MECHANISMS OF THE MIND'S EYE

A THESIS

SUBMITTED TO THE SCHOOL OF GRADUATE STUDIES
IN PARTIAL FULFILMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
MASTER OF EDUCATION

by
PETER W. JOHNSON

UNIVERSITY OF MANITOBA
WINNIPEG, CANADA

APRIL, 1978

MECHANISMS OF THE MIND'S EYE

BY

PETER W. JOHNSON

A dissertation submitted to the Faculty of Graduate Studies of the University of Manitoba in partial fulfillment of the requirements of the degree of

MASTER OF EDUCATION

ABSTRACT

Ten subjects were asked to mentally rotate six slightly varying stimuli to fixed positions of orientation. Each of these stimuli were varied according to known features utilized in visual perception. Each of these variations (outline, solid form, angled form) were further presented in an incomplete (unclosed) format. A response time measure was recorded for subjects mentally rotating these stimuli.

A linear response time measure for these various mental rotations would replicate previous research (Shepard 1971, 1973, 1975; Cooper 1975). This research assumed that mental image encoding and rotation was "quasi-perceptual" or spatial in its construct, in that the rotation of an object in real space also reveals a linear time measure.

The actual structural format of the encoded stimuli could be inferred from markedly different response time sets for different stimulus presentations.

Previous research results were upheld only marginally. Two of six stimulus presentations revealed linear response time data. Analysis of variance computation revealed that the six presentation formats of the six stimuli were not significant in the determination of different response times for mental rotation. Hence the structural format of a mental image is still unknown.

ACKNOWLEDGMENTS

I should like to thank the members of my thesis committee, Dr. Gavin Hoare, Chairman, Faculty of Education; Dr. Murray Singer, Department of Psychology; and Dr. Peter Blakey, Fort Garry School District, for their direction, concern, and patience.

I should also like to acknowledge my indebtedness to the Program Advisor, Mark Judd, of Computer Services, and Malcolm Shooter, Programmer in Computer Science, for some much needed aid in the program description of the statistics used.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

CHAPTER I

INTRODUCTION

Background: mental imagery and learning

As children grow, they acquire knowledge. Just how children come to obtain, and organize information is a problem which has puzzled teachers and psychologists for many years. Piaget at Geneva (1950), Murray at Harvard (1955), and Jerome Bruner (1957) at Oxford, did much to relate the efficacy of teaching methodology with the theoretical issues of how children learn.

Neisser (1967),has defined the acquisition of information as, those "processes by which sensory input is transformed, reduced, elaborated, stored, recovered, and used." For Neisser, all contact with the external world is somehow represented internally. This internal representation of our physical environment involves far more complex processes than what might be called the "storehouse" conception of mental life.

The mind had been thought of in early psychology, as a kind of vessel to which such things as facts, ideas, feelings, sensations, and images were to be stored and found. Teachers were quick to accept this concept of mind, and methodology was concerned with filling this vessel with the appropriate language and numerical skills. Even today, the "clean slate" concept of a child's mind is still held by some teachers.

A Changed Paradigm

More recently, education, as well as experimental psychology, has been undergoing a paradigm shift. The mind is no longer thought of as a place, but as an organ (much like a lung) with a specific function. The function of mind is the internalization and processing of information.

Seeing the mind as an information-processing body, theoretical positions (models) have been described in order to account for the vast array of stages, strategies, activities, and transformations necessary to internalize and represent man's external environment by way of his five senses and his memory.

The mind patterns man's world, and his experience of it. Cognition is an active, not a passive process. But what exactly do we mean by a mental picture? Further, just how does the mind derive and recall these patterns such that man can re-cognize, adapt, or change his environment? Lastly, are these patterns modality specific, or are they related in say the structure and organization of his linguistic and visual systems? Nowhere are these questions more closely studied today than through the phenomenon known as mental imagery.

The phenomenon of mental imagery is persistent in all mental processes. Hills (1957), and Barakat (1951), have found that an internal spatial ability is an essential quality for aptitude in mathematics. Paivio (1969), has done voluminous research on the associative power of a mental image in understanding the meanings of words; Gibson, Osser, and Hammond (1962), have shown that reading readiness is as much a mental image (grapheme) property as it is a phonemic one. Lastly, Sternberg (1969), has revealed that mental imagery is utilized in the search for the retrieval of information from memory. Mental imagery is so pervasive and ubiqitous a phenomenon that often it is referred to as "seeing with the mind's eye."

The Metaphorical Trap

Mental images are so vivid, and so much a part of other mental processes that there has been a tendency for the layman teacher or psychologist to accept the metaphorical description to see "with the mind's eye" as if it were indeed real. The exact metaphor utilized by investigators to describe this "internal seeing" has varied throughout history. This variation has always depended upon the technology for recording this most personal yet most ubiquitous experience.

Mental images have been described by Aristotle and other early Greeks as internal wax tablets. Galton (1880), described the notion of a mental image as a kind of internal photographic plate. Kerr (1932), believed imagery to be a sort of mental motion picture. Narasimham and Reddy (1967), put forward the proposition that mental imagery might be described as a sort of internal computer graphics. Most recently, Pribam, Newer, and Baron (1974), described mental imagery as akin to an internal mental holography.

The questions still remain: what "really" are mental images? what are their patterns and properties? how are images represented internally? Perhaps most importantly, what is the relationship between seeing in the perceptual sense (with the foveal system) and "seeing" images in memory?

Perceptual learning is essentially the movement from sensory input data which is patterned into an internal representation. This "patterning" usually refers to a configuration consisting of several elements that somehow fit together. Patterns consist of elements called features, attributes, cues, dimensions, or components, and as such, these words could be used interchangeably.

Eleanor Gibson (1969), has made an analysis of the cognitive processes involved in perceptual learning. Gibson proposed that the discovery of features precedes the formation of any internal representation of one's external environment. A child learns the basic sounds of a language long before he uses words. Once that child has abstracted or concretized these features, a complete inner representation of a part of his world (a sound, an image, etc.) is possible such that he can now detect other mental patterns by comparison to those already internalized.

This feature identification theory proposed by Gibson (1969), has been applied to a child's perceptual learning of visual patterns. Caldwell and Hall (1970), found that lines of different orientations, various contours, horizontal and vertical discontinuity, were all examples of features utilized by children in discriminating standards and transformations of various letter-like visual stimuli. There has been difficulty, however, in providing an acceptable theoretical framework which would allow for the construction of new patterns from "older" internal configurations, especially when attempting to account for the multitudinous relationships inherent in the two-or three-dimensional internal representation of an object--which brings us back again to mental imagery amd memory.

The fundamental problem in dealing with the question of the nature of mental imagery lies in its definition. Imaging is the "back-end" of a perceptual process, so to speak. A child can retain the visual memory of an object long after it has disappeared from view. Imagery is like perceiving in that it utilizes similar cognitive processes involved in vision. However, it need not necessarily refer to any introspective reports of picture-like mental contents. Though mental imagery is often reported as appearing like what we see, in actual fact it may not be so vivid or graphic as reported.

Neisser (1969), has pointed out:

> a subject is imaging whenever he employs some of the same cognitive processes that he would use in perceiving, but when the stimulus input that would normally give rise to such a perception is absent (Cognition and Reality, p. 129).

Defined in this way, mental imagery refers to all those quasiperceptual experiences of which we are aware (in memory) and which exist in some form despite the absence of those stimulus conditions which produce the genuine sensory or perceptual counterpart. Mental imagery then is defined as any concrete (known) representation of sensory, perceptual, affective, and other experiential states. In this sense there can be: after-imagery (visual optical persistence of retinal activity), auditory imagery (remembered sounds), eidetic imagery (remembered "pictures"), kinesthetic imagery (remembering or photographic memory movement), tactile imagery (remembering the touch of things), and visual memory imagery.

Visual memory is the commonest and most familiar form of imagery. It is also a most uniquely private event. Visual memory imagery is often described by way of its vividness to the world we actually see, as in the Betts QMI Vividness Scale, or in its perceptual brightness, as in the Gordon Test of Visual Imagery Control. Introspective measures of visual memory imagery have assumed, it seems, the reality of the metaphorical notion of an image as an "internal picture."

Empirical Research: Mental Rotation Studies

More recently, investigators have utilized more objective techniques in order to study this uniquely subjective phenomenon. These experimental design procedures have, over the last fifteen years, ranged from eye-fixation studies (Singer (1966); Gould (1969)); to pupil size studies (Paivio and Simpson (1966)); and even correlative studies of EEG readings with subjective reportage of imaging (Oswald (1962)).

The most encouraging work, however, has been done by Shepard and his associates (Shepard (1967); Shepard and Metzler (1971, 1973); and Shepard and Feng (1972)), utilizing reaction-time procedures in the measurement of subject's mental imaginal rotation of given shapes. Shepard's work has lent support to earlier introspectionist data which showed that mental images may be like "pictures in the mind." Shepard sees mental imagery as a process analogous to those utilized by the human foveal system in that the same spatial representational mechanisms operating in visual perception also operate in visual memory.

Shepard and his associates (Shepard, 1963; Shepard and Cooper, 1975; Shepard & Metzler (1971)), utilized a reaction-time experimental design. Subjects were presented (in a tachistoscope) with various two- or three-dimensional shapes in various fixed positions about their axis. Upon the disappearance of that shape, subjects were asked to imagine in their "mind's eye" that same shape rotating about its given axis to a given position, in a given direction. After a predetermined delay, the subjects were presented with a test shape to which they made a "same" or "different" determination of orientation to their internally oriented figure. This mental judgment, Shepard concluded, was made by a subject carrying out some sort of internal analog of an actually perceived external rotation. Support for this position was found in the results which revealed that reaction time for mental rotation increased linearly with angular differences in mental orientation. In other words, the subject was utilizing some sort of internal spatial representational system such that the figure was "seen" to rotate through a kind of internal trajectory.

In essence, Shepard's work endorsed a new theoretical model of the internal representation of visual memory imagery. The linearity of these reaction-time studies made him discount a discrete (propositional) representational encoding system. Such a system, Shepard argued, could not be the basis of the linear relationship between angular displacement and response time.

More than that, Shepard's work casts some doubt on the accepted models of visual pattern recognition. One theoretical model poses a "template-matching" view of visual perception. This model assumes that each new visual input (or visual memory of an input) is compared to a standard representation gained initially through direct perception. The other position, a "feature analysis" model, assumes that only particular or distinctive properties are utilized in recognition. Shepard's work favours a feature-analysis view of pattern recognition with the added notion of a perceptual anticipatory mechanism.

For Shepard, mental rotation of an image requires a schema that accepts information about the speed and direction of any internally oriented stimulus. The "image" consists of a readiness to pick up certain parts of information from a given part of a moving stimulus. Simply, subjects pick up information most quickly at the orientation they already have in mind.

Cooper (1975), proposed the anticipatory mechanism this way:

> . . . during a mental rotation [task] the internal process passes through a series of states at each one of which the subject is especially prepared for the presentation of a particular external object in a particular orientation (Cognition and Reality, p. 149).

Nowhere do these investigators however attempt to specify specifically the actual structural detail of the represented feature undergoing mental rotation.

Statement of the Problem

This study addresses itself to the question suggested by Shepard's (1971) study of how the specific spatial encoding mechanism of mental imagery works. In order to satisfy the "template vs. feature" encoding issue raised by Shepard and his collaborators, a mental rotation experiment was chosen. This study is different from previous ones in that in order to test the validity of the "features" model, the mental rotation tasks were delineated by stimulus conditions which permitted the analysis of the effects of a "features" model.

For Shepard, the "image" part of mental imagery was defined as a perceptual readiness mechanism in which certain parts of its spatial scheme were noted in different orientations. The conditions imposed upon the stimulus in the rotation task performed in this study were the isolation and highlighting of three well-known pattern-producing properties of visual perception.

A standard random shape was structured such that one of three perceptual properties of line, solidity, or angulation was prominently displayed. Each of these visually prominent displays was further reduced to an "unclosed" variation. If a particular stimulus presentation markedly reduced the reaction time for mental rotation, then it could be inferred that this particular aspect of the stimulus condition represented or contained the feature undergoing internal representation. The stimulus conditions presented to subjects for mental rotation appear in Figure 1 below.

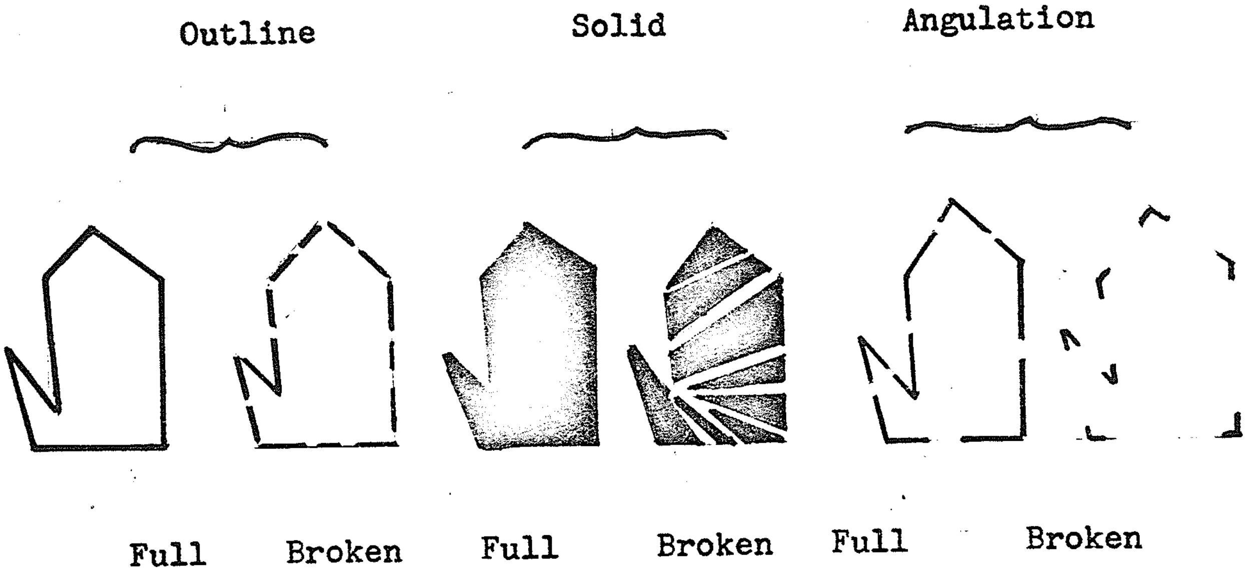


Fig. 1. The stimulus format: outline, solid, angulation, in two conditions: fullness and broken-ness.

Outline, solidity, and corners (angulation) are considered common perceptual discriminants of a given visual display, as explained in Chapter II. In order to test if the encoded features were perceptually holistic or fragmentary in nature, each of the three common conditions was further reduced to a corresponding unclosed (broken) format. A further reduced reaction time for a particular unclosed (broken) stimulus condition could be interpreted as evidence that only a small portion of a perceptual mechanism is utilized in image encoding.

Importance of the Study

The mental rotation experiment investigated in the study was undertaken to try to answer two basic questions:

1. Do mental image rotation tasks utilize spatial encoding mechanisms, such that mental imagery processes are analog in nature to the visual perception of objects rotated externally?

2. Can the analog structure of visual memory rotation tasks be specified as a feature-analytic process whose features can be known?

If any of three perceptual conditions in either whole or fragmentary portions are found to exert major effects in a mental encoding and rotation task, then the following more general questions become readily apparent for further study.

1. Is a short-term visual memory rotation task typical of other visual image processes?

2. Are the elements of image encoding universally evident and distinctive as, say, the specification of phonemic units in man's linguistic ability?

3. Are the encoding mechanisms involved in mental imagery developmental in character, as other cognitive tasks? And if so, what are the markers of such image ability and development?

Clearly, the answers to these and other questions will reveal that man possesses an internal spatial ability which is vastly different from other known propositional and linguistic mechanisms of mind. The acceptance of mental imagery as a distinct and rich quality of mind can only alert those interested in the diversity and uniqueness of all mental processes utilized in human problem-solving. This enriched understanding of a widely felt but little understood mental phenomenon can only make itself felt in improved educational methodology and curriculum design.

Definition of Terms

1. cognition. Cognition refers to all the processes by which the sensory input is transformed, reduced, elaborated, stored, recovered, and used. Cognition is concerned with these processes operating even in the absense of relevant stimulation.

2. concrete image. Reed's (1973) understanding of an internal representation of a pattern that is perceived rather than created. Once a complete concrete image is formed, subjects should be able to discriminate that pattern from other patterns.

