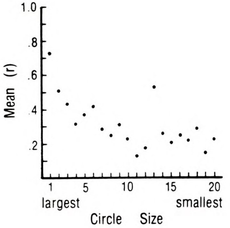


Figure 23 Mean Distance From the Origin for Each Circle in the Recall Tests.

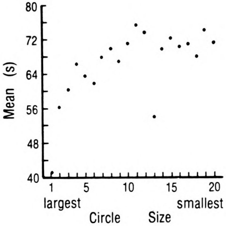


Figure 24 Mean Angular Deviation for Each Circle in the Recall Tests.

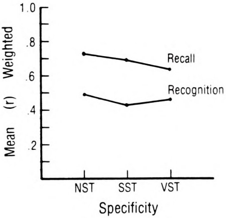


Figure 25 Weighted (r) Values for All Test Groups.

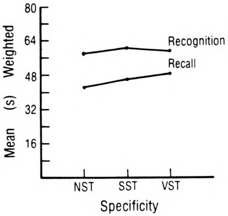


Figure 26 Weighted (s) Values for All Test Groups.

other statistical analyses used, centroids are closer to defining error for an entire pattern rather than the accumulation of error terms for individual circles. This statistic essentially defines the mean center of the response patterns in relation to that of the stimulus pattern: $\frac{Ex}{20}$, $\frac{Ey}{20}$. Like *r* and *s*, a weighting factor (area of a circle) was also computed: $\frac{Exw}{Ew}$, $\frac{Eyw}{Ew}$ where *w* equals the area of a circle. The response centroid and weighted centroid locations and their distances from the stimulus centroid are presented in figures 27a through l.

The centroids plotted for the recognition tests indicate the number of choices for each location. Because of the limited number of choices made in the recognition tests, these vectors do not appear to show as much of an orientation as do the locations of the patterns in the recall test groups. And in viewing the unweighted recall vectors, the reader may note that while the VST vectors are closer to the stimulus than either the NST or SST vectors, a more unimodal distribution of the centroids appears in the NST. When the weighting factor is added on, the orientation of the vectors shifts counterclockwise, from beneath the stimulus to the right of it. Such a shift may indicate that the larger circles, which were primarily concentrated on the left and upper sides of the stimulus map, were positioned by the subjects towards the center and the right side of their maps.

