



Exploring the role of components in object recognition.

Item Type	Thesis (Open Access)
Authors	Santhi, Nayantara
DOI	10.7275/7675975
Download date	2025-05-24 04:22:52
Link to Item	https://hdl.handle.net/20.500.14394/45824

UMASS/AMHERST



312066011260100

EXPLORING THE ROLE OF COMPONENTS IN OBJECT RECOGNITION

A Thesis Presented

by

NAYANTARA SANTHI

Submitted to the Graduate School of the
University of Massachusetts Amherst in partial fulfillment
of the requirements for the degree of

MASTER OF SCIENCE

September 1994
Department of Psychology

EXPLORING THE ROLE OF COMPONENTS IN OBJECT RECOGNITION

A Thesis Presented

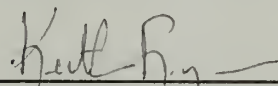
by

NAYANTARA SANTHI

Approved as to style and content by:



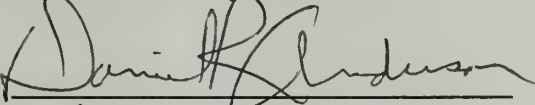
Alexander Pollatsek, Chair



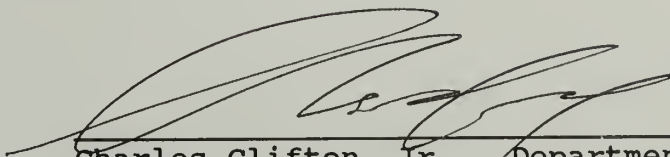
Keith Rayner, Member



Arnold Well, Member



Daniel Anderson, Member



Charles Clifton, Jr., Department Head
Department of Psychology

TABLE OF CONTENTS

	<u>Page</u>
LIST OF TABLES.....	iv
LIST OF FIGURES.....	v
Chapter	
1. BACKGROUND.....	1
Componential Approach.....	2
Theories of Object Recognition.....	5
Componential Representations.....	11
Hypothesis and Predictions.....	21
2. INTRODUCTION TO THE STUDY.....	25
Hypothesis.....	29
Experimental Design.....	32
Predictions.....	33
3. EXPERIMENT 1.....	36
Method.....	36
Results and Discussion.....	39
4. EXPERIMENT 2.....	49
Method.....	50
Results and Discussion.....	52
5. GENERAL DISCUSSION.....	61
The Parsing and Integration Principle.....	62
Types of Componential Representation.....	64
Time Course of Processing.....	72
Conclusions.....	73
BIBLIOGRAPHY.....	84

LIST OF TABLES

Table		Page
1	Response Times by SOA and Presentation Condition in Experiment 1.....	75
2	Error Rates by SOA and Presentation Condition in Experiment 1.....	76
3	Response Times by SOA and Presentation Condition in Experiment 2.....	77
4	Error Rates by SOA and Presentation Condition in Experiment 2.....	78
5	Response Times by Object/Non-Object and Presentation Condition in Experiment 2.....	79
6	Error Rates by Object/Non-Object and Presentation Condition in Experiment 2.....	80

LIST OF FIGURES

Figure		Page
1	Parts Only in Correct versus Incorrect Location.....	81
2	Parts + Interface in Correct versus Incorrect Location.....	82
3	Pieces + Interface in Correct versus Incorrect Location.....	83

CHAPTER 1

BACKGROUND

People demonstrate an amazing capacity to process and identify objects quickly and accurately despite a continually changing visual environment. Three of the biggest challenges in vision research is centered on understanding exactly what type of information is extracted from an object, understanding the consistencies and/or variations in any specific type of information that is extracted from the object under different visual conditions and understanding the representation of this information in memory. These three factors are critical in accounting for the fact that we can perceive a 3 dimensional object although the retinal image is itself only 2 dimensional as well as the visual system's ability to successfully cope with the variability in the visual environment.

There are several theoretical approaches, each with different assumptions about the kind of information that are extracted from an object during the identification process. The assumption in this study is that object identification is a componential based process where information about the constituent parts of an object is very important in its identification. Thus the primary goal of this study was to explore the role of componential information in object identification and the consistencies and variations in the componential information that are extracted (from an object) under different visual conditions.

A critical implication of the componential idea is that the process of identification requires the parsing of information from the retinal image into components followed by the integration of this componential information. If this is true then in a situation where parts of an object are presented in quick succession, people should be able to integrate the information from the parts and identify the object. Since performance in such a task can be used to assess the role of componential information in the identification process, the basic paradigm in this study involved a sequential presentation of componential information from objects that had to be identified.

Componential Approach

According to the componential approach, information from an object is considered to be separable into different levels, which is reflected in the memory representational system as a hierarchical arrangement of increasing complexity. Thus for example within the theoretical framework of this study several levels like the featural level, the componential level, semantic level and conceptual level can be distinguished. The focus of this study, however, is the information that constitutes the componential level which is considered to be an intermediate level in the object representational system.

Components or parts can be thought of as the meaningful units of information about an object. Several researchers

(Rosch, Mervis, Gray, Johnson, and Boyes-Braem, 1976; Tversky and Hemenway, 1984) have actually defined parts of an object as units that have both a structural and functional aspect. Examples of components are the handle of a cup, wheel of a car, beak of a bird, etc.

One possible description of an object is in terms of its parts and the relationships (spatial and other types) between them. In order to thus describe an object, two important units of information are necessary. They are the whole meaningful components or parts of the object and the relationship (spatial and other types) between the components or parts. Thus, for example, a componential description of a cup would include information about the handle, the body and the join between the handle and the body.

A related issue here is the distinction between the information that is available in the retinal image and its analogous representation in memory. This distinction is a reflection of the kind of mapping that exists between the information that is extracted from the retinal image and the memory representation. For example, it is possible that each type of information that is extracted from the retinal image has a corresponding memory representation or that several types of information share a common representation. Assumptions about the mapping that might exist between the information represented in the retinal image and the information represented in memory is crucial in explaining the

type of transformations that are needed to map a 2-D image to a 3-D percept and in accounting for the fact that the visual system exhibits remarkable competence in adapting to the variability in the visual environment.