3. feature extraction. A model of pattern recognition in which various parts of a pattern are identified in order to take place. For example, when a visual stimulus is exposed, feature extraction theory assumes that information about lines, angles, orientations, velocities, colour, and retinal disparity is obtained such that a recognition pattern is formed.

4. Gestalt. A school of psychology which is based upon the assumption that what we see is determined by whole

figures (on overall structure) rather than its determination by individual parts.

5. _information_. As defined by Shannon (1948), as essentially choice. In processing terms, Shannon defines information as the narrowing of alternatives, such that meaning (pattern) is discerned.

6. _memory trace_. An internal representation of information that is capable of being held in storage for determined periods of time.

7. _parallel processing_. An alternative model for information processing in which response to a particular pattern formation is begun before the search for templates or features is finished. In essence, parallel processing examines the visual input simultaneously for recognition detectors.

8. _pattern description_. The mental process by which "features" and relations are combined. A pattern description corresponds to a concrete image if it completely describes the pattern.

9. _perceptual learning_. The ability to learn to identify specific objects and classes of objects, such that subjects are able to recognize new members of the class.

10. _schema theory_. An important part of perceptual learning based upon the subject's ability to _generalize_. It is formulated on the idea that cognitive structures organize information into systems.

11. serial processing. A process by which a subject makes internal comparisons with external patterns in a sequential, one after the other, manner.

12. size invariance. The ability of a subject to discriminate between a pair of shapes even if the size of the shapes is altered. For example, a postage-sized letter "A" seen at six inches and a typeset letter "A" in a page of type such as this are both determined as "A's."

13. stimulus equivalence. The process of pattern recognition in which a new input is compared with a standard internalized form. This comparison may be in the form of "template-matching" or through the detection of specific features.

14. stimulus set. In the context of a stimulus, a stimulus set is additional information about a stimulus, such as the specification of a limited number of alternative responses.

15. template-matching. A model of pattern recognition in which a new input is identified by noting its complete coincidence, or congruence, with a basic model. Template-matching stands in contrast to the "feature detection" theory of pattern recognition.

Delimitations of the Study

The following delimitations are placed on this study:

1. The study is confined to the investigation of the structural nature of one particular mental image ability, namely a mental rotation task. Given the amount of unknowns in this area, generalizability from this study to other qualities of mental image manipulation should be done with extreme caution.

2. The programmed stimulus conditions and mechanized delivery system (tachistoscope) used in this study contrast widely to the reported "free-ranging" scope of natural image formation ability. The benefits of such closed image-evoking conditions are readily apparent when one is investigating this phenomenon empirically; yet it is only an assumption that image encoding mechanisms utilize any perceptual mechanisms used in vision at all. Hence, this presentation format may have little or nothing to do with "real" mental image generation or transformation tasks.

CHAPTER II

REVIEW OF RELEVANT LITERATURE

Historical and theoretical paradoxes in pattern recognition

In chapter one, we saw that visual imagery was defined as a mental process in which something is seen (as in the foveal sense) without the excitation of the retinal sensory detectors. Further, mental imagery was seen to be a quasi-perceptual process in that images made use of the perceptual apparatus. The beginnings then to any survey of literature regarding visual memory mechanisms must begin with questions that were posed about visual perceptual mechanisms.

The questions are: How does one recognize the difference between, say, a square and a circle? and, How is this pattern recognition process coded in memory? This chapter considers the solutions to these questions from the time of the Gestalt psychologists to the present. This necessitates briefly reviewing the major theoretical models of visual pattern recognition as well as those models of visual memory as described by Shepard (1969), and Cooper (1975).

The two major theoretical positions regarding visual pattern recognition are "template matching" and "feature analysis." Template matching holds that a new visual input is compared to a standard. Feature analysis holds that only particular parts or properties of a visual display are utilized in recognition tasks.

From 1860 to about 1912, Ewald Hering was the tour de force in visual perception theory. He argued a nativistic theory of visual space perception. For Hering, each unit of a receptor surface such as a retina, carried with it a specific tag, quality, or sign, which enabled an organism to abstract a particular attribute of, say, shape. This native endowment theory stood in direct opposition to the empiricists such as Helmhotz (1868), and Titchner (1919), who believed space perception was a learned phenomenon. As much as the great polemics against Hering's "local sign" doctrine began with Kohler (1929), and the Gestalt psychologists, it is interesting to see the theoretical similarity between the two. For the Gestaltists, perception was still dependent more upon the physical properties of the nervous system than upon the acquired properties that come about from direct experience.

"Template" Models of Visual Representation

The theory of isomorphism stated that the pattern of retinal excitation is in topological correspondence with the pattern of cortical activity which it induces; the correspondence is not quite topographical, since the cortical pattern is distorted by certain innate forces of constraint which act directly on it (supposed by Kohler to be electrical field forces). The resulting pattern was held to be isomorphic with what is perceived; indeed the cortical pattern was generally thought to be the sole determinant of perception. Thus the retina mirrors the external physical pattern of stimulation, and the brain mirrors, albeit in a distorted fashion, the events of the retina.

However, Kohler said nothing about how these patterns were <u>recognized</u>. More, he said nothing about how two patterns could be judged similar when they occurred in different places or at different times in the visual field.

Essentially, Kohler was arguing that an attribute, say, shape, was perceived whenever the pattern of the particular shape occurred in some area of the brain. What Kohler was arguing for in essence was some sort of psychophysiological parallelism. However, his theory of isomorphism as a tenable model for visual pattern recognition does not work because of faulty logic.

The recognition of spatial brain states as being equivalent to the external object being perceived implies an internal "observer" (homunculus) who recognizes the patterns of equivalences to external objects. This second order recognition (perception) needs to be again perceived by a second order "observer," and so on in an infinite regress. Simply, Kohler's theory of isomorphism pushes the problem of <u>recognition</u> right out of the brain.

From a physiological viewpoint, Kohler's theory of pattern recognition based upon the excitation of specific macroscopic patterns of electrical activity at the visual cortex was proven to be wrong by Lashley, Chow, and Semiones in 1951. However, Kohler's theory was a brilliant initial attempt to answer the question of how figures with similar physical or geometrical properties could be represented internally. It could have been also entirely possible that the internal representation of geometric patterns was based on the actual properties of their geometry. What Kohler failed to see was that the internalizing of a visual pattern and the form of that internal representation need not necessarily share anything in common with the perceptual characteristics of that pattern. Kohler's theory assumed that the coded representation of "squareness" was itself topographically square in the brain. This of course need not be so.

Kohler and the early Gestaltists did however, push the question further such that later experimenters and theorists could ask the question--what would be a satisfactory model for a system which recognized and internally represented stimulus equivalence?

The major drawback of the template notion of the Gestaltists was the various visual combinations and configurations of any given shape, say the letter "A"--which despite position in the retinal field, size, stylistic variety (typographic of humor), and rotation,--could always be recognized. Kohler had proposed a possible encoding of a "prototype" shape; he did not describe how the multitudinous varieties of this shape were matched against this prototype.

Lashley, as ardent an opponent of local sign theory as Kohler, came somewhat closer to a tenable model for pattern recognition with his theory of reduplicated interference patterns (Lashley, (1942)). The basic idea involved the notion that a retinal pattern of stimulation would generate in the brain a series of "interference patterns," which would be propagated over a large part of the visual cortex. One could think of these in terms of moving patterns of DC potentials, which might then interact much as do wave patterns generated on a smooth water surface into which objects are dropped; Lashley conceived of them in terms of the sympathetic activation of multiple series of timed resonating circuits.

The point of Lashley's model, of course, was to try to explain how one and the same brain state could be generated by a specific pattern of retinal stimulation independently of the particular retinal units excited. The idea that a stimulus pattern generated an interference pattern is the postulation of a coding process in which geometrical properties of a pattern, presumably such as the relative positions of various contours, determine what interference pattern is generated. Propagation of that process through the cortex is supposed then to ensure that the recognition of the original stimulus pattern was not position, or size-bound.

Lashley's (1942) model was a clear advance of the Gestaltist notion of isomorphism. His notion of reduplicated interference patterns set to capture certain stimulus equivalences was further a compromise between a passive and dynamic model of brain function. Lashley's model still does not answer the question as to how the various interference patterns were recognized--and so second and third order pattern-perceiving mechanisms lead his model out of the brain.

Pitts and McCulloch (1943) posed the question of visual representation by utilizing, not the language of stimulus equivalence, but the language of the computer. For Pitts and McCulloch, the question was, how can a system with a high degree of variable input, compute invariants such that, say, a square might be recognized despite size, orientation, and graphic peculiarity?

Pitts and McCulloch postulated a modular net theory. In essence, any input firing a neuron within a particular neuron net is transformed by a fixed set of variables which rotate, translate, or magnify the initial input. Every possible transformation is applied to each input, and the resulting changes are relayed throughout the parallel "sheets" of neuron nets. Thus the initial state of stimulation had no privileged status as did the isomorphic "prototype" of Kohler's model.

Without going into great detail, the class to which a pattern belongs is defined by enumerating all the instances that can be obtained by a set of linear transformations of the original. The question of recognition no longer poses a problem if each class of input patterns yields unique values for the variables which can be thought of as states of particular modules within the net. In this sense, such networks really do compute invariants of their variable inputs, and thus solve the general problem of stimulus equivalence.

What Pitts and McCulloch did not do, despite the logical functionalism of their model, was to specify just how the transformations themselves were encoded. In 1951, Lashley pointed out that histological evidence of the visual cortex did not seem to bear out a neuron-net arrangement so necessary for Pitts and McCulloch's model. Despite this physiological implausibility, this early theory--intriguing as it is--is not testable empirically; nor does it account for linear reaction-time studies (a non-discrete model) of Shepard, Metzler, and Cooper and the late sixties.

The only elaboration left for these early template models of visual representation is the model called "template-matching" as adhered to by Selfridge and Neisser (1960); Uhr (1963); and Gibson (1963). In this model, the new input was identified by noting its congruence with the internalized "prototype." However, congruence was a function of size, position, and location. This model then, called for a normalization process in which regardless of where a pattern appeared, it could be "moved," "reduced," or "expanded" systematically utilizing a mechanism first postulated by Pitts and McCulloch (1943).

This normalization process called for a further matching process even after a reduction, expansion, or rotation, etc., had taken place. This matching process determined which template or prototype best filled or overlapped with the normalized pattern most strongly. What other researchers had to do next was to test this new theoretical elaboration with experimental findings which would support or reject this new position. The question asked was, did the time required to perform a specific operation (such as a simulation "match" problem) support the template model of visual recognition processes?

When two multiattribute patterns are compared to determine whether or not they are identical, the comparisons of the patterns may be done simultaneously (in parallel) or one at a time (serially). Sternberg (1967) found that the time taken to reach a decision of "sameness" or "difference" between two patterns was greatly influenced by the number of differing attributes. Sternberg's findings are puzzling. If indeed the number of attributes influenced the decision time required to make a "match" or "mismatch" of two patterns, "match" responses should take longer than "mismatch" responses--if indeed the subject were utilizing a one-at-a-time serial scanning machine.

Sternberg's data indicated just the opposite; it seemed that subjects used an all-at-once (parallel) comparison procedure to report "sameness" of two patterns, and a serial scan to report "difference." Clearly, template matching necessitated a parallel (holistic) comparison procedure, yet Sternberg reported that only for matches of "sameness" was scanning time short enough to imply the use of a template (whole pattern) representational mechanism. Clearly, even Neisser's normalization template elaboration did not match what was actually happening in the brain.

Often in a pattern, a slight change in one small detail could significantly alter the meaning; yet in a template matching model, such an insignificant detail was likely to be lost in the overall comparison of new input and template.

Neisser and Weise (1960) have shown that small children could accurately detect differences through small, yet critical ill-defined features. Such discernibility implied even further that figures are not always recognized on the basis of their (overall) template qualities.

Feature-Analytic Models of Visual Representation

In contrast to "template" theories of visual pattern recognition, there are those models which suggest that <u>parts</u> and not wholes are the fundamental encoding mechanism of visual recognition. The "feature analysis" model suggests that specific parts of an input are tested for specific feature properties; or else a feature is detected in response to a particular weighted probablistic combination of tests at a very early input level. Further, the feature analyzers, or those triggered by earlier combinations of features were seen to work independently of each other.

The features to be analyzed may be of any desired type. In one feature-analytic model of pattern recognition such as Selfridge's computer model "Pandemonium" (1959), such features as horizontality, closed perimeters, concaveness, could be readily programmed in. Other feature models have detected roundness, or texture attributes of a stimulus presented before it.

It is unlikely that the human organism could start out with such clearly highly differentiated feature detection structures as those posed by computer models of pattern recognition. Yet it does not take long for a child to discern his environment visually.

It is not known how the features of, say, visual discrimination are programmed into a child's visual pattern-making apparatus.

The feature involved in visual memory pattern-recognition processes are not known. This study attempts to discover if those features of visual memory acuity are related to those spatial discriminants used in vision.

"Pandemonium," the model referred to above, was the first theory to systematically present a feature-analytic model for pattern recognition. In this model, weighted combinations of features were detected through their presence calling up a decision mechanism which is in response to the number or the regularity of the feature being detected. In essence, this decision mechanism in the "pandemonium" in input data identified the stimulus as that of a certain type. Sutherland (1957) also did research which favoured the feature analysis model of pattern recognition. Sutherland argued that if an animal could discriminate between two stimuli, it must therefore possess some mechanism which reacted differentially between the two.

Sutherland's original work was done with octopuses. These animals easily discriminate between vertical strokes and horizontal ones, but apparently cannot distinguish a line sloping 45° to the right from one which slopes 45° to the left. This led Sutherland to assume that they possessed analyzers for verticality (specifically, for the ratio of maximum vertical extent to square root of area) and horizontality, but not for other inclinations. The theory was subsequently elaborated to deal with differences in discriminative capacity between octopuses and rats, and to include other hypothetical analyzers as well.

The major drawback of these feature analytic theories was the inability of these models to develop features not initially "programmed in." Uhr (1963) believed it unlikely that an organism would start out with such a set of highly differentiated and well-adapted feature mechanisms. He developed a computer model for pattern recognition which changed a set of randomly chosen feature detectors if experience did not utilize them.

These early theories then were essentially computer models of pattern recognition. They did not account for the recognition ability or structure of visual memory. Hebb (1949) developed a feature-analysis model of visual pattern recognition which made use of features utilized in visual perception.

Hebb's account of visual pattern recognition in the mature individual resembles the other feature-oriented theories in many respects. The first level of processing was assumed to consist of "cell-assemblies" which act much like feature-analyzers. However, the only features extracted at this level were lines, angles, and contours. In effect, this model was a cross between a feature and a template theory: the "features" were really simple templates for parts. To solve Hoffding's problem--that response does not seem to depend on retinal locus--Hebb used spatially parallel processing. The cell-assemblies, or part-templates, were reduplicated all over the input region and corresponding ones were connected together. In this way, a square of a particular orientation excited what is effectively the same assembly wherever it happened to appear. The cell-assemblies themselves were supposedly combined by selective experience into what Hebb calls "phase sequences." These phase sequences were the beginnings of visual recognition.

Hebb's feature-analytic theory suffered from much the same problems of earlier feature detector theories. Simply, it did not account for things we actually see. Pattern recognition theory had progressed, yet it failed to explain the pattern processes operating in visual perception. For example, when we are temporarily blinded (as in an accident or an eye operation) we see objects for the first time (after the restoration of sight) with a "whole" separatedness and "whole" distinctiveness. A purely feature-analytic theory in which a parallel processing mechanism isolates features one at a time did not explain the basic "figure-ground" separation of whole figures which are noticed all too readily in visual perception.

Clearly, a feature analytic theory of visual pattern recognition must take into account the visual recognition of perceptual phenomenon and hence incorporate a perceptually based model of visual memory. The most recent theorists have tried to incorporate perceptual phenomenon into a feature theory of pattern recognition.

Perception Processes in Pattern Recognition

McKinney (1963, 1966); and Hebb *et al.* (1969), have conducted important research which has supported this feature theory from a purely perceptual position. In studying subjects under the "stopped" image technique, certain perceived figures broke into segments which disappeared in a seemingly ordered fashion. In this procedure, eye movements were compensated for and could not produce any shift of the optical image on the retina; that is, they did not change the proximal stimulus. Perceived figures soon disappeared in "wholes" or in "parts" when this was done, presumably because of "fatigue" at the retina or elsewhere in the visual system. Similar effects occurred even with ordinary ocular fixation on figures which are faint or defocused (McKinney, 1963, 1966).