NST

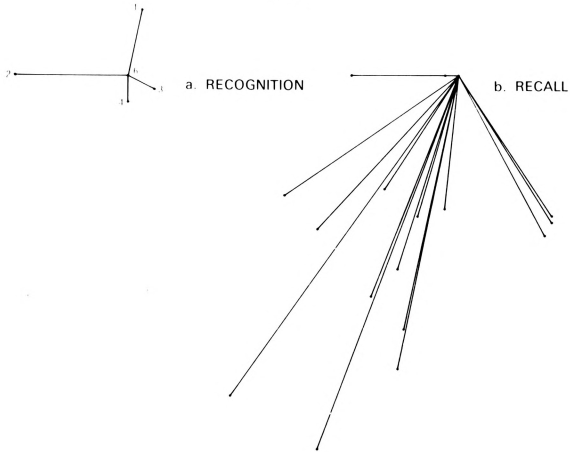
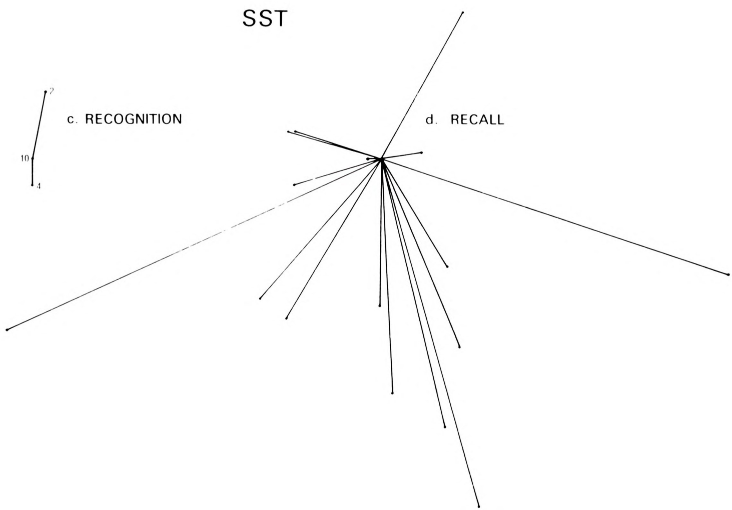
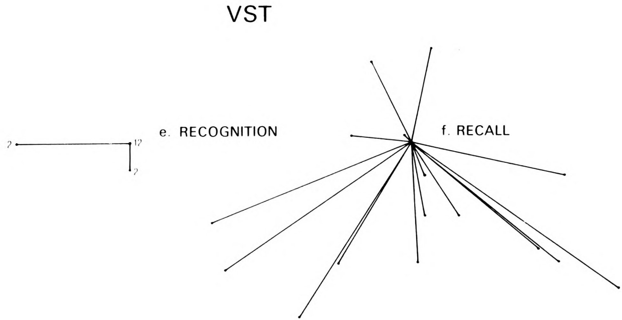
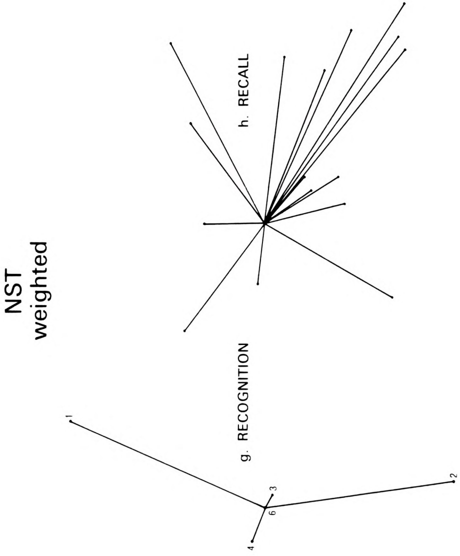
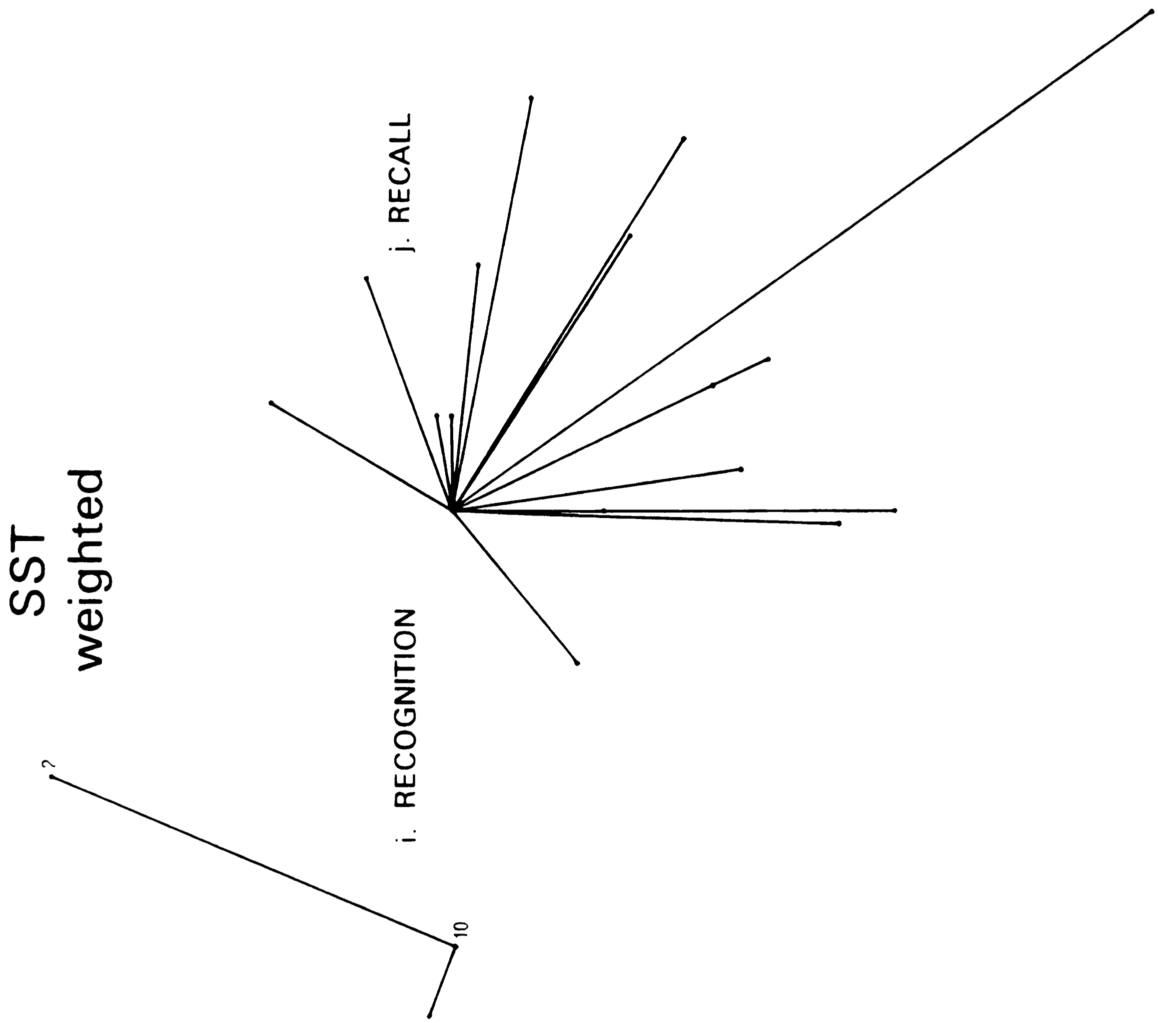