We are capable of recognizing objects despite the fact that the retinal image is viewer centered and therefore affected by variations in the visual environment. This suggests that the visual system is sensitive to and capable of extracting many different types of information from an object. Consider for example the spatial relationship between the parts of an object. It can have one of two possible descriptions. These descriptions could be based on the spatiotopic coordinates of the contour around the interface, or they could be at a more abstract level where the relationship is specified in terms of whether one part is to the left or top or front of the other. In each of these cases, a different type of information can be used to describe one aspect of the object. Furthermore, the strength or salience of each of the two types of information would depend upon the specific situation. Thus this is a reflection of how variations in a specific type of information can help the system to deal with the variability in the visual world.

Theories of Object Recognition

While most theoretical approaches to object identification propose some kind of information parsing and

representation of this parsed information, the template approach posits that the memory representation of an object is unitary. The template approach will be explored as an alternative to the assumptions of the componential approach.

RBC Theory

One theory where the central focus is the idea of componential representation is Biederman's 'Recognition by Components' theory (Biederman 1987). According to this theory an object is identified by decomposing its image into its constituent components and their spatial relationships. Implicit in this idea is the assumption that the important units of information from an object are organized as separate levels of representation in memory. The theory assumes that the lowest meaningful level of representation consists of the primitives or the fundamental components of an object. These are referred to as 'geons' in the model. A mathematical conceptualization of the representation of a geon is that it is the volume swept out by a cross section moving along an axis.

The theory postulates that in order to recover the geons from the retinal image, the visual system initially has to extract information about the edges of the object. Therefore, early in processing there is an edge extraction stage where the visual system is responsive to the surface characteristics (luminance, color, hue etc) of an object that would determine

an edge. Edge perception is a very crucial stage in the model, because certain fundamental properties that define the primitive components (geons) of any object are derived from the perception of edges. These properties are assumed to be invariant over transformations and hence non accidental in nature. The model proposes five non accidental properties: (1) Collinearity; (2) Curvilinerity; (3) Parallelism; (4) Symmetry; and (5) Cotermination. The recovery of these non accidental properties presumably leads into the next stage where the components of the object are identified. This then proceeds into the stage where the components are matched with the appropriate representation in memory and the object is identified.

A core feature of this model is that the representational system consists of a small finite number of fundamental components that can describe most of the known objects. The representational power of this theory rests on the assumption that a large number of combinations can be derived from this limited set of components. Thus the system may only need a very small number of primitives to describe the known set of objects. As an example consider the description of a pail and a cup. Both of them share the same geons but in a different spatial arrangement.

The idea of a componential analysis in object identification has some supporting experimental evidence. Biederman et al (1987) reported a series of experiments which

supported the principle of componential recovery in object recognition. Biederman, Ju and Clapper (1985) investigated the identification of objects with some of their components missing. They hypothesized that 2 or 3 components are sufficient to identify an object. Therefore they predicted that if only a couple of geons are required to identify an object then there should be no significant difference in performance between the conditions where the complete or partial object was visible. In addition, they also predicted that if a minimum of 2 or 3 geons is sufficient for identification then a complex object (by virtue of providing more componential combinations) would be identified faster than a simple object. The findings supported the above mentioned predictions.

As discussed earlier in the 'RBC' theory, edge extraction is postulated to be a very crucial factor in processing. Biederman et al (1987) tested this assumption by comparing the response times for identifying line drawings with response times for recognizing colored photographs. They predicted that if edges are really important then there should be no difference between the times to identify a line drawing and a colored photograph because the information about edges is the same in both cases. The results showed that there was no significant difference in performance for the two types of stimuli.

The other important aspect of the model is the notion of representation of information about the spatial relationships between the components. In the model, spatial relations between components are assumed to play a very important role in the process of identification. If spatial relations are indeed vital to object recognition, it becomes necessary to understand how they are represented in memory. Although the RBC theory stresses the importance of the perception of spatial relationships among components, it does not clearly specify how this information is represented in memory. There are conceivably several ways in which spatial relations between components can be described. In a recent paper Hummel and Biederman (1991) describe a neural network implementation of the RBC theory where the spatial relationships between components are represented as propositional nodes with relations like "top", "bottom", "left", "right" etc being represented by each node.

The Template Theory

One alternative to a componential representational system is a template system where the representations consist of whole objects. Identification of an object in such a framework would involve finding the best match between the retinal image and a template in memory. In order to be fairly successful, a template based model would need to develop an account of how such a system would deal with the varying

conditions under which objects are encountered. For example, the model would have to explain how despite transformations in the visual space, an object is always identified fairly quickly and accurately. Furthermore, we are also fairly good at identifying different exemplars of objects, and sometimes even unusual objects.

One possible explanation is to assume memory representations or templates for all possible instances in which an object has been encountered. In addition, this line of reasoning would also require postulating templates for the different exemplars (types) of the same object, such as different types of cats or chairs. This could cause the model to become very limited in efficiency because of the combinatorial explosion that would result. More importantly, a multiple template assumption would discount the hypothesis of systematic differences in identification due different types of partial information that vary in the recoverability of components. This is assuming that all featural information in the templates is weighted equally. Thus such an approach will not be able to account for the data in literature that demonstrates such differences.

One reason that a whole template approach would not predict a superiority for an object component over any portion of an object (assuming both segments are equated in terms of total available information) is because the identification process involves finding the correct match between the stored

template of the object and the stimulus. Therefore, differences in identification would depend more on the amount of available information and not necessarily the type of information. Thus if we have a whole object template then there should be no differences between conditions where the whole parts of an object are presented in quick succession and a condition where just any arbitrary segments of the object are presented because all the information needed is available in both cases. This point can be further illustrated by considering situations that involve sequential presentation of parts both in the correct and incorrect locations.