The disappearance of parts in these experiments was not haphazard. Lines came and went as wholes, for example, so that triangles generally lost one side at a time, while the letter "R" lost its entire crosspiece. Parallel lines tended to appear and disappear together, even at considerable separations. Curvilinear figures often underwent simplification and gap-completion. Whenever possible, the fragmentation tended to produce meaningful patterns rather than nonsensical ones. A monogram broke into recognizable letters more often than into unnameable fragments; a word characteristically lost exactly those letters which would leave another definable word behind; simply, McKinney found that one lost visual recognition capability as one perhaps constructed it--in identifiable spatial features.

The occurrence of such fragmentation supports the theory that there are functional subsystems operating in short term memory as well as in perception.

Gibson (1963), has carried out studies of pattern recognition processes in young children, utilizing a rotation and transformation task. Young children were able to discern standard shapes from confused shapes in various orientations by noting standard shape features perceived in previous tests.

Further, Gibson, strongly influenced by the Jakobson-Halle notion of "distinctive features" in spoken language, attempted to discover the critical features by which letters were identified. By isolating various perceptual segments of a set of letters which were presented to four year old subjects, Gibson found that children readily confused letters like "B" and "E" which differed in only a few features, rather than letters like "B" and "C" which differed in many.

Table 1

Gibson's Table of Distinctive Features for Letter Recognition*

Features	A	B	C	E	K	L	N	U	X	Z
Straight segment										
Horizontal	+			+		+				+
Vertical		+		+	+	+	+			
Oblique/	+				+				+	+
Oblique	+				+		+		+	
Curve										
Closed		+								
Open vertically								+		
Open horizontally			+							
Intersection	+	+		+	+				+	
Redundancy										
Cyclic change		+		+						
Symmetry	+	+	+	+	+			+	+	
Discontinuity										
Vertical	+				+				+	
Horizontal				+		+	+			+

*One possible set of distinctive features for letters (from Gibson, 1965). Each letter is characterized by those features marked "+" in its column.

The significant fact of Gibson's work is that those "features" he isolated in his letter transformation and recognition task are normally considered spatial features utilized in perceptual (foveal) processes.

In 1973, Leon Harmon began a series of experiments which investigated experimentally how much visual information was required for recognition. By assigning a brightness value to very small areas of given portrait photographs, Harmon found that such a photograph could easily be divided into a grid of small squares of measurable color or tonality. By then presenting subjects with these "block" portraits which had also been measurably blurred, Harmon explored the "threshold" of visual recognition. Surprisingly, Harmon found that the recognition of the portraits was increased in those images with a high level of blurring.

Harmon explained this phenomenon through the human ability to detect and describe conspicuous features. Like phoneme recognition which is basic to an understanding of any language (Jakobson, 1949), Harmon, like Gibson (1965), saw the perceptual mechanism operating here strongly akin to the concept of "noise" in acoustics.

A picture, like a sound, could, Harmon found, be described as the sum of simple component frequencies. In acoustical signals, pressure varies with time; in the optical signals Harmon used, the frequencies were spatial and consisted of variations of "density" (or darkness) with observed distance. Just as a musical note consists of a fundamental frequency and its harmonics, so too an optical image consists of combinations of single frequencies which make up its spatial spectrum. The significance of

Harmon's (1973) work, like that of Gibson (1965), was that he gave support to a feature-detector theory of visual pattern recognition by showing that those features may well be the same spatial features utilized in perception. What Harmon was suggesting was that visual memory utilizes spatial-perceptual processes in a feature-analytic way. This feature-analytic spatial model of visual memory has been suggested by other researchers and suggests an inner spatial process to be at the heart of visual memory encoding.

In trying to describe the process of pattern recognition in general within an organism, then, researchers were forced to examine a mechanism rich in patterns--that of vision. Accounting for mechanisms which produced patterns in visual perception via a model led experimenters inevitably to a search for those mechanisms which might be combined to form patterns in visual memory. Recent research, as shown, has suggested that visual memory utilizes the same spatial mechanisms operating when we actually perceive an object or event externally. Having then provided a theoretical basis for visual pattern recognition we must analyze further the spatial or iconic nature of visual memory.

The Structural Nature of Visual Memory

If spatial perceptual mechanisms are used as research suggests in visual memory, the questions arise, to what extent is "seeing" with the mind's eye like seeing when we use the normal mechanisms of sight? Is seeing in memory actually structured like an inner photograph? The presence of such spatial/perceptual mechanisms as contrast, verticality, and symmetry point to a form of mental pattern recognition that is iconic. The question is, what is the nature of that iconic storage mechanism? Further, if the internal representation is not exactly photographic, just what is the structural role and format of these perceptual features in visual memory?

The first task is to prove that an iconic storage mechanism exists. The second is to cite just how photographic this internal storage system is. The starting point for such a test of modality is the phenomenon of how we see and remember real objects when they are overlapped by others.

There are two theoretical viewpoints regarding how one might represent images with a distinct spatial layout such as overlap. The Helmholtz position, endorsed by most perceptual psychologists since the 19th century, is that the brain receives from the retina a two-dimensional mosaic and that inferences are made about spatial relations on the basis of the same cues one uses in perceiving relative location in a photograph, a process called unconscious inference. The contrasting view, advanced by J. J. Gibson (1955), was that one perceives the layout of objects in the world somehow directly. According to this view there is information reaching the visual system that conveys the relative locations of objects without an additional step of inferring how these objects must be arranged for the current retinal mosaic to be obtained. By the Gibson position, one can perceive that one object was hidden by another, whereas by the Helmholtz position one could not. Gibson's view was that as objects disappeared from view and emerged from behind one another, one would continue to perceive their layout in space directly; Helmholtz would maintain that a hidden object is known only through memory.

Neisser and Kerr (1973), wished to determine if mental imagery was represented somehow like a photograph, in which case overlapping objects would be invisible, or whether the visual encoding system would allow one to see "through" overlaps. One way to assess this difference would be, Neisser and Kerr theorized, to see if memory effectiveness was increased through subjects seeing images with direct versus overlapping spatial layouts.

Proof of Iconic Storage

Again, Neisser and Kerr's (1973) insight was to test the proposition by comparing the mnemonic effectiveness of images based on hidden or overlapped objects versus images based on objects not hidden. For example, it is known that when subjects are asked to learn the pair PIANO-CIGAR, performance is likely to be improved if they think of, or imagine a piano with a cigar sitting at the edge of the music rack. The question is whether the same benefit to memory occurs when subjects think of a piano with a cigar hidden from view down below the strings. If mental imagery is a mental snapshot, then there should be facilitation only when both objects are "mentally visible" and not when one of them is concealed. If imagery corresponds to a Gibsonian percept, a knowledge of spatial layout, however, then the concealed element should be an effective mnemonic element.

The task used by Neisser and Kerr consisted in reading sentences with two major concrete objects represented; one could think of these objects as the two terms of a paired-associate item. Subjects were not told they would later have to remember the pairs but were instructed only to rate the vividness of each sentence at the time of its presentation. The three conditions differed in the type of images implied by the sentences. In the pictorial condition, the two objects were portrayed in an interacting scene; in the separate condition, one of the objects was somehow hidden or overlapped by the other.

After reading and rating these materials first for the vividness of the imagery they evoked, subjects were given a surprise memory test in which they were required to respond, in terms of one of Neisser's examples, with the response STATUE OF LIBERTY, given the stimulus HARP. In other words, as long as the image suggested to the subject involved the objects to be associated in close proximity, it made no difference whether the imaginal scene would, in the real visual world, have allowed both items to be seen or whether one of the two would have been hidden. Imagining a harp inside the statue's torch was just as good as imagining it balanced on top of the torch, insofar as setting up an enduring statue-torch association. However, in the separate condition, with images in which the objects to be associated were deliberately placed in remote spatial positions, performance was impaired. Wollen et al. (1972) also found that the objects pictured in a visual image must be interacting somehow for a mnemonic advantage to occur.

It is significant that the pictorial condition led to the awareness of more vivid imagers than either of the other two conditions. Further, Neisser and Kerr found that there was a lack of correlation between vividness of imagers and their helpfulness in facilitating memory. This independence of vividness and mnemonic effectiveness is a direct violation of the snapshot metaphor: the pictorial and separate conditions were quite an adequate manipulation for producing mental pictures of different vividness, yet they do not produce different amounts of memory facilitation; therefore, the memory facilitation must not be a product of the photographic clarity of images.

Neisser and Kerr showed, in further support of this conclusion, that for any given subject, considered separately, his more vividly rated sentences were not the ones he recalled best. Across subjects, moreover, those who generally reported getting vivid images did not perform better in memory than those subjects who reported less vivid images.

In summary, it seems from these data that images elicited by verbal materials were not like snapshots. Objects that would be invisible on a snapshot nonetheless, received the full memorial advantage of imagery instructions.

The Neisser-Kerr experiment suggested a somewhat less peripheral and more central concept of visual imagery in memory than some of the earlier studies reviewed. If Gibson is correct about perception in general, however, then these data support the exact parallel between concurrent visual experience and visual experience in memory.

The Neisser and Kerr findings place objections to the concept of imagery on the part of artificial-intelligence workers (Anderson and Bower, 1973; Pylyshyn, 1973) who claimed that imagery was not an isomorphic analog process, but a digital process, much like that of computer mechanism for pattern recognition. These critics, especially Pylyshyn, object that the photographic metaphor is a fatal weakness of the concept of imagery. From an isomorphic position, Neisser and Kerr (1973) have themselves shown that visual memory is <u>not</u> photographic in nature--yet spatial in design.

The real issue of contention, once one has dismissed the snapshot metaphor along with the wax-tablet metaphor, is whether or not information maps onto two fundamentally different representation systems, one propositional-verbal and one somehow imaginal. Neisser and Kerr's findings gave not only strong support for a separate modality process operating in memory, they also showed that although visual memory is not exactly photographic in nature, it is spatial in its construct.

Finally, a startling piece of evidence against the mental snapshot metaphor comes from a study by Jonides, Kahn, and Rozin (1975). These authors compared performance of normal college students in a design involving manipulation of word concreteness and of imagery building instructions (see also Paivio, (1971a), pp. 518-520). The results showed no effect of blindness. The facilitation from concrete words, and the facilitation from instructions to form mental images were just as great for blind subjects as for normals. Because the blind subjects tested in this study had never had any visual experience, whereas they showed essentially normal patterns of facilitation in memory, the lesson of Neisser and Kerr is confirmed and extended by the Jonides et al. (1975) results. These authors had no special insight as to what the coding processes involved in the memory facilitation were for their blind subjects, but, as they mentioned, the possibility would be strengthened that something similar to an abstract spatial mode of cognition was being tapped. Neisser (1972) has been the foremost spokesman for this view, that the distinction being groped for among workers in imagery is between the verbal and spatial modes rather than between the verbal and visual. Blind subjects would certainly need some modality for dealing with objects in space and for imagining such objects in the absence of direct sensory experience. The suggestion is that it is this spatial modality that underlies the mnemonic effect.

Iconic storage then does exist. If mental imagery is not an exact photographic replica of our external environment, the question arises, just what is the nature of these spatial mechanisms used in visual memory?

Visual perception utilizes mechanisms of organization such as figure-ground, and contrast. It is not simply a process of building up sensations into higher-order units. Visual memory utilizes perceptual mechanisms, yet it does so in ways that do not exactly copy our external visions of the world. The results of such research have led experimenters to a wider understanding of the process of perception, and in so doing, have led to experimental designs which seek to clarify further the spatial nature of visual memory mechanisms.

The most recent understanding stated by Segal and Fusells (1970), and confirmed by Brooks (1972), and Shepard (1971-1973), held that perception is a cyclic cognitive activity which includes an anticipation stage, as well as a recognition stage. Imagery for these experimenters was thought to be the anticipatory stage of perception occurring alone. Images, according to Shepard (1973), were not pictures in the head, but plans for obtaining information from potential environments.

Mental Rotation Studies: the Structure of the Construct

The most inventive of the experimental work which explored the possibility of mental imagery being perceptual anticipations were those studies conducted by Shepard, Cooper, and their collaborators (1971-1976). In the first of these studies by Shepard and Metzler (1971), subjects were shown two pictures of geometrical objects and asked if the same object were depicted in both. The two objects were oriented differently, so that subjects had to rotate one of them mentally before the identity task could be performed. The resulting reaction time for such tasks, was a linear function of the degree of mental rotation required to bring the objects into coincidence. The greater the angular difference between the orientations of the two objects, the slower was the response. In essence, Shepard concluded that subjects were apparently carrying out mental rotations at a fixed rate of speed.

The major result of Shepard's work was the assumption that the internal representation underlying the execution of mental rotation tasks was structurally analogous to the operations of the foveal system when one rotated a stimulus externally.

In arguing for a mental rotation analog theory of the structure of the "mind's eye," Metzler and Shepard claimed that their reaction-time experiments revealed that it was not the difference between the two pictures (internal and test shape) which determined reaction time; rather that it was the time required that the subject took to mentally rotate a given shape through a particular path or trajectory in degrees of arc. This cumulative linearity of their time studies pointed indeed to an analog internal rotation structure of visual memory imagery rather than a feature-by-feature comparison theory, as some of their critics have cited.

By analog, Shepard meant a continuous process in which the final mental orientation to a predesignated position of rotation was the result of an additive process whereby the rotated mental image passed through a spatial trajectory, as it were, of measurable degrees of arc. For Shepard, this increasingly linearity of the reaction-time with specific measurable rotational positions made him discount a discrete imaginal processing system in which RT data of such mental orientations would not be cumulative in nature.

What was important in Shepard's work was that the internal representation of an object being manipulated must somehow preserve a kind of structural isomorphism to that real object throughout all intermediate stages of processing, or else the structural information would not be available for comparison with the presented second stimulus. Thus, as their time-linearity studies infer, subjects had no choice but to carry out mental imaginal rotation in a number of small adjustments, each of which preserved the essential structure until the desired orientation was achieved.

Several researchers tried to specify the actual structure of the internally rotated form. John Gould (1972), observed that the mental rotation of images seemed not to vary with the complexity of the stimulus. Most recently, Just and Carpenter (1976), have shown that the greater reaction time in identifying a stimulus with a mentally rotated one, was due perhaps to the angular discrepancies of the compared images. Hochberg and Gellman (1976), have revealed that rotation rates depended on the availability of "landmark features" in the stimulus display, though such features did not necessarily vary with the so-called perceptual complexity of the form.

In 1975, Lynn Cooper carried out further studies on the mental rotation of random two-dimensional shapes which further replicated the findings of Shepard's studies. For Cooper, the reaction time for determining whether a rotated test shape increased linearly with the angular departure of that shape from a previously "seen" orientation. Hence, Cooper (1975), as well as Shepard (1971), and Shepard and Feng (1973), believed that subjects mentally rotated some kind of isomophic internal representation of the test shape into the given required new orientation such that a comparison with a new test shape could be made.

Cooper's unexpected findings regarding the standardization of internal rotation speed despite the visual complexity of the presented shape further enhanced the notion that visual memory imagery was a perceptual process. In furthering the knowledge about the actual internal representation (visual or schematic) of the rotated shapes, it would be only necessary to make very minor subtle distinctions between standard presented shapes and test forms to determine the actual perceptual universals.

Such is the task of this thesis. Specifically, what is needed is the certainty that the "subtle distinctions," as Cooper (1975) calls them, are indeed perceptual in nature. That is, that the distinctions are chosen with careful attention to the processes of perceptual learning in, say, a visual discrimination task. Though indeed Cooper (1975), Gould (1972), and Carpenter (1976), did find that mental rotation speed was standard, despite the given complexity of a stimulus, they did not turn to visual pattern recognition theory in order to be certain that such complexity was judged on a perceptual and not merely a graphic level.

How can one be certain that the design of a given stimulus is perceptually altered? Simply by isolating known features of a perceptual task, such as those identifiable spatial features utilized in a visual discrimination task.

Weaknesses of Previous Rotation Studies

One of the major drawbacks of Cooper's (1975) study was her misunderstanding of the complexity of the presented stimulus. For Cooper, "perceptual complexity" was based solely upon the number of points which determined inflections on the perimeter of the form. Quite simply, Cooper assumed that these points were landmark features, and, like Hochberg and Gellman's (1976), study, their quantity signified complexity. Cooper's notion of complexity was based on Attneave's 1957 data which claimed there was a linear relationship between the logarithm of the number of points and judged complexity.