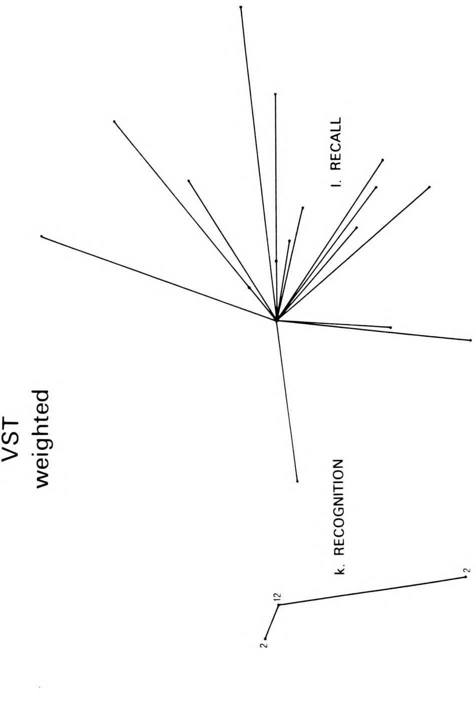
Figure 27 Response Centroid Locations in Relation to the Stimulus Centroid for all Test Groups (Recall x 7; Recognition x 11).


SST
c. RECOGNITION
d. RECALL
2
10
4


VST
e. RECOGNITION
f. RECALL


NST
weighted
h. RECALL
g. RECOGNITION
1
2
3
4
6


SST
weighted
j. RECALL
i. RECOGNITION
2
10
4


VST
weighted
l. RECALL
k. RECOGNITION
12
2
2

The values of the mean vector lengths of the unweighted and weighted test groups are plotted in Figures 28 and 29, respectively. As predicted, the difference between NST and SST is less than that between SST and VST for recall. This statistic also shows that the recall error in VST is sufficiently less than NST to be significant at the .05 level. Recognition, on the other hand, again exhibited a tendency for the error terms to decrease from NST to SST and then rise from SST to VST, although the difference at the .05 level could be detected between any of the recognition test groups while recall and recognition did differ significantly at the .01 level.

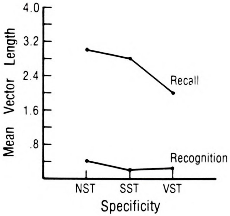


Figure 28 Mean Vector Lengths for All Test Groups.

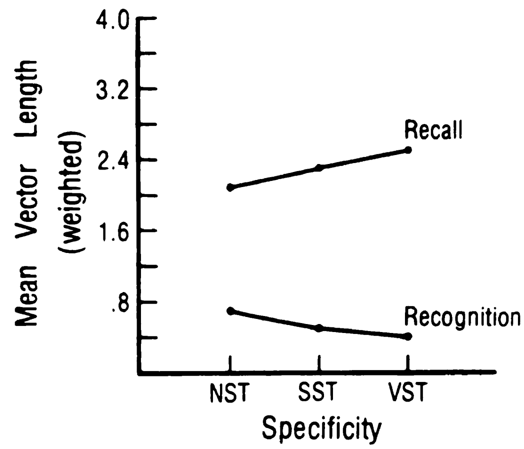


Figure 29 Weighted Mean Vector Lengths for All Test Groups.