Since there are no discrete representations of componential information in the template model, information about the spatial relations between components do not play an important role in processing especially in situation that require specific use of this information. For example in a situation where the parts of an object are presented in an incorrect spatial location (see Figure 2) any kind of abstract spatial relational information might not be extracted to resolve the ambiguity.

Recently however, there have been attempts to model visual processing using a revised template approach that overcomes some of the above problems. One such attempt is Arnold Trehub's (1991) neural model of shape recognition. It is a modified template model with parallel processing, parsing and componential representation. This model is better able to

account for at least some of the data involving the role of part information on the identification of an object. However an important difference between this model and the componential approach presented in the study is that in the Trehub's model componential representations are arbitrary parts of objects and not necessarily the actual meaningful parts.

Componential Representations

As mentioned earlier the basic theoretical assumption in this study is that object recognition is a componential based process where recognition involves the parsing of the retinal image into components. Furthermore, several types of componential information will be differentiated, based on a concrete (visual) or abstract distinction.

The parts of a cup could be represented as the handle and the cylindrical body. In addition, the parts can be conceptualized as a 3-D model of the two parts or as a more conceptual representation of each of the parts that has both abstract and visual features. Taking into account the fact that there is a lot of redundancy in the visual representational system, it is plausible that we have both kinds of representations for parts of objects.

There is some evidence in the literature for the importance of information about components in object identification. The logic of some of these studies is that if identification of an object is assumed to proceed via the

recovery of its components, then even partial information that allows for the recoverability of the parts, should afford successful identification. Therefore, presenting different portions or segments of an object that vary in the recoverability of components should systematically affect the efficacy of identification thereby demonstrating the role of componential information.

Biederman et al (1985) conducted experiments using two kinds of contours, recoverable and non recoverable. When contours are deleted at points of concavity such that when extended they bridge the concavity, the recovery of components is prevented. Such contours are called non recoverable contours. On the other hand, deleted contours (that have intact regions of concavity) which allow the components to be recovered are called recoverable contours. In other words, regions of concavity help the system to determine if a join between two components exists at that point or conversely points of concavity could be used to parse an object into its various parts. Consequently certain contours become critical for object recognition.

Several experiments using deleted contours have shown that under conditions where contextual inference is not possible, certain contours are essential for object identification. Contours of this type could be considered partial information. The data from the experiments by Biederman et al (1985) showed that the error rate

inidentifying non-recoverable objects was almost 100% while this was not true for the recoverable contours.

Apart from information about the parts themselves, another important type of componential information is the spatial relationship between the parts of an object. Spatial relations can be differentiated into visual and abstract relations with different levels of specificity of information included in each representation. For example, the relationship between the handle and the body of a cup could have one of the two possible abstract descriptions: the handle is attached to the side of the cup or the handle is attached in the middle to the left side of the cup. In this example there are clearly two separate levels of specificity of information in the two descriptions. Furthermore, it is possible that relational information that is not necessarily spatial can also be used in integrating componential information under some conditions. Thus we might actually have several types of representations of the relations (spatial and other) between parts of an objects. Three sources of relational information are hypothesized to be available in the retinal image.

Relations Between Object Parts

The description of the relationship between two parts of an object is a complicated issue because there are conceivably many different ways in which relations can be

described. Furthermore, while there may be many sources of relational information in the retinal image of an object, there may or may not be a corresponding memory representation for each type. Thus the relationship between the relational information (between parts) in the retinal image and the representation in memory might be complex. Three sources of relational information between the parts of objects are hypothesized. While there is evidence in current literature for the role of relational information in object recognition, this study is an attempt to extend this idea by providing a concrete description of the different types of possible relational information. Furthermore, this study also aims to experimentally test this idea. Two of the proposed types of relational information directly involve the actual spatial relations between the parts while the third type is proposed to have both a structural and functional component and can be used in deriving the spatial relation. The three types of relational information are concrete geometric relational information, abstract spatial information and potential relation information. These will be discussed in turn below.

Concrete Geometric Relational Information. One of the key sources of information about the spatial relationship between two parts is the region of join between them. Thus the interface between two parts of an object could be considered concrete geometric relational information that is available in the retinal image. Since the retinal image is

always two dimensional, the actual interface might appear different depending on the orientation of the object. Therefore it is proposed that the information that is extracted from the interface is the object centered coordinates of the contour around this interface which is then matched to an analogous object centered 3-D representation in memory. Thus there is a one to one correspondence between the representation in the image and the memory representation for this type of spatial relational information.

The memory representation of the concrete geometric relation is conceptualized as an encoding of the coordinates of the contour around the interface in a spatiotopic map. While the information in the retinal image is viewer centered the memory representation of the concrete geometric relational information is considered to be object centered.

Abstract Spatial Relation Information. The abstract spatial relation is the relational information that specifies whether one part is to left or right, front or back or top or bottom of the other. In other words it is abstract relational information about the spatial configuration of two parts with respect to each other. In the retinal image, this would be the information about whether one part is to the right or left or top or front of the other. The information that is extracted from the retinal image is proposed to be propositional in nature. Furthermore, the memory representation is also assumed to be propositional. However,

there are two possible ways in which this information can be represented in memory. There could be a viewer-centered propositional representation of the abstract spatial relational information. For example, in the case of a cup, the memory representation would be the handle goes to the left or right or front or back of the cylindrical body. Depending on the orientation in which the object is encountered the abstract spatial relationship that is extracted from the retina can be matched to the appropriate representation in memory. Secondly, there could be an object-centered propositional memory representation of the abstract spatial relation. For example, the relationship between the parts of a cup would be 'the handle goes to the side of the cylindrical body' or in the case of a pail it would be 'the handle goes to the rim on the top'.

In the case of a partially occluded object (where one part is visible) the spatial location of the visible part along with the spatial location of the occluding surface can enable one to extract information about the spatial locations of the parts relative to each other.

For example, in Figure 2 the position of the handle and the occluding boundary implies that there is a part 'X' on top of the cylinder. Another instance of a situation where the abstract spatial relationship can be used is in the case of an object that is broken at the join between two parts.

The broken area can provide some information about the spatial configuration of two parts with respect to each other.