Further, Attneave's data revealed that the number of points accounted for 80% of the variance of the judgment. To suggest even in 1957 that perceptual complexity is judged accoratedly upon a "points count" in the light of Gestalt perceptual theory is somewhat short-sighted. To suggest in 1975 the same notion is, in the light of more recent perceptual theory--ludicrous. That Hochberg and Gellman should specify the operation of "landmark features" in visual memory encoding tasks is quite a valid statement. However, it is encumbent upon Cooper in 1975 to specify the internal representation of these features in terms far more specific than mere "points."

John Kennedy (1975), in his work on the psychology of picture recognition, has shown that line, edge, occluding bound, occluding edge, texture, color, shadow, and form are all identifiable parts of any given optic array. Furthermore, that line itself is so powerful a visual mechanism that often other features such as texture, color, contrast, brightness, and occlusion may be depicted by it alone. Figure-ground illusions in modern perceptual theory have also pointed to the importance of actual form (as opposed to <u>out</u> line) in visual recognition tasks as did Harmon's studies on the perceptual threshold basis of picture recognition.

Harmon (1973) and Kennedy (1975) have specified the spatial configuration of those features we utilize in a visual discrimination task. It is logical then, to test the perceptual nature of visual memory by isolating a given stimulus with various known spatial configurations used in another closely related perceptual task, that of visual pattern recognition.

The Parameters of This Study

What parameters then need to be isolated in this study of the structural nature of mental image processes?

1. That mental imagery is a perceptual task. That is, mental imagery like other perceptual processes utilizes discriminatory mechanisms which can be specified.

2. That a mental rotation task would be one way to isolate suspected perceptual features.

3. That the mental rotation of these features, isolated in a given stimulus presentation, be tested for the effects of linearity such that the analog (spatial) nature of image encoding be assessed.

If variations in reaction time occur in a mental rotation which utilizes the subtle perceptual distinctions of, say, outline, angulation, and form-fullness, it should be possible, more accurately than in past studies, to specify something of the actual perceptual encoding basis of visual memory.

CHAPTER III

METHOD

Specific Concerns

The problem under investigation in this study concerns the representation mechanisms involved in mental imagery. Specifically it seeks answers to the following questions: What are the structural encoding mechanisms of mental image processes? Are these mechanisms related to other pattern-making processes of the human cognitive system? Can these units of such internal representation be specified?

Because mental imagery has been found to consist of a multiplicity of processes from image transformation to image movement, these general questions need to be specified according to one particular controlled manipulation of one image process--mental rotation.

More directly, this thesis is concerned with the following more specific questions:

1) Do rotated mental images have an analog relationship to externally observed spatial rotation?
2) Does this internal spatial representation utilize the same known perceptual features used when one records one's environment usually?

Hypotheses

The following hypotheses were tested in this investigation:

1) For the main effects of stimulus format, the reaction times for the mental rotation of the stimulus conditions of outline, solidarity, and angulation will not be different.

2) For the main effects of stimulus level (full-ness and broken-ness), the reaction time for the full-ness condition will not be superior to the reaction time for the broken-ness condition.

3) For the main effect of orientation, there will be no linear increase in reaction time, as orientation of the stimulus increases.

4) The interaction effect of stimulus format with orientation will not reveal any greater reaction time data than with the interaction of stimulus level and orientation.

5) The interaction effect of stimulus level with orientation will not reveal any difference in reaction time for the full-ness and broken-ness conditions.

6) The combined interaction effects of stimulus format, stimulus level, and orientation will not reveal any difference in reaction time data from the interaction effects revealed in orientation and each of the stimulus conditions (format and level) separately.

Hebb (1937), has shown how the outline of a figure from its background is immediate at the first moment of stimulation. Deutsch (1953), like Just and Carpenter (1976), has suggested that pattern-coding for shape entailed a mental computation which noted the relative positions of contour (re-positions of corner angles) with the internal representation of that shape.

Previous research then has revealed that outline, corners, and shape are important attributes of visual perception. If mental imagery is spatial in its construct, it too must utilize some of the same building blocks of a spatial schematic as used in vision.

The stimulus used in this study was a randomly chosen shape slightly altered according to those elements of visual perception noted by previous research. Hence, outline, solid, and contour-makers (angulation), became the three variations of a standard stimulus to be mentally rotated by subjects.

These three stimulus variations were further reduced to two further conditions in order to take into account Hochberg and Gellman's (1976) recent findings that mental rotation operated by way of the encoding of specific "Landmark Features," rather than the complete representation of the figure undergoing rotation.

Hochberg and Gellman determined that highlight conditions such as points, corners, etc., were the operative features manipulated in mental rotation tasks. By utilizing a "complete" and "incomplete" version of each of the three stimulus types, Hochberg and Gellman's hypothesis can be tested.

The resulting reaction time for the internal rotation of these six stimuli will presumably reveal if the internal encoded structure is anything like the perceptual format of the externally presented form.

Subject Sample

Twelve subjects were secured from a first-year psychology class subject pool at the University of Manitoba. Of these twelve, the first two acted as pilot subjects such that the experimenter could familiarize himself with the procedural routine.

Previous studies by Shepard and Metzler (1971); and Cooper (1975); have utilized small (ten-subject) samples in order to gain information about image processes. The generalizability of such small samples to a whole population seemed to be based on the belief that image processes are universally spatial in man. Replication of previous studies on the analog nature of mental imagery seems to bear out this fact. Hence, sample size is important only in that an average reaction time for an image rotation task be obtained. In previous studies, as well as this one, reaction time for mental rotation was surprisingly similar enough among subjects to discount different image rotation processes in different people.

Stimuli

In all, six stimulus conditions were presented to subjects. The six stimulus variations of a standard form were:

1. the complete + incomplete outline stimulus;
2. the complete + incomplete solid stimulus;
3. the complete + incomplete corner (angles) stimulus.

These six conditions were felt to be varied adequately enough to test if three common features of visual perception were in any way utilized by image encoding mechanisms. The six stimuli appear in Figure 2 below.

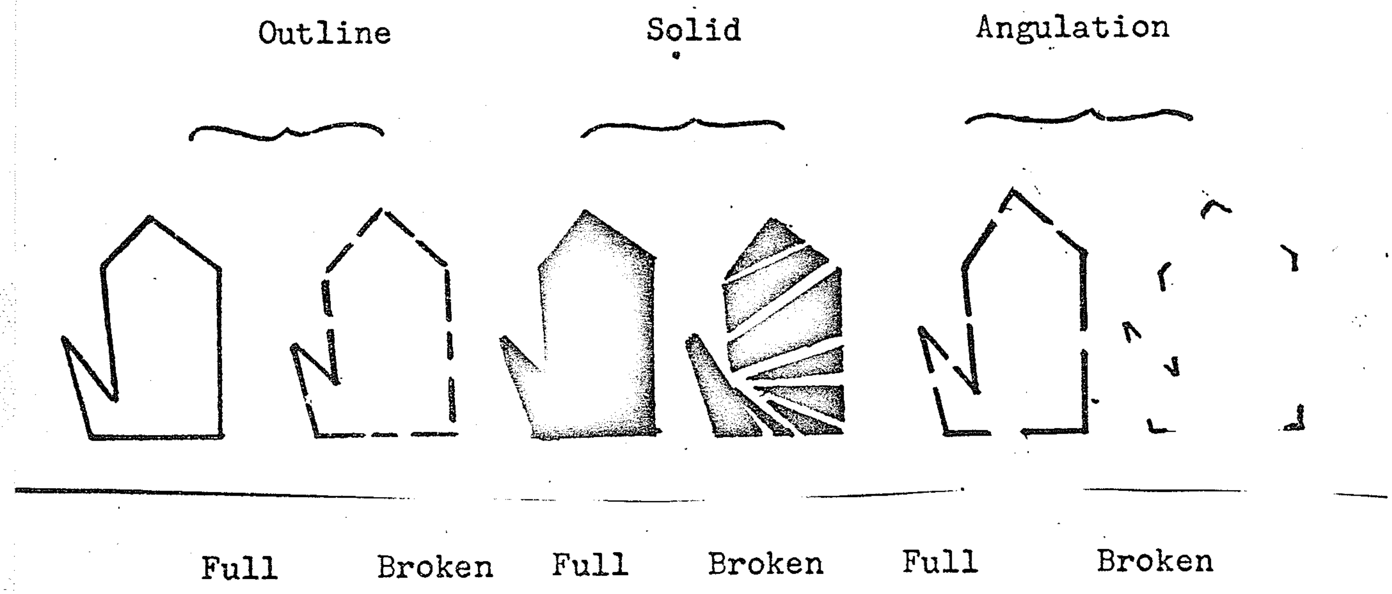


Fig. 2. The stimulus format: outline, solid, angulation, in two conditions, fullness and broken-ness.

The six conditions were presented briefly to subjects in a two-field tachistoscope. After initial top-field presentation of a given random stimulus, subjects mentally rotated an "image" of the stimulus to a pre-determined orientation. This orientation was determined by subjects rotating this mental image to congruence with a bottom-field presentation of the same stimulus at a given orientation. The reported achievement of this rotated congruence with the externally presented orientation was recorded by subjects stopping a timer (see control features).

The orientations which subjects must rotate a given input to, in this study, were 30° variations from 0° through 180°. Previous research revealed that beyond a rotation of 180° there was a pronounced drop in linear reaction time, due perhaps to an internal transformation of an image through its reciprocal. Thirty degree variations in mental orientation were felt to be clear and distinct enough positions, such that a moving internal trajectory could be adequately timed.

Dependent Measures

If subjects are asked to mentally rotate six different stimuli to various positions of orientation, the dependent measures of such rotations are twofold:

1) mental image rotation time (reaction time);
2) error rate. In this forced-choice procedure, sub-

jects could incorrectly choose the wrong presented stimuli as the one which presumably matched their internally rotated one. Such an error could signify a loss of a feature during rotation, and hence loss of the schema, or else a failure to represent and rotate an "image" at all.

Control Features

The measurement of internal rotation time was achieved through a two-field tachistoscope procedure. This tachistoscope procedure enabled the experimenter to control various aspects of the design such that the data obtained would be freer from outside influence and hence more meaningful. The control features imposed in this design were: 1) the presentation format of the stimulus; 2) the measure for internal orientation; 3) the reduction of learned or "guessed" responses.

1. The presentation format of the stimulus. Subjects were shown a total of thirty-six different stimuli for a given (2-sec.) period in the top field of a two-field tachistoscope. These thirty-six different presentations represented six stimulus conditions (outline, solid, and corner aspects of a standard form and their incomplete form), at six positions of orientation. The order of presentation was randomized through a random number table.

Subjects were presented with five blocks of these same thirty-six presentations in order to: a) reduce the effects of individual error in, say, one specific orientation; b) achieve an average reaction time for such internal rotation which might be generalizable to image rotation processes in man.

2. The measure for internal orientation. This was controlled in this design by a forced-choice test for such mental orientation in the lower field of the tachistoscope. The tachistoscope procedures is as follows: Each subject was seated in front of a two-field tachistoscope. One of six stimulus configurations (three variations--line, solid, and angulation; and two sub-formats--full-ness and brokenness) was presented in the top field at 0° orientation for a fixed time period of two seconds. This same stimulus was then presented in one of six rotated positions (from 30° to 180°) in the lower field along with a mirror or incorrect version of the test stimulus in that same rotated position. This second lower field presentation started a digital timer. The duration of this lower field was controlled by the subject. The subject was to choose the correct rotated version of the two presented choices by pressing the corresponding left/right button located below the tachistoscope and directly beneath the left-right choice area of the lower field. The subject's corresponding left/right index fingers were placed on these "choice" buttons at the outset of the experiment. The subject's pressing of

any left/right button stopped the digital timer. A schematic of the experimental procedure appears in Figure 3 below.

3. The reduction of learned or "guessed" responses. This was controlled through the randomized presentation format of the thirty-six top-field tachistoscope presentations. Since subjects could not predict upcoming stimulus presentations and hence evoke short-term memory, the learning (remembrance) factor was reduced. In order to minimize mental fatigue and hence guessed response or outright error in the forced-choice lower-field test, subjects had a short (3-min.) break between each set of randomized presentations.

Design of the Study

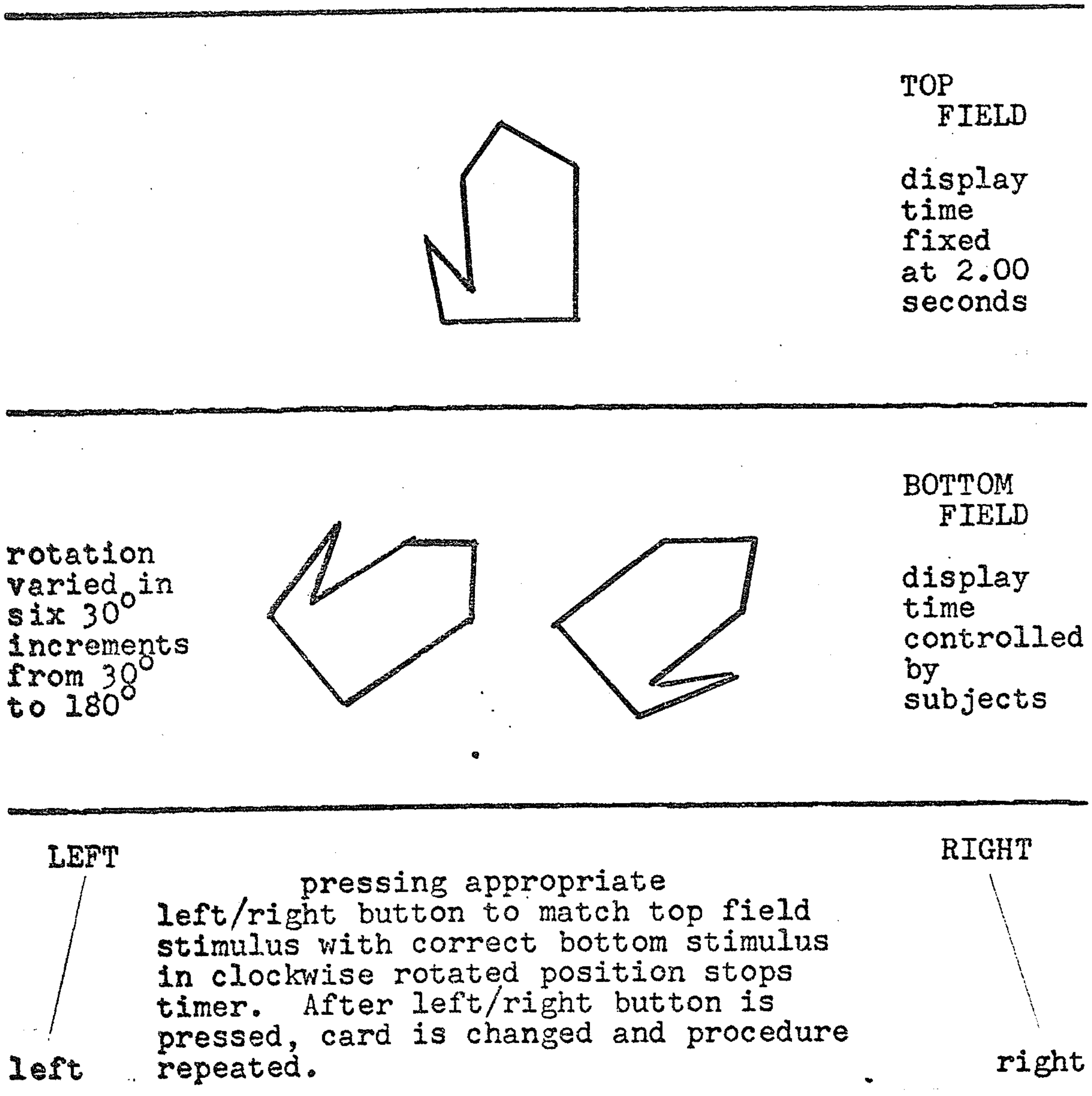
The design model utilized by this study is a 2x3x6 factorial design. The independent variables were: three formats of a given stimulus shape, i.e., an outline shape, a solid shape, and an angulated shape. These stimuli were further broken down into two levels, full and broken-ness for each format. These six stimulus variations were then presented to subjects at 30° increments from 30° to 180°. The independent variables then were: stimulus format, stimulus level, and stimulus orientation. The dependent variables were reaction time and error rate. The design is outlined in Table 2.

Statistical Procedures

The 2x3x6 design of this study permits the measurement of the reaction time for the mental rotation of six stimulus conditions. This measurement of the various stimulus formats (three levels of features times two levels of broken-ness) when crossed with six levels of mental orientation, will presumably yield information on the encoding structure and spatial nature of rotated mental images.

The statistical method used to reveal the interaction effects of reaction time and stimulus type is analysis of variance with repeated measures.