The addition of a weighting factor affected the recall mean vector lengths in much the same way that the weighted r and s terms behaved, i.e., the progression of error terms from NST to VST are reversed from the unweighted terms. Surprisingly, the weighted error terms for recognition did not behave in the same fashion. In fact, the progression resembles that for the mean deviation of the percent black per quadrant (Figure 17). The relationship between the two slopes could mean that since the larger circles are given greater weight in the location of a pattern centroid, their influence also generates a greater percentage of black for a particular quadrant; and since, in VST, the stimulus pattern is chosen more often than in either NST or SST, the "pull" of the larger circles to the stimulus centroid is that much stronger. But no significant difference for either the

weighted recall or recognition test groups could be seen at the .05 level. As in the other analyses, the two memory tasks differ significantly at the .01 level.

Summary

Table 3 summarizes the relationship between error and consistency for the recall and recognition test groups given the above statistical analyses. All of the tests demonstrated that the error and consistency in recall and recognition are significantly different at the .01 confidence level. The difference between the instructional sets, as expected, was not as great, and in most cases was not statistically significant. Between NST and VST, however, a significant difference can be noted at the .05 level for the mean squared distance after rotation and the mean vector lengths and at the .1 level for the mean squared distance before rotation while no significant differences were found between any of the recognition tasks by any of the above statistical procedures. Due to the choices made by two of the subjects in the very specific recognition task, all of the statistical procedures, with the exceptions of the frequency of choices, the mean deviation of the percent black per quadrant and the weighted centroids, showed VST to possess a larger degree of error and a smaller degree of consistency than SST. But the three recall tasks exhibited either a nearly linear decrease in error terms (mean squared distance after rotation, mean $\underline{r}$) or an accelerated decrease in error terms from NST to SST to

VST (mean squared distance before rotation, mean deviation of the percent black per quadrant, mean vector lengths). Consistency, on the other hand, decreased linearly (mean $\underline{s}$) or at a decelerated rate (frequency of choices) from NST to SST to VST.

Table 3

Summary Table of the Mean Error Terms for all Test Groups

	Recall			Recognition		
	NST	SST	VST	NST	SST	VST
$\bar{D}\%b/q$	5.706	5.361	4.220	1.872	1.491	.944
$\bar{X}sdbr$	8973.9	8735.7	7813.1	1860.4	915.4	1401.2
$\bar{X}sdar$	33.56	31.72	39.84	7.44	3.38	5.75
$\bar{X}r$	.2865	.3249	.3604	.8420	.9275	.8797
$\bar{X}s$	68.0	65.9	65.1	34.2	24.6	28.5
$\bar{X}rw$	.7278	.6827	.6346	.4867	.4287	.4662
$\bar{X}sw$	42	46	49	58	61	59
$\bar{X}v$	3.05	2.80	1.99	.46	.23	.26
$\bar{X}vw$	2.13	2.29	2.49	.73	.53	.42

$\bar{D}\%b/q$ - Mean deviation of the percent black per quadrant.
$\bar{X}sdbr$ - Mean squared distance before rotation.
$\bar{X}sdar$ - Mean squared distance after rotation.
$\bar{X}r$ - Mean distance from the origin.
$\bar{X}s$ - Mean angular deviation.
$\bar{X}rw$ - Mean distance from the origin (weighted).
$\bar{X}sw$ - Mean angular deviation (weighted).
$\bar{X}v$ - Mean vector length.
$\bar{X}vw$ - Mean vector length (weighted).

Chapter IV

Discussion

Conclusions

Error

Several questions should be readdressed at this time: First, did the two different memory tasks of recall and recognition produce significantly different results? Regardless of the statistical measure used, there was a highly significant difference between recall and recognition tasks: recall always produced greater response error than recognition. Two basic reasons account for this consistently large difference in error: (1) In order to make the recognition test sufficiently difficult, the distractor map patterns had to be constructed which were similar to the stimulus pattern; so the error of the recognition tasks was expected to be small whereas the subjects in the recall tasks could possibly create a map having close to 100 percent error. (2) The variation in recognition error was limited to the eight maps presented as potential targets; in any recall task, on the other hand, the variance is unlimited because an infinite number of subjects conceivably could create an infinite variety of patterns, given the confines of the map.