Potential Relation Information. The potential relational information is conceptualized to be a representation that is very general and includes all types of relational information (both structural and functional), that would enable the viewer to integrate the parts of an object. In the case of objects the set of potential relations is conceivably constrained by various sources of information that is available, such as the physical structure of the parts or knowledge about the parts themselves (functional knowledge) or whether the central axis of the parts are in a horizontal or vertical position which would determine whether the parts are in a horizontal or vertical configuration with other parts.

The physical structure of any part of an object offers some information about the set of potential spatial relations it could afford. For example, the outside surface of a hollow cylinder could potentially have other smaller parts attached to it. Thus the structural property (ie, the outside surface) of a cylinder can provide the system with the information about the potential relations it can afford. While this is the structural component, the functional component of a hollow cylinder would be the information that it could potentially be a container. Thus this function can provide the system with the information that the cylinder can

have a 'handle' relation with another part. This information can reduce the potential relations to a part on the outside surface of the cylinder. Based on all this information, the system can extract the potential spatial relation between two parts.

The potential relational information is always available under all conditions and is the only source of information under conditions where the concrete geometric and abstract spatial information are not available. Described below is an example of situations where the potential relational information is the only source of relational information that is available. For instance a dimly lit surrounding with moving objects (like animals) would be a condition where only the potential relational information would be available. In this situation the orientation of the object is continually changing due to constant movement making it difficult to extract any stable concrete geometric or even abstract spatial relational information. However the potential relation information that would still be available can enable the integration of the parts. Thus under extreme circumstances where the available componential information is unstable, the set of potential spatial relations seem to be an easier way of getting the information needed to integrate the parts.

As mentioned earlier, the concrete geometric relational information is hypothesized to have an object centered 3-D representation in memory. On the other hand two alternatives

are considered for the representation of the two kinds of abstract relational information. Each of the types of abstract relational information might have a separate representation in memory or they may share a common memory representation.

There is some evidence in the literature for different types of representations of spatial relationship information. Kosslyn et al (1991) reported a series of experiments using the divided visual field methodology to investigate two spatial representational systems. The main hypothesis of the Kosslyn et al study was that there are two spatial representational systems, an abstract or categorical one and a specific metric one. It should be noted that these spatial relationships have been typically discussed with respect to relations between objects rather than within an object nevertheless can be extended to within an object. An example of a categorical relationship between a cup and a table on which it has been placed would be the following. The cup is on top of the table. This is an example of a categorical relationship between the cup and a table. A metric relation for the two objects would be 'the cup is on top of the table about 5" from the center'. It is evident from this example that a categorical relation can be considered analogous to the abstract spatial relation that has been proposed in this thesis, while the metric relation can be considered analogous to the concrete geometric relation. Furthermore, the

investigators hypothesized that the primary processing area of the categorical system was in the left hemisphere and the metric information was initially processed in the right hemisphere. The results showed that subjects evaluated and encoded categorical spatial relations faster when the stimuli were initially presented to the left hemisphere while coordinate spatial relations were encoded faster when the stimuli were initially presented to the right hemisphere. Thus the data from their study seem to suggest that abstract spatial relational information is processed in the left hemisphere while the concrete spatial relational information is processed in the right hemisphere. The most interesting aspect of their study is the evidence for the processing of different types of spatial relational information.

Both the theoretical proposals and experimental evidence seem to suggest that information about the components of objects and the spatial relations between them play an important role in processing and furthermore also have corresponding representations in memory. Two experiments were conducted in this study to explore the assumptions of the componential theory of object recognition that have been discussed so far. One of the experiments in the study used a naming task and the other used an object decision task (Kroll and Potter 1984).

The basic paradigm involved presenting the componential information (parts and their spatial relationships) from an

object in quick succession. A trial consisted of the presentation of two or more segments, each of which was either the whole component or a partial component of an object and the interface between the components. In both experiments the primary dependent measure was the response time (the time required to name the object or decide if it was an object) and the accuracy of response. The details of the methodology are described in the following chapter.

Hypotheses and Predictions

Three hypotheses were tested. They are:-

(1) The representation (in the retinal image and in memory) of the components of objects consists of whole meaningful parts and is highly visual in nature. Furthermore, the information extracted from the retina is compared to an object centered 3-D representation in memory.

To test this hypothesis, two contrasting conditions were used where the segments consisted either of line drawings that had a complete contour (whole meaningful part: figure 1) or line drawings that had a small portion of the contour missing (arbitrary segment: figure 3). If components are represented as whole meaningful parts then a situation where the parts are presented in quick succession (figure 3) should enable better object recognition than a situation where just arbitrary segments of an object are shown in quick succession (figure 1 and 3).

An object template approach on the other hand would predict no differences between the experimental conditions (figure 1 and 3) since the two types of segments are equated in terms of total information and finding the closest match in either case should take the same amount of time. The only situations where the template theory would predict a significant difference is in cases where the segments differ in the total amount of available information.

(2) There are three different types of relational information available in the retinal image with analogous representations in memory. They are (a) Concrete geometric information (b) Abstract spatial information (c) Potential relational information.

In the two experiments, the relational information was afforded by the interface segment, the spatial positioning of the parts as they appeared on the screen, the position of the occluding boundary which acted as an indirect source of relational information and the parts themselves which are sources of the potential relational information.

Two experimental manipulations were designed to explore the three types of relational information. The first manipulation involved two contrasting situations. The parts were presented along with a segment that consisted of the interface between the two parts or the parts were presented without the interface segment. If the concrete geometric spatial relationship between parts of objects plays a crucial

role in the integration of componential information, then the interface segment condition should enable quicker integration than the non-interface segment because the interface segment will enable quicker extraction and access of the concrete geometric representation. In the latter case, this geometric information needs to be extracted from the positioning of the parts and will slow down the process. An example of interface information versus no interface information is Figure 2 versus Figure 1.