A test for the linearity of data provides a basis for assessing the analog nature of image-encoding processes operating in a mental rotation task. If the reaction time for the mental rotation of any or all stimuli increases in a regular additive manner as orientation increases, then the internally rotated structures are analogous to the structure of real objects rotated in real space, as the reaction time for the visual perception of this real rotation is also linear in nature. A test for the linearity of data therefore reveals the spatial origin of image encoding structures used by subjects.



pressing appropriate left/right button to match top field stimulus with correct bottom stimulus in clockwise rotated position stops timer. After left/right button is pressed, card is changed and procedure repeated.

Fig. 3. Forced-choice 2-field tachistoscope experimental procedure.

Table 2

2x3x6 Factorial Design for Independent Variables

Format	Level	Orientation					
		30°	60°	90°	120°	150°	180°
Outline	Full	$\bar{x}$ RT, 10 subjects	"	"	"	"	"
	Broken	"	"	"	"	"	"
Solid	Full	"	"	"	"	"	"
	Broken	"	"	"	"	"	"
Angulation	Full	"	"	"	"	"	"
	Broken	"	"	"	"	"	"

CHAPTER IV

ANALYSIS OF THE DATA; AND DISCUSSION

Restatement of the Problem

The problem studied in this thesis concerns the interval-encoding mechanisms involved in mental image rotation tasks. Specifically it seeks to clarify the most recently held theoretical position that image representation processes are spatial in construct, and analog to those processes utilized when one visually records the rotation of an object in real space.

The actual spatial nature of the internal image construct was investigated by subjects mentally rotating stimuli which were varied according to three known discriminants of visual perception. If spatial/perceptual features are at the basis of mental image encoding processes, then reaction time measures for the internal encoding and rotation of these perceptually based stimuli would presumably yield clues to the structural nature of this internal spatial construct.

Type of Data Derived in This Study

The data collected in this study are reaction times (in milliseconds) for subjects mentally rotating six different stimuli through six orientations from 30° to 180°. This reaction time data is significant in the specification of mental image rotation mechanisms in two ways.

1) If visual memory rotation tasks are analog to our visual recognition of externally rotated objects, then one could expect the data to indicate a steady and progressive increase in reaction time for the increasing orientation of any given mentally related stimulus. Such a linearity of reaction time data would reveal that, as in real space, time is a function of movement, i.e., more time passes as an object moves farther through space.

If the reaction time data is related linearly to angle of displacement, then it is highly suggestive that the internal representation is spatial in construct in that the encoded "image" does seem to pass through an external-like trajectory. If, on the other hand, the data does not reveal linearity for the mental rotation of a given stimulus, then it can be inferred that digital or other propositional processes are at the heart of mental image mechanisms. Hence the recent theoretical position of the iconic nature of image rotation processes must be accounted for in the weaknesses of the experimental design, or discounted altogether.

2) If the reaction time data <u>is</u> linear in nature and hence indicative that mental image mechanisms are spatial in construct, how can the structure of this construct be further revealed from the data?

Whenever a stimulus presentation is to be committed to visual memory, it is currently believed (Neisser 1973), that certain distinctive features are isolated from the presentation and encoded. John Gould (1972), has shown that rotation time does not seem to be a function of stimulus complexity. However, strict encoding measures and vastly different stimulus presentations were not evident in his experimental procedure.

If presentation stimuli are sufficiently different in perceptual as well as graphic terms, it is inferred that the reaction time data (the sum of the encoding and rotation process) will reflect this complexity through different sets of linear time data for the mental rotation of different stimuli.

In acoustics, a pure tone is often buried in a multiplicity of other sounds. These extraneous sounds are referred to as "noise." It is entirely possible that a given visual stimulus which is to be internally represented suffers from graphic "noise," so to speak, such that the distinctive feature must be pulled from its background. If careful attention is paid to stimulus differences, it should be possible that encoding features be more readily accessible for encoding in certain stimulus presentations. This resultant shorter reaction time set for any given stimulus configuration may be inferred as perhaps indicative as the structure of the encoded iconic construct.

Design Features

The reaction time data are mean scores of five trials per subject at each of six rotated positions for six different stimulus conditions. This reaction time measure is a result of subjects choosing one of two displays, one of which corresponded to their internal orientation. If subjects made the wrong choice in the forced-choice matching test, they were told an error had been made. Error rate in the resultant reaction time data was 4.16%.

Statistical Procedures

The reaction time data collected in this study were subject to two statistical procedures, analysis of variance and linear trend analysis:

1) The 2x3x6 factorial design repeated measures analysis of variance test was carried out with the intention of investigating the interaction effects of orientation and stimulus configuration. Such interaction will presumably yield information on the internally represented form of visual memory, and hence provide clues to the structural nature of such processes.

2) Linear trend analysis provided a basis for assessing the analog nature of image encoding processes operation in mental rotation tasks by revealing whether or not this internal analog is spatial in construct.

Analysis of Variance

Main effects. The Source Table (Table 3) for the 2x3x6 repeated measures analysis of variance revealed:

significant main effects for stimulus format.
($F(2,8) = 23.4$, $p < .001$).
significant main effects for stimulus level.
($F(1,9) = 37.7$, $p < .001$).
significant main effects for orientation.
($F(5,45) = 4.2$, $p < .001$).

The analysis of variance procedure revealed very significant differences for each of the main effects. What do these main effects mean in terms of mental encoding and rotation processes? Simply, the main effect which revealed the greatest amount of variance (full/broken-ness) did, it can be inferred, exert the greatest differential in mental encoding and rotation time.

The implication of this effect is quite clear. One should expect to find different slopes of reaction time within the stimulus level condition, because that level full/broken-ness exerted the greatest variance in the RT measure. If this indeed is so, then the main effect of stimulus level might be considered to be an important attribute of the encoding construct of a mental image.

Other important differences beyond this critical one can be specified from the main effects through the derivation of the mean scores of these effects as shown in Table 4.

From Table 4, within the main effect of stimulus format, the largest difference in response time appeared in the stimulus which accented angulation. This difference was an increase in .96 seconds for response time for mental rotation over the stimulus which accented outline as a possible spatial encoding

Table 3

Source Table for 2x3x6 Design: Biomed P2V Analysis of Variance

Repeated Measures

	Source	Sum of Squares	Degree of Freedom	Mean Squares	F-Ratio	Tail Probability
1.	MEAN	1595.24878	1	1595.24878	114.01	0.000
	ERROR (SUBJECT)	125.92761	9	13.99196		
2.	STIMULUS FORMAT	74.98016	2	37.49037	23.42	0.000
	ERROR (SUBJECT)	28.81383	18	1.60077		
3.	FULL/BROKEN-NESS LEVEL	51.66205	1	51.66205	37.67	0.000
	ERROR (SUBJECT)	12.34382	9	1.82323		
4.	STIMULUS FORMAT x FULL/BROKEN-NESS	71.80757	2	35.90398	19.69	0.000
	ERROR (SUBJECT)	32.81818	18	1.82323		
5.	ORIENTATION	12.97823	5	2.59365	4.17	0.000
	ERROR (SUBJECT)	27.96445	45	0.62143		
6.	STIMULUS FORMAT x ROTATION	3.08429	10	0.30843	0.60	0.812
	ERROR (SUBJECT)	46.44159	90	0.51602		
7.	FULL/BROKEN-NESS x ROTATION	5.77870	5	1.15574	2.20	0.071
	ERROR(SUBJECT)	23.63885	45	0.52531		
8.	STIMULUS FORMAT x FULL/BROKEN-NESS x ROTATION	3.03819	10	0.30382	0.47	0.904
	ERROR (SUBJECT)	57.93190	90	0.64369		

Table 4

Means of Main Effects*

Main Effects	
Stimulus format	
1. Outline	$\bar{x}_1 = 1.79$
2. Solid	$\bar{x}_2 = 1.62$
3. Angulation	$\bar{x}_3 = 2.75$
Stimulus Level: Full/Broken-ness	
1. Full	$\bar{x}_1 = .70$
2. Broken-ness	$\bar{x}_2 = 1.145$
Orientation	
1. 30°	$\bar{x}_1 = 1.92$
2. 60°	$\bar{x}_2 = 1.91$
3. 90°	$\bar{x}_3 = 2.03$
4. 120°	$\bar{x}_4 = 2.23$
5. 150°	$\bar{x}_5 = 2.08$
6. 180°	$\bar{x}_6 = 2.46$

format, and a 1.13 sec. increase over a stimulus which accented a solid shape as the structural mechanism of representation and interval rotation.

---

*It should be noted that the derivation of these mean scores was done arithmetically. A more precise measure would have resulted through the use of a multiple comparisons test.

Of the stimulus level, level 2, ...brokenness, revealed the greatest increase in reaction time for mental rotation. Presumably it took .445 sec. longer to mentally rotate this stimulus condition than it did to rotate a stimulus which was not divided in any way from its "full" graphic description.

The main effect of orientation revealed that the angular displacement of 180° resulted in the greatest average increase of response time for mental rotation over other specific positions of angular displacement. On average, there was a .54 sec. increase in response time for mental rotation of a stimulus from a 30° to a 180° angular displacement. Though response time for mental rotation did increase as the position of angular displacement increased, this increase was not regular.

How are these specific differences in the means of the main effects to be interpreted; and how are these results to be viewed against the response time data which is subjected to further statistical procedures, such as a test for linearity, and interaction effects of an analysis of variance?

From the means table (Table 4), it is apparent that the greatest increase in response time for mental rotation occurs in the stimulus format of angulation. In terms of the internal representation mechanism, it can be inferred that this stimulus format presents the greatest difficulty for subjects to encode and mentally rotate. Similarly, the stimulus level "broken-ness" revealed the greatest increase in response time because it too presumably presented the greatest visual ambiguity from which an internal representation must be taken. Had the angulation and broken stimulus conditions been encoded as presented, it can be inferred that there would be less difference from the mean response times cited for the rotation of the other stimulus conditions.

One other possibility is that the internal representation of these two stimulus conditions did occur as presented. Yet this particular internal construct was such that a slower mental rotation speed was necessary in order to prevent "loss" of the construct while mental rotation was in progress. If this second assumption were true, one could expect a different slope for the full/broken interaction of response time for mental rotation, with orientation for these two stimulus conditions. The results of the interaction effects in the analysis of variance procedure, as well as a linear trend analysis, do not categorically uphold this view.

Interaction Effects of Special Interest

Of particular interest in this study are the interaction effects of the stimulus conditions (format and level) with orientation. The interaction of stimulus format and orientation revealed an F-value of .60 at the .812 level of significance. Clearly, the level of significance of such interaction is well above the .05 level which would be meaningful in this study. Therefore, the format of those stimuli which accented the perceptual feature of outline, solid-ness, and angulation asserted no special role in affecting response times in a mental rotation task.

The computed F value for the interaction of the stimulus level (full/broken-ness) with orientation was given in Table 3 as 2.20 at the 0.07 level of significance. For the purposes of this study, this interaction effect is well above the acceptable 0.05 level of significance and is, therefore, an interaction which is not significant. However, at the 0.07 level this interaction is showing an interactive effect which might be inflated due to the large degree of hidden variance revealed in the stimulus conditions themselves.

How can this interaction effect, which is approaching significance, be specified such that the degree of interaction be more readily observed?

A graphical depiction of the interaction of "full/broken-ness" and orientation was found by first finding the mean RT (reaction time) score for every point of orientation for one stimulus designation, say fullness, as it occurred in all three stimulus formats (line, form, angle).

This mean RT score was then plotted for each of the orientation positions and the line joined. The same procedure was followed for the "broken-ness" stimulus level, and plotted. Figure 4 reveals the results of this interaction.

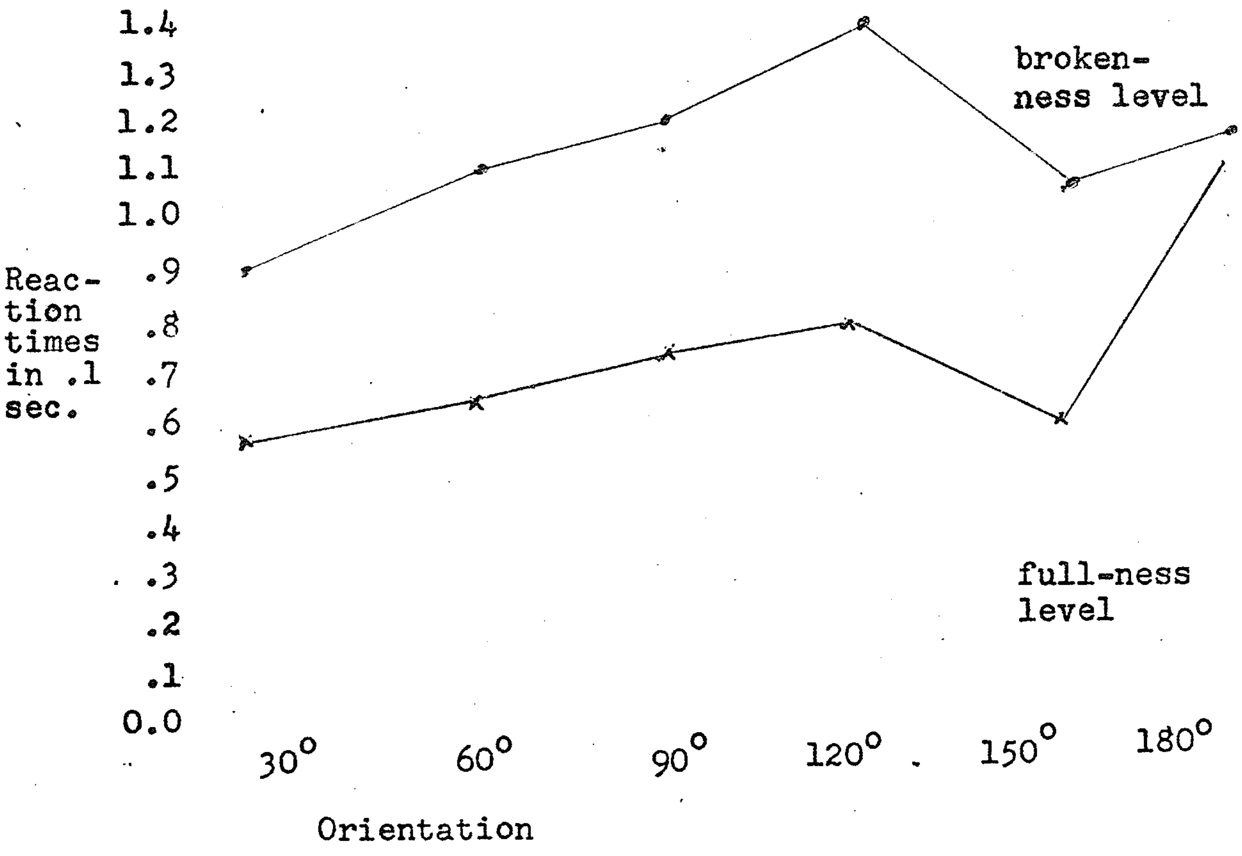


Fig. 4. Interaction plot of stimulus level with orientation

What we are interested in through this figural depiction of stimulus level with orientation, is the depiction of different slopes for the "full-ness" and "broken-ness" stimulus levels. Different slopes for each level would be indicative of a difference in reaction time for the internal representation and rotation of "full" stimulus formats (line, solid, angulation) over "broken" stimulus formats.

Figure 4 does clearly suggest two different slopes for each of these two stimulus levels, at least to the 120° orientation. Beyond this orientation, the clarity of the separate slopes is lost, though a strange parallel reduction of RT occurs at the 150° and 180° orientation positions for both slopes. Hence the degree of interaction <u>is</u> seemingly more readily clarified by Figure 4. However, the interaction is not statistically significant, and though it presents some interesting highlights for further analysis, it must be concluded that neither the full nor broken variants of the stimulus formats of line, solidness, and angularity directly affect the response time for mental rotation.

Yet the problem still persists! Subjects are obviously encoding and rotating a reported graphic or spatial internal construct. Thus far we have failed to validate that a variance in stimulus presentations according to common visual discriminants of outline, solidness, and angulation have anything to do with the internally represented and rotated schema. That this schema is iconic in nature is inferred by a linear trend analysis.

Linear Trend Analysis

The linearity of reaction time for mental rotation data is indicative of the analog processes of rotation in visual memory. Linear reaction time data is highly suggestive of the iconic nature of image encoding mechanisms, or at least indicative of similar processes operating in visual memory when one records the external rotation of an object via the human foveal system. The computations and results of the linear trend analysis are given in Tables 5 through 10.