Second, did changes of specificity in the instructions create significantly different map reader responses? While

the statistical analysis has proven that no significant difference in the mean error terms exists between NST and SST or between SST and VST, a definite progression from high to low error does exist from NST to SST to VST. One may also assert that there is a significant difference in mean error terms between NST and VST for the recall test groups. But due to the problems concerning the one map that was chosen by two of the subjects, no empirical statement, strong or weak, will be recorded about the difference of error between the recognition test groups.

Based upon the results obtained in the data analysis, hypothesis Ia, error will be less for recognition tasks than for recall tasks, is confirmed. And while the error terms for each level of instruction may not be significantly different, a trend does exist which supports hypothesis Ib: error will decrease from a non-specific to a somewhat specific to a very specific instructional set.

Consistency

Finally, did the subjects consistently remember the circle patterns to be a particular shape or in a particular position on the map? From those who responded, apparently a variety of shapes were consciously or subconsciously utilized by the subjects as an aid in remembering the pattern of circles. From the vectors plotted in the circular distributions (Figure 22) and from the vectors plotted to the centroids (Figure 29), one might note that the pattern

locations, while probably not being significantly consistent, are at least relatively consistent within a given area. Most of the circles in each pattern were congregated on the left side of the map--where the greatest amount of blackness was present on the stimulus map. Obviously, the mental coding and organization of the stimulus pattern varied from subject to subject while the regurgitation of that pattern was somewhat consistent within a given test group because the mean vectors indicate that the level of specificity does indeed partially influence the position of the pattern.

Given the frequency of responses in the recognition tests, the mean angular deviations and the centroid locations of the patterns for all test groups, one should note that a trend does exist which supports hypotheses IIa: the stimulus pattern will always be more consistently recognized than it will be reproduced; and IIb: the consistency with which the stimulus pattern is recognized or reproduced will increase as the task becomes more specific.

Recommendations

Given the variables considered in this study, one may put forth the following guidelines for cartographic psychophysical testing methodology: (1) Even though a strong case cannot be made on the differences between task orientations, a trend did exist such that one may state that the instructions in any testing procedure should be as specific as possible.

Indirectly, one might assume that clarity, conciseness and consistency should also be maintained because otherwise, when the tasks are read to all segments of the test population, the level of specificity will either differ from subject to subject or from sample to sample. Thus, inasmuch as a very specific task is required to prevent biased results, it should not be written or verbalized to the extreme that clarity is sacrificed for it.

(2) Although recall and recognition are rarely intentionally used together in any psychophysical test, one should be aware of the exact type or combination of memory response(s) being elicited from the subjects on a particular test. Even though some degree of overlap is present between recall and recognition, they are still two different memory responses and the variables that affect them react accordingly. And while the use of recognition is greater than recall in cartographic psychophysical tests, the latter may have greater applicability to everyday map usage; but recall is a much harder response to study given the difficulty of the task for a subject.

Further Research

Additional testing should be conducted on the types of memory response and instructional set now present in testing methodologies. Concerning this topic alone, the following suggestions are put forth: (1) the sample population should be large enough so that individual variations in the responses do not significantly affect the variations in the mean response; (2) the use of eye movement technology as reported by DeLucia (1974) and Steinke (1979) would present important quantitative and qualitative additions to the measures already used in this study; (3) additional variables, i.e., the complexity of the test maps, environmental perception, exposure time, and so forth, could be added into the testing procedure to see what, if any, effects their presence would have on map reader responses. In other words, while recommendations have been made based upon the research conducted within this study, the evaluation of cartographic testing methodology is not complete; and until the effects of all test situation variables are examined, cartographers should exercise caution in both their tests and in the conclusions that they reach.

BIBLIOGRAPHY

Anastasi, Anne. Psychological Testing. New York: Macmillan, 1976.

Archer. E. James. "Identification of Visual Patterns as a Function of Information Load," Journal of Experimental Psychology, Vol. 48, No. 5 (November 1954), pp. 313-17.

Attneave, Fred and Malcolm Arnoult. "The Quantitative Study of Shape and Pattern Recognition," in Uhr, Leonard, ed., Pattern Recognition. New York: John Wiley & Sons, 1966.

Attneave, Fred. "Symmetry, Information and Memory for Patterns," American Journal of Psychology, Vol. 68, No. 2 (June 1955), pp. 209-22.

Baddeley, Alan D. The Psychology of Memory. New York: Basic Books, Inc., 1976.

Blumenthal, Arthur L. The Process of Cognition. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1977.