The second manipulation involved contrasting conditions. The parts are presented sequentially in the right spatial location or the parts are presented sequentially in an incorrect spatial location. The correct or incorrect location was based on the actual positioning of the parts with respect to each other as they appeared on the screen. An example of correct versus incorrect spatial relation information is Figure 3. In the condition where the parts are in the correct location it should be possible to extract some kind of concrete geometric information and abstract spatial information from the positioning of the parts and the occluding boundary while in the incorrect location only the abstract information is available from the position of each part relative to the occluding boundary.

If componential integration is differentially affected by the different types of spatial relational information, then responses in the condition with the correct location should be

more efficient than the condition with the incorrect location because in the former case both the concrete geometric and abstract spatial relational information is available while in the latter case only the abstract spatial relational information is available.

Once again, an object template approach would predict no differences since the total amount of available information is the same in all conditions.

(3) Componential information is accessed and used very early in processing.

Since componential analysis and integration occurs very early in processing it was felt that manipulating the timing might be one possible way of getting some evidence for this. Three different SOA conditions were used. They were 75 ms, 250 ms and 500 ms.

If componential analysis and integration occurs early in processing then effects of the manipulations designed to reflect componential processing should be seen only in the two short SOA conditions because in the 500 ms SOA presumably the system could use inferential processes to resolve the various ambiguities (e.g. the incorrect location).

The following section describes the details of the methodology used in both the experiments.

CHAPTER 2

INTRODUCTION TO THE STUDY

Recent trends in the theoretical and experimental research on object recognition show an emphasis on the importance of componential information in the identification process (Biederman 1987, Hummel and Biederman 1991). The basic idea underlying this viewpoint is that the early stages of object recognition involve the parsing of the retinal image into information about the components and the relationships (spatial and other) between them.

In the componential approach, information that is important in object identification is conceptualized as having a hierarchical arrangement of different levels of complexity with componential information constituting the intermediate level. Furthermore within each level there could be different types of information that can describe one or more attributes of an object. For example, the spatial relationship between two parts can be described in terms of concrete (visual) or abstract information. This is a critical assumption in this study, because it is the underlying basis of the hypothesis that is designed to explore the role of different types of componential information in the identification process.

The issue of what qualifies as componential information is debatable since there are conceivably several ways in which an object can be parsed into smaller units of information. In this study componential information is assumed to be

information about the whole meaningful parts (a part that has both a structural and functional component to it) of an object and the relationships (spatial and other) between them. This type of information can be considered to be distinct from other possible basic units of information like featural information which are also processed during identification.

The primary aim of this study was to explore the role of componential information in object recognition by testing some of the assumptions of the componential approach. There were four critical assumptions to this study: (1) Componential information refers to the information about the parts and the spatial relationships between them and furthermore they are vital to the identification process; (2) The componential approach implies a parsing and integration stage where it is the information about the spatial relations between the parts that are important to the integration of componential information; (3) The abstract/visual distinction can be used to describe different types of relational information between parts of an object; (4) There is a difference between the different types of componential information available in the retinal image and their representation in memory. For example the information in the retinal image is always 2 dimensional and viewer centered while this may not be true of the memory representations. They could be either 2 or 3 dimensional and viewer or object centered. Both alternatives are considered

in this study. In this study four different types of componential information have been explored. They are parts of objects and the three types of relationships that can exist between the parts. These will be discussed in turn.

Parts of Objects. One of the theoretical assumptions in this study is that the visual system parses the retinal image into whole meaningful components and not just any arbitrary segments. A possible description of a component can be conceptualized as the area defined by the contour of a part. Conversely, an area defined a contour with some missing portions can be considered analogous to arbitrary segments.

Based on the above definition of whole parts, one of the experimental conditions in this study involved manipulating the amount of contour that was presented on the screen, to see if the wholeness of a part affected componential integration. That is if the segments with missing contours are analogous to arbitrary parts, then integration of parts in this condition should be affected because the visual system is considered to parse an object into whole meaningful parts.

As mentioned earlier, the most crucial information in the integration of components is the information about the relationships between the parts of an object. Therefore in addition to information about parts, three types of relational information between parts of objects were also explored.

They are:

Concrete geometric Relation Information. One obvious source of information about the relation between two parts could be the join or the interface between them. In the retinal image, this is proposed to be the area defined by the contour around the interface between two parts and the actual information that is extracted from the interface is assumed to be the coordinates of the contour around this interface. The memory representation analogous to this type of relational information is conceptualized as an encoding of these coordinates in a spatiotopic map (see chapter 1 for details).

Abstract Spatial Relation information. Another source of information about the relation (which is in a propositional format) between two parts could be derived from the spatial configuration of the parts with respect to each other. This is assumed to be the abstract spatial relation information since it is propositional in nature. In the retinal image, this is postulated to be the information about whether one part is to the right or left or top or front etc of the other. Thus information that is extracted from the retinal image is proposed to be propositional in nature (see chapter 1 for details).

Potential Relation Information. The potential relational information is conceptualized to be a representation that is very general with all types of relational information (both structural and functional), that can enable the viewer to

integrate the parts of an object. By definition it then follows that the set of potential relations for any part is fairly large. However, in the case of objects this relationship would be conceivably constrained by various sources of information that are available, like the physical structure of the parts or knowledge about the parts themselves (functional knowledge) or whether the central axis of the parts are in a horizontal or vertical position which would determine whether the parts are in a horizontal or vertical configuration with other parts (see chapter 1 for a detailed discussion).

Assuming that our representations of componential information (both the parts and the relations between them) are as described above, then presumably under some conditions one should be able to access or use these representations. Particularly in a situation where components of an object and their spatial relations are presented one at a time in quick succession, the componential representations should function as salient units of information. Some examples of the different possible experimental conditions were discussed in the earlier chapter.

Hypotheses

There were three primary hypotheses that were explored in this study. The first hypothesis was that the representation (in the retinal image and in memory) of the components of objects consists of whole meaningful parts and

is highly visual in nature. Furthermore the information extracted from the retina is compared to an object centered 3-dimensional representation in memory.