In summary, the linear trend analyses revealed that only two of the six stimulus conditions (full solid and broken solid) showed any clear linearity of reaction time for mental rotation. The presentation formats of the other stimulus conditions when subjected to mental rotation were presumably sufficiently ambiguous so as to prevent linearity.

In all, the data from this experimental design supported the two null hypothesis. That is,

1) there was <u>no</u> significant difference in the mean sets

TABLE 5

Computational Methods Used in Testing Linearity of Trend: Full Line Stimulus Condition

	Orientation Positions in Degrees						Sum Orientation
	30	60	90	120	150	180	
Mean RT's for Full Line Orientation (10 Subjects)	15.9	16.9	15.6	16.6	17.6	22.1	104.7
Linear Coefficients	-5	-3	-1	1	3	5	
Coefficients x RT	-79.5	-50.7	-15.6	16.6	52.8	110.5	34.1

Sum squares for full line = $\frac{(34.1)^2}{700} = \frac{1162.81}{700} = 1.66$

Mean squares for full line = $\frac{104.7 - \frac{(104.7)^2}{60} = 104.7 - 182.7 =}{54}$

$-78 = 1.44$

F Ratio $\frac{1.66}{1.44} = 1.152$

Tabled value of F for linearity at .05 level of significance is 4.08.

∴ mental rotation with full line stimulus configuration is non-linear.

TABLE 6

Computational Methods Used in Testing Linearity of Trend: Broken Line Stimulus Condition

	Orientation Positions in Degrees						Sum Orientation
	30	60	90	120	150	180	
Mean RT's for broken line orientation (10 subjects)	17.2	14.6	16.7	20.0	19.0	21.3	108.8
Linear Coefficients	-5	-3	-1	1	3	5	
Coefficients x RT	-86.0	-43.8	-16.7	20.0	57	106.5	37.0

Sum squares for broken line = $\frac{(37.0)^2}{} = \frac{1369}{700} = 1.96$

Mean squares for broken line = $\frac{108.8 - \frac{(108.8)^2}{60} = 108.0 - 197.29 =}{54}$

$-88.49 = 1.64$

F Ratio = $\frac{1.96}{1.64} = 1.195$

Tabled value of F for linearity at .05 level of significance is 4.08.

∴ mental rotation with broken line stimulus configuration is non-linear.

TABLE 7

Computational Methods Used in Testing Linearity of Trend: Full Solid

	Orientation Positions in Degrees						Sum Orientation
	30	60	90	120	150	180	
Mean RT's for full form orientation (10 subjects)							103.6
Linear Coefficients	-5	-3	-1	1	3	5	
Coefficients x RT	-65.5	-43.8	-16.5	17.1	55.8	118.5	65.6

Sum squares for full form = $\frac{(65.6)^2}{700} = \frac{4303.36}{7000} = 6.15$

Mean squares for full form = $\frac{103.6 - \frac{(103.6)^2}{60} = 103.6 - 178.88 =}{54}$.

F Ratio = $\frac{6.15}{1.39} = 4.42$

Tabled value of F for linearity at .05 level of significance is 4.08.

$\therefore$ mental orientation with full form stimulus configuration is linear.

TABLE 8

Computational Methods Used in Testing Linearity of Trend: Broken Solid

	Orientation Positions in Degrees						Sum Orientation
	30	60	90	120	150	180	
Mean RT's for broken form orientation (10 subjects)	17.5	18.0	17.9	19.0	20.5	20.3	113.2
Linear Coefficients	-5	-3	-1	1	3	5	
Coefficients x RT	-87.5	-54	17.9	19.0	61.5	101.5	91.1

Sum squares for broken form $= \frac{(91.1)^2}{700} = \frac{8299.21}{700} = 11.86$

Mean squares for broken form $= \frac{113.2 = \frac{(113.2)^2}{60} = 113.2 - 213.57 =}{54}$

$-100.37 = 1.86$

F Ratio $= \frac{11.86}{1.86} = 6.38$

Tabled value of F for linearity at .05 level of significance is 4.08.

. . mental orientation with broken form stimulus configuration is linear.

TABLE 9

Computational Methods Used in Testing Linearity of Trend: Full Angle

	Orientation Positions in Degrees						Sum Orientation
	30	60	90	120	160	180	
Mean RT's for full angle orientation (10 subjects)	15.7	14.8	15.6	16.7	16.7	22.3	101.8
Linear Coefficients	-5	-3	-1	1	3	5	
Coefficients x RT	-78.5	-44.4	15.6	16.7	50.1	111.5	39.8

Sum squares for full angle = $\frac{(39.8)^2}{700} = \frac{1584.04}{700} = 2.26$

Mean squares for full angle = $\frac{101.8 - \frac{(101.8)^2}{60} = 101.8 - 172.72 =}{54}$

$-70.92 = 1.31$

F Ratio = $\frac{2.26}{1.31} = 1.73$

Tabled value of F for linearity at .05 level of significance is 4.08.

∴ orientation with full angle stimulus configuration is non-linear.

TABLE 10

Computational Methods Used in Testing Linearity of Trend: Broken Angle*

	Orientation Positions in Degrees						Sum Orientation
	30	60	90	120	160	180	
Mean RT's for broken angle orientation (10 subjects)	39.2	35.7	46.6	40.6	45.1	39.3	246.5
Linear coefficients	-5	-3	-1	1	3	5	
Coefficients x RT	-196	-107.7	-46.6	40.6	135.3	196.5	22.7

Sum squares for broken angle = $\frac{(22.7)^2}{700} = \frac{515.29}{700} = 0.74$

Mean squares for broken angle = $\dfrac{246.5 - \dfrac{(246.5)^2}{60} = 246.5 - 1012.70 =}{54}$

$766.2 = 14.19$

F Ratio = $\frac{0.74}{14.19} = 0.052$

Tabled value of F for linearity at .05 level of significance is 4.08.

$\therefore$ orientation with broken angle stimulus configuration is non-linear.

---

*Test for Linearity of Trend from Roger E. Kirk, *Experimental Design Procedures for the Behavioral Sciences* (Belmont: Cole Publishing, 1969), p. 120.

of reaction time for six different stimulus conditions undergoing mental rotation;

2) in a test for linearity, the mean reaction time for mental rotation will be less than the given F value. Hence the mean will not be significant and the reaction time progression was not linear.

Yet, as revealed previously, the rejection of the null hypotheses was not consistent, nor was it entirely conclusive. The linearity of the full and broken solid stimulus conditions upholds Shepard's (1971; 1973) research that mental rotation tasks are effected by an encoding system which is spatial in its construct. Further, though the stimulus level with orientation interaction was not statistically significant, it did graphically reveal a difference in slopes for the "full" and "broken" stimulus conditions. Though this difference in slope for the two conditions was not entirely parallel or consistent throughout all orientation, it does bear some discussion in the light of the inconclusive data presented in this study.

How then might these two curious anomalies in the data be interpreted? And what do these inconclusive results mean, both in terms of this experimental design and in terms of current mental rotation theory?

Interpretation of Some Anomalous Results

Full and broken solid stimulus conditions showed a clear linearity of reaction time for increased rotation. Why should this linearity appear in this stimulus format and not in others?

Subjects revealed in a post-test questionnaire that they found the recognition of "solid" stimuli far easier to "remember" and "rotate" than other stimulus conditions. One subject reported the broken angle stimulus condition beyond the 30° orientation as completely unrecognizable. Clearly, isolation of a feature in a mental encoding and rotation task is due to the ambiguity threshold of a particular stimulus. It seems that stimuli which present to the subject the least amount of visual ambiguity (lines, corners, edges, pieces, etc.) are readily represented. A solidly depicted stimulus presents very little choice in terms of encoding possibilities other than its holistic depiction. Perhaps the notion of "no choice" is important in visual encoding. Where stimulus presentations permit a variety of edges, corners and outlines which might be featured, this very variety mitigates against the memory of a feature which must be "held" during rotation. It is possible therefore, that mental image encoding mechanisms are not feature-analytic in nature. The encoding feature may be any depiction which presents the least possible spatial ambiguity.

An obvious but unproved interpretation of the linearity of the solid stimulus conditions reaction time is that solidness is a feature of image encoding which is utilized more often than other presentation formats. Though visually, we schematize images by the utilization of outline drawings, it is a distinct possibility that image-encoding mechanisms utilize actual solid spatial structural representation mechanisms. Allport (1928), recorded subjects noting solid visual memory displays in after-image experiments. Further support of the use of this structural format is mentioned by McKinney & Hebb (1953), whose subjects also reported "solid" areas of visual loss in an experimental design which stopped saccadec eye movement and effected image loss.

At least these early introspective reports do point to the need for further empirical research on the distinctive encoding property of fully defined shapes, as opposed to less clearly defined stimulus conditions.

Though in Figure 4, no exact differentiation of separate slopes was depicted, there was strong graphic evidence of the beginnings of such slopes. This, it can be assumed, is indicative of different encoding times for the two conditions of the stimulus level. This differentiation of the two slopes is however lost at the 150° and 180° positions of orientation.

Why should this be so? The fact that the reduction of reaction time for both full and broken conditions at the 150° orientation can be assumed to be due to an embedded feature of each stimulus condition assuming primacy at this orientation. Obviously, the whole figure seems not to be rotated to this orientation. What may be occurring is that a "parts-analysis" is overriding the initial presented Gestalt such that the angular displacement of the stimulus is sped up through the "loss" of other unnecessary features.

The graph also reveals that the stimulus condition of broken-ness took longer to encode and rotate than the fullness condition. Again, the visual ambiguity of the broken stimulus conditions would suggest that a sorting procedure may be being used until an acceptable encoding pattern or representation is derived from the ambiguous presentation. The reduced slope of the "fullness" conditions can be assumed to be due to the lesser amount of presented ambiguity of such figures. Hence an encoding format is more quickly pulled from this condition.

Though this is entirely speculative, the apparent difference in the slope conditions of Figure 4 does suggest that the encoding and rotating representation of an image in memory does seem to be affected by the configuration of the presented test stimulus. The problem of reducing the presentation image to the encoded representation still remains undone.

Though this speculative consideration is in no way meant to override the statistics revealed in this study, it does point to weaknesses in the design which could account for the only marginally positive data.

Weaknesses in the Study

The interpretations above pointed to weaknesses not only in the statistical procedures used in this study but also in the nature of design procedure itself. All of these weaknesses could presumably affect the nature of the response time data which was collected. Weaknesses occurred in: 1) the tachistoscope presentation procedure; 2) the depiction of the presentation stimuli; and 3) the initial grouping of response times for the tachistoscope trials into mean scores. These criticisms are dealt with more specifically as follows:

1) The tachistoscope presentation procedure

Previous studies (Cooper, 1975) revealed that a mental rotation task consisted of two distinct internal mechanisms, the encoding of a stimulus, and the separate rotation of that encoded "image." Other studies (Cooper, 1975) made accommodation for these separate mechanisms by utilizing distinct response time measures. This study utilized only one measure of response time as being adequate in accounting for encoding and rotation tasks.

The difficulty in the use of only one measure of response time for two distinct mental processes is such that a feature search of a stimulus which is readily available in one stimulus presentation may not be available at all in another stimulus presentation. Thus before a change in feature processes mechanism could take place, search time for known features apparent in other stimulus displays is being recorded.

In the post-test questionnaire, subjects did report a search procedure in new stimulus displays for isolated parts "remembered" from previous displays. Due to the changing nature of the stimulus presentations, these memory feature searches were often fruitless and used up time before a new feature extraction process was activated. Hence, the one recorded reaction time measure could hypothetically be a combination of two completely isolated and highly inordinate internal encoding time tasks.

A more serious weakness of the one response time measure lay in the forced-choice task which was to record not only the response time for the angular displacement of an internally rotated image, but also the "correctness" of that displaced internal representation.

Subjects were presented after an initial presentation time for a particular stimulus, with a lower field tachistoscope display of <u>two</u> rotated images, one a mirror image of the other. Only one of these two lower field displays was the "correct" rotated version of the initially presented test stimulus. Presented with two displays simultaneously, it is possible that subjects could compare a given display with the internally rotated "image," reject it as being wrong, and then proceed to match and confirm the other display as being the "correct" rotated form.

However, it is also possible that subjects could be "correct" in their first choice of any one of the two lower field displays. Such ambiguities of a 1) search and accept, and 2) search--reject + search again and accept--procedure could conceivably distort all response time data from any semblance of linearity at all.

In order to eliminate the possibility of a double choice and hence a two-search response time measure, it would be necessary to utilize only those response times which indicated correct matches in one left or right choice. However, though this weakness is an important one, there is evidence in the recorded error times that this dual-choice problem was not overly responsible for distorting response times collected in this study.

Subjects pressed the appropriate left/right button which stopped a digital timer. This recorded on which side of the lower field display their choice was presented. If subjects chose an incorrect "mirror" image as similar to the internally rotated image in that orientation, they were immediately told that an ERROR had been made.

Overall error rate for 180 trials by ten subjects was 4.16%. Response time for "correct" matches of internal orientation with the externally rotated stimulus ranged from 0.668 sec. at 30° orientation to 3.00 sec. at 180° orientation. The response time for ERRORS was significantly higher than for correct choices. A sampling of the ERRORS recorded for the rotation of the reported most difficult angulation stimuli revealed an error response time of 2.414 for a 30° orientation, and an error response time of 5.012 sec. for an incorrectly chosen match of a rotated stimulus at 180°.

Given that this error level was, in all cases, far above the response time recorded for correct matches, it can be assumed that the dual search procedure was carried on only in those responses which were designated as Error. Those response times which showed an inordinate increase of response time which could be indicative of a dual search process, though being "correct" choices, were counted and treated as error. Even these excessive, though "correct" response times still only raised the overall error rate in this study to 6.91%.

Therefore though the forced-choice lower field display could readily distort the response time data, it seemed not to do so inordinately in this particular study.

2) The Presentation Format of the Stimuli

The configuration of the test stimulus in the forced-choice, two-field tachistoscope procedure always appeared at 0° orientation. The consistent orientation of a test stimulus allowed subjects to disregard the presentation feature of a new test stimulus, in favour of an internal representation of a previous stimulus stored in short term memory. In two cases, the post-test questionnaire revealed that subjects remembered features from previous stimuli in order to encode and rotate a "present" stimulus configuration. In one case, a subject reported utilizing features not designed for an experimental stimulus at all.

3) Sample Size

The small sample size (ten subjects), and the use of a mean score for the five subject trials introduced an error factor in the design procedure. Though subject error was recorded at 4.16%, the error score was still utilized in the computation of a mean score. Hence, the mean score was not a strictly accurate measure of the correct reaction time for the mental rotation of a given stimulus.

In spite of these weaknesses of the design, data were recorded which confirmed a linear time function for mental rotation in one stimulus condition. The post-test questionnaire revealed that the other two stimulus conditions were highly ambiguous visually. Hence, it is concluded that the subtlety of the presentation stimuli in two of three cases was not subtle enough perceptually to isolate singular features which could have been internally represented in a mental rotation task.

Data was recorded which revealed a possible interaction between "full" and "broken" stimulus levels with orientation, if the speculated perceptual "cross-over" phenomenon were removed. However, it must be accepted that this study did not clearly reveal any statistical support for the full/broken-ness stimulus distinctions as having a direct effect on response time for mental rotation.

The analog nature of image rotation tasks was again supported by this design only marginally. No clear indication of a proportionate increase in reaction time for increased mental rotation of a given set of stimuli was found. Hence the various stimulus conditions presented in this thesis were not encoded as presented. Linearity was recorded only in one of three stimulus conditions.

However slight, the inconsistency of the data revealed in this experimental design cannot entirely disprove the analogical understanding of imaging processes. With limited success, as revealed earlier in this chapter, the spatial encoding organizational mechanism for mental imagery is upheld. It is the belief of this writer that if some of the weaknesses cited earlier in this chapter for this study were removed in further experimental procedures, data much more consistent to this theoretical position could be achieved. Further implications of the results of this data for current image rotation theory are considered in the next chapter.

CHAPTER V

DISCUSSION AND CONCLUSIONS

Overview of Study and Results

The purpose of this study was to verify empirically something of the structural nature of visual memory. Introspective reports on mental imagery likened it to a process of "seeing with the mind's eye." Our subjective experience of mental images has been compared to wax tablets, painted portraits, photographs, and even internal motion pictures--complete with living colour! Just as a photograph or a motion picture is itself a representation of reality, so too mental images were considered internal reproductions, passively recorded upon one's internal memory trace. This subjective, metaphorical understanding of mental imagery downgraded the role ofimaging as a functional mechanism of mind, and created within our everyday language system, a way of talking about visual memory which made the metaphor assume a literal interpretation.