Brown, John. "Some Tests of the Ceday Theory of Immediate Memory," Quarterly Journal of Experimental Psychology, Vol. 10, No. 1 (1958), pp. 12-21.

___________. "An Analysis of Recognition and Recall and the Problems of their Comparison," in Brown, John, ed., Recall and Recognition. New York: John Wiley & Sons, 1976.

Castner, Henry. "The Role of Pattern in the Visual Perception of Graded Dot Symbols in Cartography," unpublished Ph.D. dissertation. Madison, Wisconsin: University of Wisconsin. 1964.

Castner, Henry and Arthur Robinson. "Dot Area Symbols in Cartography: the Influence of Pattern on Their Perception," American Congress on Surveying and Mapping, Technical Monograph #CA-4, 1969.

Cooper, Aldwin and Andrew Monk. "Learning for Recall and Learning for Recognition," in Brown, John, ed., Recall and Recognition. New York: John Wiley & Sons, 1976.

DeLucia, Alan. "The Nap Interpretation Process: Its Observation and Analysis Through the Technique of Eye Movement Recording," unpublished Ph.D. dissertation, University of Washington, 1974.

Dobson, Michael W. "The Map in the Mind's Eye," Auto-Carto II, (1975), pp. 225-32.

________________. "The Adaption of an Eye Movement Recording Technique to a Cartographic Experiment Involving the Graduated Symbol," unpublished Ph.D. dissertation, University of Kansas, 1977.

Dodwell, Peter C. Visual Pattern Recognition. New York: Holt, Reinhart and Winston, Inc., 1970.

Downs, Roger M and David Stea, eds. Image and Environment: Cognitive Mapping and Spatial Behavior. Chicago: Aldine Publishing Co., 1973.

Downs, Roger and David Stea. Maps in Minds: Reflections on Cognitive Mapping. New York: Harper & Row, 1977.

Eagle, Morris and Eli Leiter. "Recall and Recognition in Incidental and Intentional Learning," Journal of Experimental Psychology, Vol. 68, No. 1 (January 1964), pp. 58-63.

Fitts, Paul, Meyer Weinstein, Maurice Rappaport, Nancy Anderson and J. Alfred Leonard. "Stimulus Correlation of Visual Pattern Recognition; A Probability Approach," Journal of Experimental Psychology, Vol. 51, No. 1 (January 1956), pp. 1-11.

Flannery, James J. "The Graduated Circle: A Description Analysis and Evaluation of a Quantitative Map Symbol," unpublished Ph.D. dissertation, University of Wisconsin, 1956.

French, Robert J. "Identification of Dot Patterns from Memory as a Function of Complexity," Journal of Experimental Psychology, Vol. 47, No. 1 (January 1954), pp. 22-26.

________________. "Pattern Recognition in the Presence of Visual Noise," Journal of Experimental Psychology, Vol. 47, No. 1 (January 1954), pp. 27-31.

Gerlach, Arch C. "Visual Impact of Thematic Maps as a Communication Medium," International Yearbook of Cartography, Vol. 11 (1971), pp. 169-76.

Groop, Richard E. and Daniel G. Cole. "Overlapping Graduated Circles: Magnitude Estimation and Method of Portrayal," Canadian Cartographer, Vol. 15, No. 2 (December 1978), pp. 114-22.

Hake, Harold W. "Form Discrimination and the Invariance of Form," in Uhr, Leonard, ed., Pattern Recognition. New York: John Wiley & Sons, 1966.

Herriot, Peter. Attributes of Memory. London: Methuen and Co., 1974.

Jenks, George F. "The Evaluation and Prediction of Visual Clustering in Maps Symbolized with Proportional Circles," in Davis, John C. and Michael J. McCullagh, eds., Display and Analysis of Spatial Data. New York: John Wiley & Sons, 1975.

Kaufman, E.L., M.W. Lord, T.W. Reese and J. Wolkmann. "The Discrimination of Visual Number," American Journal of Psychology, Vol. 62, No. 4 (October 1949), pp. 498-525.

Lloyd, Robert and Theodore R. Steinke. "Visual and Statistical Comparison of Choropleth Maps," Annals of the Association of American Geographers, Vol. 67, No. 3 (September 1977), pp. 429-36.

Leewenberg, E.L. and H.F. Buffart, eds. Formal Theories of Visual Perception. New York: John Wiley & Sons, 1978.