To test this, two different types of segments called parts and pieces were used in the experiment. Parts (figure 1) were segments where the component of an object with the complete contour was presented and pieces (figure 3) were segments that had a slight missing contour.

If the representations of the components in memory are whole meaningful parts then the condition with parts should enable better recognition than the condition with pieces.

The second hypothesis was that there are three different types of relational information available in the retinal image with analogous representations in memory. They are the concrete geometric information, the abstract spatial information and the potential relational information.

Two manipulations were designed to test the importance of the three hypothesized spatial relation representations. One of the manipulations involved either presenting the object components in the correct spatial relation or incorrect spatial relation (this was discussed in the earlier chapter). When the components are presented in quick succession in the correct spatial location the system should be able to extract all three types of spatial relational information.

On the other hand when the components are presented in the incorrect spatial location only the abstract spatial and potential relational information are available.

If the two experimental conditions afford different spatial relational information then the condition with the correct spatial location of parts should be the more efficient condition in terms of recognition than the condition with the incorrect spatial location because in the correct location condition the availability of the concrete geometric information should allow faster integration of the part information.

The second manipulation involved either presenting the components and interface (Figure 2) between them or not presenting only the components without the interface (Figure 1). If the concrete geometric relation plays a role in the integration of componential information then the condition with the interface segment should enable more efficient recognition than the condition without the interface segment because in the interface condition there is redundancy since more than one type of relational information is available and the interface segment should enable faster integration of componential information.

A third hypothesis was that the access of componential representations occurs very quickly and fairly early in processing. One test of this hypothesis would be if evidence for the use and access of these representations can be seen at

a short SOA (where the time between each presentation frame is extremely small) and not a long one. Thus the experimental manipulation for this involved varying the SOA.

Three different SOA's were used in this experiment. They were 75 msec, 250 msec and 500 msec. However each segment was on the screen for 70 msec in all the three SOA's. If the access of componential representations occurs very early in processing, then effects of the experimental manipulations designed to explore componential analysis should be seen in the two short SOA's (75 msec and 250 msec) while in the 500 msec SOA they should cancel out because in the 500 msec SOA the presentation of the components are spread out in time and inferential processes can set in cancelling out the lower perceptual effects.

Experimental Design

In order to investigate these hypotheses the following experimental manipulations were designed. Subjects either saw trials with just components of an object or they saw trials which included components of an object and the interface between the components. Specifically the trials with the interface either included whole component and the interface called 'parts + interface' condition or they included the component with a portion of the contour missing and the interface called the 'pieces + interface'. Furthermore, the components were presented in the correct

spatial location on some trials and in the incorrect spatial location on other trials. Altogether there were 6 experimental conditions.

- (1) Parts only correct location (Figure 1).
- (2) Parts + Interface correct location (Figure 2).
- (3) Pieces + interface correct location (Figure 3).
- (4) Parts only incorrect location (Figure 1).
- (5) Parts + Interface incorrect location (Figure 2)
- (6) Pieces + Interface incorrect location (Figure 3).

Predictions

The six experimental conditions generated in five predictions. Foremost, if there is a difference in the efficacy of integration of componential information between the three types of relational information between the parts of an object (with the concrete geometric being most efficient) then the conditions where the segments were in the correct spatial location (1,2,3) should have a significantly faster response time and lower error rate than the conditions where they were in an incorrect location(4,5,6). In the former case there is access to the concrete geometric spatial relational information which would enable faster integration than the latter where only the two types of abstract relational information are available.

Secondly if information about the spatial relationships between parts of an object is represented in a concrete

geometric manner and affords faster integration of those parts then the conditions with the interface frame (2,3 and 5,6) should have a faster response time than the conditions without the interface frame (1,4) because the interface segment should enable the system to extract the concrete geometric relation.

The third prediction was that if the concrete geometric spatial relationship enables the most efficient integration of componential information then the condition where the segments are in the incorrect location (5 and 6) should show a greater facilitatory effect of the presence of interface frame than the condition where the segments are in the correct location (2 and 3). In the former case the interface segment is crucial to resolving the location ambiguity whereas in the latter case the interface is merely an extra source of information about the relationship between the components. In other words the greater importance of the interface segment in the incorrect location condition should result in an interaction between the interface and location conditions.

The fourth prediction was that if the representation of a component of an object is a whole meaningful part and if the efficiency of recognition depends on the closeness of match between the input and the appropriate representation then the condition with the parts and interface frames (2 and 4) should have a faster response time and lower error rate than the conditions that have the pieces and interface frames (3 and 6).

Lastly if the different componential representations are accessed and used early in processing, then the various effects of the experimental manipulations should only be seen in the 75 and 250 msec SOA and not in the 500 msec SOA because in the 500 msec SOA where the presentation frames are spread out in time, inferential processes could set in and cancel out all lower level perceptual effects.

The stimuli consisted of line drawings of object segments and the interface between two components. The parts of the objects were constructed by separating the objects into their constituent components (see Figure 1). The pieces were the components that were left when the interface portion was separated out (see Figure 3). The interface was the segment of the object that closely approximated the concrete geometric representation of spatial relationship between components (see Figure 2). This was constructed by taking a small portion of the object on either side of the join between two components.

CHAPTER 3

EXPERIMENT 1

This experiment used an object naming task.

Method

Subjects

Eighteen University of Massachusetts students (Graduate and Undergraduate) were recruited to participate in this study. They were all fluent in American English and had normal or corrected vision. The subjects were either paid or given class credits for participating in the experiment.

Stimuli

The stimulus set consisted of line drawings of 30 common inanimate and animate objects (see Appendix 1). The segments consisted of either the part or piece or interface of an object (see Figure 2 and 3). The objects either had vertically arranged or horizontally arranged components. 90% of the objects had a total of two components while the remaining 10% had a total of 3 components. The two-component objects had one interface while the three-component objects had 2 interfaces.

Design

The experiment employed a mixed design with two within and one between subject's factors. The within subject factors were: (1) The spatial order of presentation of components

(correct versus incorrect) and (2) The interface versus non interface conditions (parts only versus parts + interface versus pieces + interface). The between subjects factor was the SOA (75 msec, 250 msec and 500 msec). The three factors were crossed, which resulted in 18 experimental conditions.