Recent research, as well as that obtained by this study, found the metaphorical idea that images are picture-like reproductions, as inappropriate. Instead, the results of this and earlier studies view mental imagery as an active, dynamic process, much like perceptual-motor activity. Further, some of the properties of this dynamic process were inferred from objective performance in an image rotation task, rather than from introspection.

Subjects were presented with six variations of a standard shaped stimulus, which had been varied according to three known features of visual perception--namely, an outline format, a "solid" format, and a format in which only corners (angles) were featured. These three presentation formats were further reduced into incomplete versions of the standard presentation stimuli, such that subjects were presented with a total of six stimulus conditions.

Upon presentation of the separate stimuli, subjects were asked to mentally rotate them to known positions of orientation. Reaction times for encoding and rotation of all presented stimuli were recorded by way of a forced choice test. The designation of the correct "choice" format with the mentally rotated stimulus of the same orientation stopped a digital timer.

It was predicted that different presentation stimuli would present different encoding features to the subject and hence result in various reaction times for mental representation and rotation. Such differences could be used as clues to the actual structural nature of the mental image rotation mechanism. It was further predicted that reaction times would increase linearly as a function of the number of degrees of mental rotation.

Such predictions were based on the theory that mental images were analogical representations of perceptual information.

Summary of Emergent Trends

The results of the data upheld the analog representation theory of mental imagery only marginally. All presented stimulus conditions did not result in linear reaction times for mental rotation. However, those stimuli that did not, were felt to be poor, ill-defined stimulus conditions which were difficult to encode, and even more difficult to rotate through an internal trajectory.

The assumption that the encoding feature of mental image manipulation tasks such as rotation, were perceptual in origin was not supported by this study. It was clear, however, from a graphic extrapolation of recorded significance levels of complete (full), and incomplete (broken), stimulus conditions that the presentation format of a given stimulus which is to be represented internally, is a definite attribute of that internalization. However, the degree of comparison between external stimulus presentation, and internal code was not specified.

Mental imagery, then, utilizes encoding and transformational mechanisms which are distinct from mechanisms utilized by language or the manipulation of numbers. The linear data, even though not consistent throughout all mental rotations involved in this study, supports the idea that mental images are spatial in origin. The questions as to the specific structural format of the internal spatial organization of mental images and their determinant features still remain unanswered.

Implications in the Data for Image Rotation Theory

Mental rotation theory as advanced by Shepard, 1971, 1973, and Cooper, 1975 suggests that the internal representation underlying the execution of mental rotation tasks was structurally analogous to the operations of the foveal system when one rotated an image externally.

In arguing for this analog theory, Shepard claimed that their reaction time experiments revealed that it was not the difference between the two images (internal and test shape) which determined reaction time; rather, it was the time required for the subject to mentally rotate a given spatially encoded image through a particular trajectory to some inner orientation. Shepard concluded that the linearity of reaction time data pointed to this internal spatial construct which like its externally rotated counterpart increased in response time as orientation increased. Other processes, such as a feature comparison theory, would not, Shepard believed, yield data which increased linearly as the internal angle of displacement increased.

This study did, I feel, uphold Shepard's hypothesis, though it did not uphold it categorically. Linear reaction time data was found to exist for two of six stimuli which were mentally rotated. The failure of other stimuli to yield similar linear data was felt to be largely a fault of the experimental design.

Other than a speculative hunch, no real clue as to the spatial construct of these internally rotated images was found.

Suggestions for Further Study

If the basic encoding mechanisms of visual memory are spatial in origin as evident through this study, how can this spatial representational system be specified? More importantly, are the determinants of this internal organization perceptual in origin? That is, does mental image encoding and transformation mechanisms utilize features which are recognized as basic to the recognition of one's environment foveally?

The presentation of stimuli to be mentally rotated by subjects must be very carefully delineated. The simple isolation of a perceptual feature such as "outline" (occluding edge) or "angulation" (corner, or landmark feature) which is to be encoded for mental rotation is ineffectual since one has <u>no</u> way of knowing if embedded unknown features are being internally represented, or whether indeed utilized by encoding mechanisms of visual memory at all. A better way of getting at the root of the structural nature of mental imagery processes might be to utilize stimuli that are varied, not by the distinct isolation of visual percepts, but stimuli which utilize perceptual "threshold phenomena" as the basis for their distinctiveness.

If images fade from short term memory in an ordered fashion, as Hebb has shown, and if recognition thresholds for photographs can be specified by the amount of contrast a given photograph contains (Harmon, 1973), then the distinctive features of mental imagery might be specified more readily through the use of presentation stimuli which are "graded" in their recognizability. Such "threshold stimuli" would clearly remove embedded features that otherwise could contaminate a given presentation stimulus display.

A different way of approaching the problem of the structural nature of mental imagery might be from the more recent expression of the organizational distinction found in information processing approaches to perception and memory. Neisser (1967), and Paivio (1971), have defined two processing models, that is, parallel and serial processing, and the subdivisions of these into spatially parallel, operationally parallel, serial, and sequential processing systems.

The defining property of spatially-parallel processing is simultaneity of functioning. The visual system is an example in that simultaneously given information can be processed over a broad area of the retina. Similarly in visual memory, in a description of, say, a living room from memory, the information carried by the imagery system is assumed to be simultaneously available for processing, though of course the verbal description of that memory image would be sequential in its delivery.

It might be possible to present to subjects, instead of the one stimulus format, a whole array of stimulus conditions simultaneously. If these "large scale" stimuli presented to subjects for encoding into visual memory were "scaled" or "ranged" in terms of specific visual features, those large scale displays which were most easily remembered from a stack of such presentations might provide clues as to the structural unit utilized in visual memory.

Conclusions and Implications for Education

Mental imagery is a distinct mechanism of mind. Though as yet, the specification of its structural nature other than spatial or quasi-perceptual in origin is unknown, mental imagery encoding and transformation mechanisms are characterized by: remarkable speed (milliseconds for rotation), accuracy (less than 5% error in some 360 difficult and confusing matching tests), and great flexibility (rotation).

Apparently, mental images can be encoded quickly into synchronously-organized integrated spatial compounds that function somehow as units in memory. Moreover, these spatial layouts can be retrieved by other modalities, such as language.

It is known (Miller 1957), that groups of words cannot be so easily integrated into memory. Further, words are encoded into linear informational structures and are subject to the sequential constraints of syntax upon retrieval.

While it is too early to make clear statements about the distinctive features of imaginal encoding and transformational mechanisms, it is clear that mental imagery is a powerful and dynamic mechanism of mind.

Why then have educators been slow to incorporate the use of this mental activity into school curricula? The answer lies more in an outdated theory of learning which prevails in education than in teachers' ignorance of the concept of spatial ability.

In 1951, Barakat began a series of experiments in English public schools which recorded that a "spatial factor" was utilized by adolescents who exhibited proficiency in algebraic and geometric abilities. Barakat also found that at purely mechanical arithmetic problems, those students who reported poor imaginal processes, usually did better.

Blout (1971), replicated educational research of the 1950's which linked reading difficulty in young children to their inability to encode spatial (graphic) symbols as presented in alphabets. Blout went further and revealed that the order of teaching concepts to young children was performed with little attention to the spatial ability found in these same children. He proposed that map reading (i.e., spatial recognition ability) should precede reading instruction. Presenting a child with a multi-dimensional set of symbols so readily adaptable to his spatial system would, Blout believed, provide a basis for the child's learning a set of very complex, linearly ordered symbols, which is the alphabet.

Given this educational research, why should teachers still largely neglect a child's spatial ability? Simply because visual perception is considered by most educators a process not as rigorous as cognition, and hence not as important a learning tool to be included in the schools' curricula.

Most educators see perception as an activity of the senses. Cognition, according to the old paradigm, begins where the work of the senses ends. Yet this study has shown that such operations as selection, simplification, abstraction, comparison, synthesis, and putting into context, are not the sole domains of concept formation, but the active ingredients of perception itself.

For teachers, "cognition" needs to include the word "perception," for no thought process operates outside perception. Thinking is perception, and thinking with the mind's eye is still a special kind of spatial cognitive process. Before teachers adopt the special non-linear processes involved in mental imagery as unique and powerful tools to be utilized in schools, they must first see visual perception and mental imagery in this new paradigmatic light.

Mental imagery has great influence on other cognitive modalities. The fact that there is a remarkable interplay between imaginal and verbal processes, and the fact that mental imagery is free from the sequential constraints of language gives it a unique position in the generation of new ideas.

It must be remembered that the description of the benzine molecule, as well as the bonding mechanism of the D.N.A. helix, are direct results of image processes, the results of which changed radically the understanding of molecular bonding. The further specification of the "building blocks" of imaginal organization can only result in a more direct access and use of this cognitive ability for solving even more pressing problems.

Reading, writing, and arithmetic are cognitive abilities which are sequential and logical in nature. It is important that these abilities be developed in educational institutions. However, ideation is the result of intuitive leaps of the imagination; that is the result of flexible non-sequential transformations of perceptual information--all attributes of one's image system.

The person who knows how to organize a visual pattern or who knows the variety of forms and techniques for depicting such patterns graphically or in memory must surely be one with an extra problem-solving ability beyond language and propositional mechanisms.

It is not enough then for teachers to merely turn on the movie projector in the classroom, or to pay lip service to the doctrine of visual aids. What is needed is the systematic understanding and training of a visual sensitivity as an indispensable part of any teacher's preparation for his profession.

It is fitting that the etymological root of the word "idea" derives from the Greek <u>idein</u>, to see. Perhaps when psychologists reveal more of the spatial representational mechanics of vision and visual memory, will teachers then accept the responsibility for monitoring and nurturing all the problem-solving capacities of children? At this time it is hoped adequate curricula will be developed which will train students to recognize and utilize this powerful mental process more fully.

BIBLIOGRAPHY

BIBLIOGRAPHY

Allport, G. W. Eidetic Imagery. British Journal of Psychology, 1924, 15 19-110.

Allport, G. W. The eidetic image and the after image. American Journal of Psychology, 1928, 40 418-425.

Aserinsky, E., and Kleitman, N. "Regularly occurring periods of eye motility and concomitant phenomena during sleep." Science. 1953, 18 274-284.

Bartlett, F. C. Remembering. London: Cambridge University Press, 1932.

Baylor, G. W. A Treatise on the Mind's Eye: An empirical investigation of visual mental imagery. Doctoral dissertation, Carnegie-Mellon University, Ann Arbor, Michigan. University Microfilms, 1972. No. 72-12 699.

Baylor, G. W. A treatise on the mind's eye: An empirical investigation of visual mental imagery. Unpublished doctoral dissertation. Carnegie-Mellon University, 1971.

Berry, W. The flight of colonies in the after image of a bright light. Psychological Bulletin, 1922, 19 307-77.

Berry, W. Color sequences in the after-image of white light. American Journal of Psychology, 1927, 38 548-96.

Betts, G. H. The Distribution and Function of Mental Imagery. New York: Columbia University Teacher's College, 1909.

Bower, Gordon H. Mental imagery and associative learning, in Lee W. Gregg, Cognition in Learning and Memory. New York: John Wiley & Sons Inc., 1972.

Bower, G. H. Narrative stories as mediators for serial learning. Psychonomic Science, 1969, 14 181-82.

Brooks, L. R. The suppression of visualization by reading. Quarterly Journal of Experimental Psychology, 1967, 19 289-99.

Brooks, L. R. Spatial and verbal components of the act of recall. Canadian Journal of Psychology, 1968, 22 349-68.

Barakat, M. K. A factorial study of math ability. British Journal of Psychology, 1951, #4 137-56.

Biddle, W. E. Image therapy. American Journal Psychiat., 1969, 126 408-11.

Carey, N. Factors in the mental processes of school children. I. Visual and Auditory Imagery. British Journal of Psychology, 1915, 7 453-90.

Chase, William G., and Clark, Herbert H. Mental operations in the comparison of sentences and pictures, in Gregg, Lee, ed. Cognition in Learning and Memory. New York: John Wiley & Sons Inc., 1972.

Chase, William G. ed. Visual Information Processing. New York: Academic Press Inc., 1973.

Chomsky, N. Aspects of the Theory of Syntax. Cambridge: MIT Press, 1965.

Cooper, Lynn and Shepard, Roger N. Chronometic studies of the rotation of mental images, in Chase, William G., ed. Visual Information Processing. New York: Academic Press Inc., 1973.

Cooper, Lynn. Mental rotation of random 2-dimensional shapes. Cognitive Psychology, 1975, 7 20-43.

Deutsch, J. A. A theory of shape recognition, in Uhr, Leonard, Pattern Recognition. New York: John Wiley & Sons, Inc., 1966.

Dodwell, P. C. Coding and learning in shape discrimination, in Uhr, Leonard, Pattern Recognition. New York: John Wiley & Sons, Inc., 1966.

Feinbloom, W. A quantitative study of the visual afterimage. Archives of Psychology, 1936, 33 (No. 233).

Fernald, M. R. The diagnosis of mental imagery, Psychological Monographs, 1912, 14 (No. 58).

Galton, F. Statistics of Mental Imagery. Mind, 1880, 5 300-18.

Gibson, A. Feature Analysis of visual stimuli. Journal of Experimental Psychology, 1963, #1, 233-58.

Gould, John. "Eye movements during visual search and memory scan," in The Journal of Experimental Psychology. 1973, Vol. 98, No. 1, 184-195.

Gregg, Lee W. Cognition in Learning and Memory. New York: John Wiley & Sons, Inc., 1972.

Haber, Ralph Norman, and Herehenson, Maurice. The Psychology of Visual Perception. New York: Holt Rinehart and Winston Inc., 1973.

Hebb, D. O. The Organization of Behavior. New York, 1949.

Hill, D. S. An experiment with an automatic mnemonics system. Psychological Bulletin, 1918, 15 99-103.

Hochberg, Julian E. Perception. Englewood Cliffs: Prentice-Hall Inc., 1964.

Hochberg, J. "In the mind's eye," in R. N. Haber, ed. Contemporary Theory and Research in Visual Perception. New York: Holt Rinehart and Winston, 1968.

Holt, Robert R. "On the nature and generality of mental imagery," in Sheehan, Peter W. The Function and Nature of Imagery. New York: Academic Press, 1972.

Hull, Clark L. Goal attraction and directing ideas conceived as habit phenomena. Psychological Review, 1931, 38 487-506.

Huttenlocher, J. Constructing spatial images: A strategy in reasoning, Psychological Review, 1968, 75 550-560.

Jaenech, E. R. Eidetic Imagery. London: Kegan Paul, 1930.

Kennedy, John M. A Psychology of Picture Perception. London: Jossy-Bass Publishers, 1976.

Lyons, John. Introduction to Theoretical Linguistics. Cambridge: Cambridge University Press, 1968.

Martin, C. J., Cox, D. L., and Boerma, F. J. The role of associative strategies in the acquisition of P-A material. Psychonomic Science, 1963, 3 463-64.

Miller, G. A. The magical number seven, plus or minus two: Some limits on our capacity for processing information. Psychological Review, 1956, 63 81-87.

Miller, George A. The magical no. 7 ± 2. Psychological Review, 1953, 63 81-97.

Mower, O. H. Learning Theory and the Symbolic Processes. New York: Wiley, 1960.

Myers, J. L. Fundamentals of Experimental Design. Boston: Allyn & Bacon, 1966.

Neisser, Ulric. "Changing conceptions of imagery," in Sheehan, Peter W., The Function and Nature of Imagery. New York: Academic Press, 1972.

Newell, A., and Simon, H. Information processing analysis of perceptual processes in problem solving. Psychological Review, 1969, 76 473-83.

Newell, A. "On the analysis of human problem solving protocals," in Gregg, L. ed., Cognition in Learning and Memory. New York: Wiley & Sons, 1972.

Norman, Donald A. "Memory, knowledge, and the answering of questions," in Solso, Robert C. Contemporary Issues in Cognitive Psychology: The Loyola Symposium. New York: John Wiley and Sons, 1973.

Norman, Donald A. Memory and Attention. New York: Wiley and Sons, 1969.

Osgood, C. E. The similarity paradox in human learning. Psychological Review, 1953, 56 132-143.

Osgood, C. E., Suci, G. J., and Tannenbaum, P. H. The Measurement of Meaning. Urbana, Ill.: University of Illinois Press, 1961.

Oswald, I. Sleeping and Walking. New York: Elsevier Inc., 1962.

Paivio, A. Imagery and Verbal Processes. New York: Holt Rinehart and Winston, 1971.