Loftus, Geoffrey R. and Elizabeth F. Loftus. Human Memory - The Processing of Information. New York: John Wiley & Sons, 1976.

Mandler, George. "Organization and Recognition," in Tulving, Endel and Wayne Donaldson, eds., Organization of Memory. New York: Academic Press, 1972.

Mardia, K.V. Statistics of Directional Data. New York: Academic Press, 1972.

McGeoch, John A. "Forgetting and the Law of Disuse," Psychological Review, Vol. 39, No. 4 (July 1932), pp. 352-70.

Mendenhall, William, Lyman Ott, and Richard Larson. Statistics: A Tool for the Social Sciences. Boston: Duxbury Press, 1974.

Monmonier, Mark S. "Measures of Pattern Complexity for Choropleth Maps," American Cartographer, Vol. 1, No. 2 (October 1974), pp. 159-69.

Muehrcke, Phillip. "Visual Pattern Analysis: A Look at Maps," unpublished Ph.D. dissertation, University of Michigan, 1969.

Muller, Jean-Claude. "Objective and Subjective Comparison in Choropleth Mapping," Cartographic Journal, Vol. 13, No. 2 (December 1976), pp. 156-66.

Neisser, U. Cognitive Psychology. New York: Appleton, Centry and Crofts, 1967.

Norton, David and Lawrence Stark. "Scanpaths in Saccadic Eye Movements while Viewing and Recognizing Patterns," Vision Research, Vol. 11, No. 9 (September 1971), pp. 929-42.

Olson, Judy. "The Effects of Class Interval Systems on Visual Correlation of Choropleth Maps," unpublished Ph.D. dissertation, University of Wisconsin, 1970.

__________. "Class Interval Systems on Maps of Observed Correlated Distributions," Canadian Cartographer, Vol. 9, No. 2 (December 1972), pp. 122-31.

__________. "Autocorrelation and Visual Map Complexity," Annals of the Association of American Geographers, Vol. 65, No. 2 (June 1975), pp. 189-204.

Pellegrino, James W., et al. "Short Term Retention of Pictures and Words as a Function of Type of Distraction and Length of Delay Interval," Memory and Cognition, Vol. 4, No. 1 (January 1976), pp. 11-15.

Phillips, W.A. "On the Distinction Between Sensory Storage and Short Term Visual Memory," Perception and Psycho-Physics, Vol. 16, No. 2 (October 1974), pp. 283-90.

Reyment, R.A. Introduction to Quantitative Paleoecology. New York: Elsevier Publishing Co., 1971.

Schulman, Arthur I. "Recognition Memory and the Recall of Spatial Location," Memory and Cognition, Vol. 1, No. 3 (June 1973), pp. 256-60.

Sneath, P.H. "Trend Surface Analysis of Transformation Grids," Journal of Zoology, Vol. 151 (1967), pp. 65-122.

Steinke, Theodore R. "The Optimal Thematic Map Reading Procedure: Some Clues Provided by Eye Movement Recordings," Auto-Carto II, 1975.

Taylor, Peter J. Quantitative Methods in Geography. New York: Houghton-Mifflin, 1977.

Tulving, Endel and Zena Pearlstone. "Availability vs. Accessability of Information in Memory for Words,"

Journal of Verbal Learning and Verbal Behavior, Vol. 5, No. 4 (August 1966), pp. 381-91.

Tversky, Barbara. "Encoding Processes in Recall and Recognition," Cognitive Psychology, Vol. 5, No. 3 (November 1973), pp. 275-82.

________________. "Eye Fixations in Prediction of Recognition and Recall," Memory and Cognition, Vol. 2 No. 2 (April 1974), pp. 275-78.

Warrington, Elizabeth K. and Carol Ackroyd. "The Effects of Orienting Tasks on Recognition Memory," Memory and Cognition, Vol. 3, No. s (March 1975), pp. 140-42.

Wathen-Dunn, Weiant, ed. Models for the Perception of Speech and Visual Form. Cambridge, Massachusetts: M.I.T. Press, 1967.

Whiteside, John A. "Visual Display Recognition and the Duplication of Inspection Sequences," Perception and Psychophysics, Vol. 23, No. 4 (April 1978), pp. 350-55.

Wicklegren, Wayne A. and Donald A. Norman. "Strength Models and Serials Position in Short Term Recognition Memory," Journal of Mathematical Psychology, Vol. 3, No. 2 (July 1966), pp. 316-47.