Each subject was presented with a total of 180 trials. The trials were divided into 6 blocks with 30 trials in each block. The trials were blocked by the correct versus incorrect location condition such that every subject saw 3 correct and 3 incorrect blocks of trials (in either the CICICI or ICICIC order). Furthermore, the order of presentation of these blocks was counterbalanced across subjects. The interface versus non interface conditions were present in every block. The stimuli were equally divided among the conditions such that across the 6 blocks each object appeared once in every condition. The order of presentation of the interface versus non interface conditions within each block was randomized. The 18 subjects were equally divided between the three SOA's resulting in 6 subjects per SOA.

Apparatus

The stimuli were presented on a Megatek monitor (the resolution factor is 4,096 * 4,096 points) that was interfaced with a Vax 11-730 computer. The center of the screen was approximately 33.5 inches from the eyes of the subject. The visual angle for the stimuli varied between 3 - 5 degrees. A

voice key interfaced to the computer was used to record the latency of the subject's responses.

Procedure

The procedure involved a masked sequential presentation. A mask appeared on the screen to signal the start of a trial. Simultaneously a small cross appeared either at the top or bottom or left or right of the screen to precue the subject about whether the segments were in the correct or incorrect order and whether the components were vertically arranged or horizontally arranged. The subjects were precued in order to prevent any confusion regarding where to look since there were objects with both vertically and horizontally attached parts. Although the correct and incorrect location trials were blocked it was felt that precuing the location would further strengthen evidence for the for any effects of the location condition. Therefore a cross on the top meant the segments would be displayed from top to bottom in the incorrect location. A cross on the bottom meant that the segments would be displayed from bottom to top in the correct location. A cross at the left indicated that the segments would be displayed from left to right in the correct location while a cross at the right indicated that the segments would be displayed from right to left in the incorrect location.

Each segment was visible on the screen for 70 msec. The SOA was the time from the appearance of the first segment to

the start of the next segment. Thus the 75 msec SOA had an ISI of 5 msec, the 250 msec SOA had an ISI of 180 msec and the 500 msec had an ISI of 430 msec. During the ISI the mask that appeared at the beginning of the trial was presented.

The sequence of events in a trial consisted of the appearance of a mask at the start of a trial followed by the first segment, the mask, next segment and the mask in the non interface condition. In the interface condition the sequence was the same. There was an alternating sequence of the mask and the segments. Since the presentation was a masked sequential one, every display was followed by a mask. The end of the trial was signalled by the appearance of the whole mask, and as soon as this mask appeared the subject was required to name the object. The response time was recorded from the beginning of the last segment to the start of the response. In addition the errors were also recorded.

Results and Discussion

The main focus of Experiment 1 was the role of components in the identification process. As discussed in the earlier chapter the experimental conditions were designed to test specific assumptions about the different types of componential information that are involved in the recognition process. The results reported here are primarily the significant effects. However any non significant results of theoretical interest are also discussed. The two measures

that were analyzed were the response times (naming latencies) and the error rates.

Two important concerns in Experiment I necessitate that the results be interpreted with caution. One concern was the possible effects of the familiarity of objects used as stimuli. All the stimuli in Experiment 1 were line drawings of very common objects which when combined with repetition over several blocks of trials might have enabled identification on the basis of a single part. Therefore it is possible that any effects of the experimental manipulations might have been diminished.

A second concern in Experiment 1 involved the difference in the total exposure time between the trials that had the interface segment and the trials that did not have this segment. Therefore the possible effects of the presence or absence of the interface are confounded with the differences in the time between the onset of the first segment and the onset of the last segment.

An important aspect of this study was to test two possibilities: (1) three types of relational information (concrete geometric, abstract spatial and potential relational) between parts of objects are available in the retinal image and also represented in memory and (2) the efficacy of the integration of componential information is contingent on the type of relational information that is available (with the concrete geometric being predicted to

be the most efficient). The manipulations designed to test these hypotheses were the location condition (the parts either in the correct or the incorrect spatial location) and the interface (seeing the interface segment versus not seeing the interface) condition.

As predicted, there was a main effect of the location manipulation. Overall, the correct location trials had a faster response time than the incorrect location trials (784 vs 832) $F(1,30)= 8.53, p <.01$ (see Table 1). There was a main effect of the location condition in the error data with the correct location trials (19%) having a lower error rate than the incorrect location trials (26%), $F(1,30)=15.36, p<.001$. The direction of this effect was the same as the direction of the location effect in the response time data. This indicates that concrete spatial information is used in integrating the part information. Both conditions, however, afford the extraction of the abstract spatial and potential relation information.

It is also possible that redundancy facilitates integration of componential information and that extracting three types of relational information (afforded when the components are in the correct location) enables a more efficient integration than having two types of relational information. This is not a surprising finding if object recognition is in part conceptualized as a threshold process. This conceptualization would suggest that the more information

there is, the faster the system would reach the threshold for a particular stage in processing.

Further support for the idea that the visual system is sensitive to more than one type of relation information between parts of objects comes from the fact that the difference between the correct and incorrect location trials is rather small (48 ms and 7% errors). If the system almost completely relied on the concrete geometric information, performance in the incorrect condition should have been very slow and/or had many errors.

Surprisingly the location manipulation did not interact with SOA $F(1,30)=.04$, $p < .95$. There were differences of 55 ms (in the 75 ms SOA), 46 ms (in the 250 ms SOA) and 48 ms (in the 500 ms SOA) between the correct and incorrect location trials. While the interaction in the error data was also not significant, the pattern of error does suggest that the location differences in the 500 ms were smaller than the location differences in the 75 or 250 ms SOA (see Table 1). This might be due to the effects of the additional processing time in the 500 ms SOA that would have helped subjects to resolve the location ambiguity more efficiently.