Paivio, A., and Yuille, J. C. Word abstractness and meaningfulness, and paired-associate learning in children. Journal of Experimental Child Psychology, 1966, 4 81-89.

Paivio, A., and Yuille, J. C. Mediation instructions and word attributes in paired-associate learning. Psychonomic Science, 1967, 8 65-66.

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

Paivio, A., and Yuille, J. C. Changes in associative strategies and paired-associate learning over trials as a function of word imagery and type of learning set. Journal of Experimental Psychology, 1969, 79 458-63.

Paivio, A., Yuille, J. C., and Rogers, T. B. Noun imagery and meaningfulness in free and serial recall. Journal of Experimental Psychology, 1969, 79 509-514.

Paivio, A., Yuille, J. C., and Smythe, P. C. Stimulus and response abstractness, imagery, and meaningfulness, and reported mediators in paired-associate learning. Canadian Journal of Psychology, 1966, 20 362-77.

Paivio, A. Meaning, mediation, and memory. Paper presented at the meeting of the Canadian Psychological Association, Ottawa, May 1967. (a).

Paivio, A. Paired-associate learning and free recall of nouns as a function of concreteness, specificity, imagery, and meaningfulness. Psychological Reports, 1967, 20 239-43. (b)

Paivio, A. Effects of imagery instructions and concreteness of memory pegs in an economic system. Proceedings of the 76th Annual Convention of the American Psychological Association, 1968, 3, 77-78. (a).

Paivio, A. A factor-analysis study of word attributes and verbal learning. Journal of Verbal Learning and Verbal Behavior, 1968, 7, 41-49. (b).

Paivio, A. Mental imagery in associative learning and memory. Psychological Review, 1969, 76, 241-63.

Paivio, A., and Csapo, K. Concrete image and verbal memory codes. Journal of Experimental Psychology, 1969, 90, 279-85.

Paivio, A., and Madigan, S. A. Imagery and association value in paired-associate learning. Journal of Experimental Psychology, 1968, 76, 35-39.

Paivio, A., and Olver, M. Denotative-generality, imagery and meaningfulness in paired-associate learning of nouns. Psychonomic Science, 1964, 1, 183-84.

Paivio, A., Rogers, T. B., and Smythe, P. C. Why are pictures easier to recall than words? Psychonomic Science, 1968, 11, 137-38.

Paivio, A., and Rowe, E. J. Noun imagery, frequency, and meaningfulness in verbal discrimination. Journal of Experimental Psychology, 1970, in press.

Paivio, A., and Simpson, H. M. The effect of word abstractness and pleasantness on pupil size during an imagery task. Psychonomic Science, 1966, 5, 53-56.

Paivio, A., Smythe, P. C., and Yuille, J. C. Imagery versus meaningfulness of nouns in paired-associate learning. Canadian Journal of Psychology, 1968, 22, 427-41.

Paivio, A., and Yarmey, A. D. Abstractness of the common element in mediated learning. Psychonomic Science, 1965, 2, 231-32.

Paivio, A., and Yarmey, A. D. Pictures versus words as stimuli and responses in paired-associate learning. Psychonomic Science, 1966, 5, 235-36.

Paivio, A. Learning of adjective-noun paired associates as a function of adjective-noun word order and noun abstractness. Canadian Journal of Psychology, 1963, 17, 370-79.

Paivio, A. Abstractness, imagery, and meaningfulness in paired-associate learning. Journal of Verbal Learning and Verbal Behavior, 1965, 4, 32-38.

Paivio, A. Latency of verbal associations and imagery to noun stimuli as a function of abstractness and generality. Canadian Journal of Psychology, 1966, 20, 378-87.

Palyshyn, Zenor W. What the mind's eye tells the mind's brain, Psychological Bulletin, 1973, 80, 1-23.

Pear, T. R. The place of mental imagery in mental processes. Bulletin John Rylands Library, 1937, 21, 193-214.

Pelton, L. H. Mediational construction vs. mediational perception in paired-associate learning. Psychonomic Science, 1969, 17, 220-21.

Persensky, J. J., and Senter, R. J. An experimental investigation of an economic system in recall. Psychological Record, 1969, 19, 491-99.

Persensky, J. J., and Senter, R. J. The effect of subject's conforming to economic instructions. Journal of Psychology, 1970, 74, 15-20. (a).

Persensky, J. J., and Senter, R. J. An investigation of "bizarre" imagery as an economic device. Psychological Record, 1970, 20, 145-50. (b).

Peterson, L. R., and Peterson, M. J. Short-term retention of individual verbal items. Journal of Experimental Psychology, 1959, 58, 193-98.

Piaget, J. Play, Dreams and Imitation. New York: Morton, 1952.

Posner, M. I., Boies, S. J., Eichelman, W. H., and Taylor, R. L. Retention of visual and name codes of single letters, Journal of Experimental Monographs, 1969, 79, No. 1, 1-16.

Raser, C. A., and Bartz, W. H. Imagery and paired-associate recognition. Psychonomic Science, 1968, 12, 385-86.

Richardson, Alan. Mental Imagery. London: Routledge and Kegan Paul, 1969.

Richardson, Alan. "Voluntary control of the memory image," in Sheehan, Peter W. The Function and Nature of Imagery. New York: Academic Press, 1972.

Rimm, D. C., Alexander, R. A., and Eiles, R. R. Effects of different mediational instructions and sex of subject on paired-associate learning of concrete nouns. Psychological Reports, 1969, 25, 935-40.

Segal, S. J. "Imaging and perceiving in two sensory modalities," in Gregg, L., ed., Cognition in Learning and Memory. London: Wiley & Sons, 1972.

Sheehan, Peter W. The Function and Nature of Imagery. New York: Academic Press, 1972.

Shepard, R. N. The analysis of proximities: Multidimensional scaling with an unknown distance function. I. Psychometrika, 1962, 1, 125-40.

Shepard, R. N. Attention and the metric structure of stimulus space. Journal of Mathematical Psychology, 1964, 1, 54-87.

Shepard, R. N. Recognition memory for words, sentences, and pictures. Journal of Verbal Learning and Verbal Behavior, 1967, 6, 156-63.

Shepard, R. N., and Chang, J. J. Forced-choice tests of recognition memory under steady-state conditions. Journal of Verbal Learning and Verbal Behavior, 1963, 2, 93-101.

Shepard, R. N., and Metzler, J. Mental rotation of three-dimension objects. Science, 1971, 171, 701-03.

Shepard, R. N., and Chipman, S. Second-order isomorphism of internal representations: Shapes of states. Cognitive Psychology, 1970, 1, 1-17.

Shepard, R. N., and Feng, C. A chronometric study of mental paper folding. Cognitive Psychology, 1972, 3, 228-43.

Shiffrin, Richard M., and Geisler, Wilson S. "Visual recognition in a theory of information processing," in Solso, Robert C., Contemporary Issues in Cognitive Psychology: The Loyola Symposium. New York: John Wiley and Sons, 1973.

Short, P. L. The objective study of mental imagery. British Journal of Psychology, 1953, 44, 38-51.

Simon, Hebert. How big is a "chunk"? Science, No. 183, 1974, 482-88.

Simon, H. A., and Feigenbaum, E. A. An information-processing theory of some effects of similarity, familiarization, and meaningfulness in verbal learning. Journal of Verbal Learning and Verbal Behavior, 1964, 3, 385-96.

Simon, H. A., and Barenfeld, M. Information-processing analysis of perpetual processes in problem solving. Psychological Review, 1969, 76, 473-83.

Simon, H. A. "What is Visual Imagery?" An information Processing Interpretation, in Gregg, Lee W. Cognition in Learning and Memory. New York: John Wiley & Sons Inc., 1972.

Singer, J., and Antrobus, J. S. Eye movements during fantasies. Archives of General Psychiatry, 1965, 12, 71-76.

Skinner, B. F. Science and Human Behavior. New York: MacMillan, 1953.

Solso, Robert C. Contemporary Issues in Cognitive Psychology: The Loyola Symposium. New York: John Wiley and Sons, 1973.

Sperling, G. The information available in brief visual presentations. Psychological Monograph, 1960, 74, (11, Whole No. 498).

Sperling, G. A model for visual memory tasks. Journal of the Human Factors Society, 1963, 5, 19-31.

Standing, Lionel. Learning 10,000 pictures. Psychonemtric Science, No. 19, 1970, 73-74.

Sternberg, Saul. High speed scanning in human memory. Science, No. 153, 1966, 652-54.

Sternberg, S. Memory scanning: Mental processes revealed by reaction time experiments. American Scientist, 1969, 57, 421-57. (b).

Teasdale, H. H. A quantitative study of eidetic imagery. British Journal of Education Psychology, 1934, 4, 56-74.

Titchener, E. B. A Textbook of Psychology. New York: MacMillan and Co., 1919.

Tolman, E. C. Purposing Behavior in Animals and Men. New York: Century, 1932.

Uhr, Leonard. Pattern Recognition. New York: John Wiley & Sons, Inc., 1966.

Watson, J. B. Psychology as the behaviorist views it. Psychological Review, 1913, 20, 158-77.

Watson, J. B. Behaviorism. Chicago: University of Chicago Press, 1930.

Wohwill, Joachin. "Developmental studies in perception," in Sheehan, Peter W. The Function and Nature of Imagery. New York: Academic Press, 1972.

Wundt, G., and Bolt, M. Mediation and mediation time in paired associate learning. Journal of Experimental Psychology, 1968, 78, 15-20.

Appendix 1

Post-experiment questionnaire
(introspective comments upon
mental rotation task)

APPENDIX I

THE SUBJECT QUESTIONAIRE

Question 1

What did you think of the task you have just completed? Did you find it easy or difficult? Please explain.

Answers

It wasn't easy at first to think, but I didn't find it difficult either. It was a challenge. All I had to do was to create a good enough image of the first picture so that I could compare it with the other two. At first I had a problem getting my fingers coordinated and pushed the wrong button, but after the first trial, I pushed the button on the side I wanted except for a couple of clumsy errors.

That's interesting. I find it very easy because there are actually three shapes to remember and they are repeated and repeated to appear to your eyes, but you need to concentrate very much and remember the shapes.

It was easy. It was a matter of comparing the orientation, configuration of the original object, "image," with the second set.

Easy, it got easier each time. The only difficulty was the angles. I had great difficulty relating them.

Most of it was quite easy, but it was tiring to repeat 36 cards five times. It was frustrating to get an error on something that seemed so easy.

I found the task somehow difficult with some of the sets . . . in the first part (out of the 5) then after the second set I found it pretty easy, particularly for parts 4 and 5. Once I learned the tricks by which to match the stimuli it was pretty easy. Namely, the outlined and the totally filled in "hand" stimuli. I didn't get tired of it or bored with it. I didn't find it hard to concentrate but I found it easy not to concentrate (i.e., think of the previous set while already looking at the next one), and therefore, in 1 or 2 sets I ended up guessing.

Easy. I visualize the glove pattern, then turn it until it fits either the left or right bottom image. With the dot pattern I try to remember the sequence, if there are three, two out or whatever. Later, the dot pattern moves in my mind and I get more to the right. The lines I don't turn, I just try to remember what order they appear in (e.g., thick, thin, thick).

Question 2

Was the initial stimulus displayed in the top part of the screen long enough?

Answers

For me it was fine for the solid shapes but too short for the dots. However, once I got used to the experiment and asked what to do, I had little problem.

Yes, that's long enough. Once you recognize it, you would get it straight into your head.

Yes, probably even longer--slightly.

Yes, for the most part.

Sometimes. Initially I thought there wasn't enough time for the stimuli described in Question 4, especially the dots (v's) weren't shown long enough but after the first set was over it became easier (i.e., it seemed easier with each successive set). With some of the easier stimuli I found that once familiar with them I would just glance at them and shift my eyes to the lower half of the screen before the two seconds were up (i.e., when the lower screen was still black).

Yes. In fact, after I had done the test a few times, I found myself waiting for the lower half of the screen to appear.

Yes, I look at it, then at the bottom of the screen so it is positioned properly off the turn.

Question 3

Can you explain the process by which you were able to match the related stimuli in the lower half of the screen with the target stimulus displayed in the top half?

Answers

I memorized the shape of the target stimuli, then when the related stimuli came on the screen, I rotated them in my mind's eye until they fit or matched the outline of the target stimuli which was upright.

When the shape appears on the screen, remember at once its shape, concentrating in your mind. When the rotated stimuli in the low half appears, you seem to rotate the target stimuli displayed in the top half in your mind in order to match the shape you want.

By orienting or overlapping the two.

By turning top image sideways or upside down to match bottom image.

With the hand stimuli that had nonsymmetrical top peak I would concentrate on the unsimilarity, i.e., which side was shorter; steeper and almost ignore the rest of the stimulus. With the ones that had the top peak symmetrical I concentrated on the positions of the "thumb" (i.e., whether it was clockwise or counter-clockwise) for the hand in the circle not completely filled in, I concentrated on one part only (namely a thin line across the centre). The first time I saw this stimulus I had, of course, no idea what the lower stimuli was going to look like but then (two or three tries later) I noticed that the corresponding line on the incorrect stimulus was much thicker and therefore, for all the consecutive sets I concentrated on this line only, paying no attention to the rest of the stimulus or its clockwise position.

With some of the figures, it was just a matter of looking on which side the arrow appeared. Then, when the lower screen appeared, I would rotate the two figures, or just one of them, until I could tell which one had the arrow on the same side. Sometimes I rotated the figure in the top half (in my mind) and tried to match its rotation with one of the lower figures.

I would take the part of the target stimulus and focus on it, then turn it until it fits one of the related objects on the bottom.

Question 4

Which of the stimuli appearing in the upper and lower half of the screen were hardest to match against the original target stimuli?

Answers

The dots were much more difficult. When I tried to rotate them, they sort of drifted out of their positions, making it very difficult. So I had to memorize their relative spatial relationships to each other.

I don't quite understand your question. Anyway, I would guess an answer for this. I think it is easier to match the target stimulus with the rotated stimuli because you remember right away the shape of the initial stimulus and you get such an image in your mind. But when you see the rotated stimuli, you have got to choose.

One with the dots.

Disjointed images were harder to match. Also harder to match with an upside down image.

The ones in the circle
1. the dots (v's) arranged in a circle
2. geometrics (triangles etc.-- in black) in the circles

The progress of 2) explained in answer to question 3. 1) it took me longer to recognize the arrangement of the V's and the clockwise position always made a difference in contrast to 2) I usually had to imagine the stimulus in the upper half and rotate it until it matched one of the ones in the bottom.

The broken figure with dots was the hardest to match. When I rotated that figure in my mind as it appeared in the top half, I was unable to see any corresponding difference or rotation in the bottom half. Both figures in the bottom seemed to be the same. I tried filling in the lines between the dots as well as counting the dots.

The wing-sequence was hardest to match. They seemed to have gone to totally different positions.

Question 5

Was any form of mental imagery used in this task? If so, how?

Answers

Yes, Once I formed the mental image of the target image I retained that picture of its shape in my mind. Then when the two bottom images appeared, I just took each one in turn, envisioned it in my mind, and then turned it until it was upright to compare with the target. When it matched, I pushed the button; when it didn't match, I did the other image. The dots took much longer because I had to use their spatial relationships to each other and moving these around without disturbing the dots (in my mind screen) was difficult and had to be done slowly.

Yes. When you try to remember the shape, you get such an image in your mind. When the rotated stimuli in the lower half appear, you would be able to recognize the shape immediately.

Yes. By storing the image of the original figure, then later comparing with the bottom figures.

For each image, I looked for one certain point and then matched that point to the bottom image. All that was necessary to do was to turn the image to match the top image.

I would imagine the upper stimulus (namely the V's in the circle and some of the "hands") and rotate it to match the bottom stimuli. I found that when the bottom stimuli were at "6 o'clock" (even between 6 and 9 o'clock) it took me longer to match them up than when they were between 12 and 6 or 9 and 12 (i.e., hardest when upside down).
Points of interest I found:
1. when I was told my answer was wrong, I tended to take more time with the next set in order not to make a mistake again.
2. sometimes I matched the stimuli but it took me a while to push the button (longer than at other times).
3. sometimes (maybe 5 to 7) I pushed the wrong button (meaning to push the other one) and sometimes I would push a button, and immediately realize I was wrong. Other times I had no idea I was wrong until I was told.
4. I found the stimuli one and two (of answer to question 4) the ones of the most interest.

Yes. I imagined what the top figure looked like. Because I could look at it for only 2 seconds, I tried to retain in my mind what it looked like so that I could refer to the picture in my mind when the lower half of the screen appeared. I suppose the act of rotating figures in one's mind is a form of mental imagery.