As discussed earlier, concrete geometric spatial relationship information was hypothesized to enable the most efficient and fastest integration. Therefore the interface segments were predicted to facilitate integration of the segments since the interface segments were assumed to

approximate the concrete geometric information and hence were expected to provide direct access to the concrete geometric spatial representation in memory.

Overall the two interface conditions had a faster response time than the non interface condition (768 vs 778 vs 879) $F(2,30)=8.53$, $p<.001$. Furthermore, a contrast performed on the combined data of the two interface conditions with the non interface condition for the correct location (867 vs 742) $F(1,30)=21.75$, $p<.0001$ and incorrect location (890 vs 803) $F(1,30)= 15.60$, $p<.0001$ was found to be significant.

The interface by SOA by location interaction was not significant $F(4,30)=1.23$, $p <.32$. Furthermore, the location by interface interaction was not significant $F(4,30)=1.12$, $p <.34$ indicating that the correct location trials did not have a significantly larger interface effect than the incorrect location trials.

There was, however an interface by SOA interaction $F(4,30)=12.16$, $p <.0001$ (see table 1). This interaction is somewhat surprising because the greatest facilitation for the interface segment is in the 500 ms SOA (see table 1). This is contrary to the prediction that any effects of the interface were expected to be minimal or non existent in the 500 ms SOA, because the subjects were expected to use some kind of problem solving strategies at long SOA's due to the availability of additional processing time.

Also of interest is a comparison of the non interface condition with the two interface conditions (parts only versus parts + interface + pieces + interface) for the correct location and incorrect location trials for the three SOA's. In the 75 msec SOA for the correct location trials (see table 1 for means) the contrast between the parts only and the parts +interface and pieces+interface is significant $F(1,30)=8.22$, $p<.03$. In the incorrect location trials the contrast is not significant $F(1,30)=.65$, $p<.45$. In the 250 msec (see table 1 for means) the correct location trials show a significant difference between the parts only and parts+interface and pieces + interface condition $F(1,30)=7.6$, $p<.04$. In contrast the incorrect location trials show no significant difference between the parts only and the parts +interface and pieces + interface conditions $F(1,30)=.40$, $p<.56$. The 500 msec SOA (see table 1 for means) shows a very different pattern of results. Here both the correct location and incorrect location trials show a significant difference between the interface and the two non interface conditions. In the correct location trials the contrasts between the parts only and parts + interface and pieces + interface are $F(1,30)=76.9$, $p<.001$. For the incorrect location trials the contrasts for the parts only and parts + interface and pieces + interface conditions are $F(1,30)=117.31$, $p<.001$.

The contrasts reported above are surprising and contradictory because it appears the interface had a facilitatory effect in the incorrect location trials only in the 500 ms SOA. According to the hypotheses and predictions outlined in chapter 2 the interface segments were expected to be maximally facilitatory in the incorrect location trials for the 75 and 250 ms SOA while the 500 ms SOA was predicted to show no significant effects of the experimental manipulations.

It would appear that if indeed there is a concrete geometric relationship information which enables efficient integration then the effects of the interface should be especially helpful when the spatial relationship was not given by the actual positioning of the parts (as in the incorrect location trials). However the lack of interface effect in the incorrect location trials for the 75 and 250 msec SOA seems to suggest otherwise.

Thus, in conclusion, it appears that although there was an overall location and interface effect it is hard to consider this result as unequivocally supporting the idea that componential information (elaborated in the introductory chapter) plays a significant role in the parsing and integration stages of object recognition.

Furthermore, the various analyses described so far seem to indicate that evidence for the role of concrete geometric relation information in the integration of componential information is somewhat inconclusive. While there appears to

be a facilitatory effect of the interface segment, it is difficult to dissociate this from the possible effects of additional exposure time.

One interesting finding in this study was the lack of a significant difference between the parts and pieces condition. It was predicted that if our representation of a component of an object consists of whole meaningful parts and if the efficiency of recognition is contingent upon the closeness of match between the input and the appropriate representation then the pieces condition should have shown a slower response time. Subjects were however equally fast and accurate in this condition. It is possible that the lack of difference here is because the appropriate components could have been quite easily recoverable from the pieces. For instance less distinct portions of an object would probably show a significant slow down in response time and perhaps the pieces were not a sufficiently sensitive test for this slow down.

Another important issue in the study was the time course of processing. The question is when does componential information become available. This issue in part motivated the three SOA conditions in the experiment. From the results it appears that componential analysis if it occurs, happens quite early in processing. This is reflected in the fact that even in the 75 msec SOA (807 vs 745 vs 782) there was an effect of the interface/non interface manipulation. There was

an interface/non interface * SOA interaction (see table 1) $F(4,30)=12.16$, $p < .0001$. However these results need to be viewed with caution. The timing manipulation did not yield data that could lead to unequivocal conclusions about the time course of processing of componential information. Although some of the data indicates that componential processing occurs very early in processing, it is contradicted by the fact that the greatest facilitatory effects of the interface were seen in the 500 ms SOA.

Overall, it seems difficult to come to any firm conclusions about the assumptions of the componential approach to object recognition. While it appears that in order to name the object whose components were presented sequentially, subject's were to some extent processing the parts, the results make it hard to generalize the performance in this task to object recognition in the real world because here the parts are not really presented sequentially most of the time. On the other hand the error data (see table 2) seem to suggest that the subjects performance in the experiment was rather poor considering they had to name familiar objects. The error rates suggest that perhaps componential information does not play a significant role in the parsing and integration stages of recognition.

Perhaps object recognition is a more complex and flexible process than was outlined in the introductory chapter. Therefore the experimental manipulations may not have been

sensitive enough to reveal the subtle effects. Experiment 2 was designed to see if the basic pattern of results in Experiment 1 (particularly the interface effects) could be replicated after eliminating the time and familiarity problem of Experiment 1.

CHAPTER 4

EXPERIMENT 2

Experiment 2 involved an object decision task and essentially had the same conditions as Experiment 1. There were two important concerns in Experiment 1. The first concern was that possible effects of the familiarity of objects used as stimuli which might have diminished any effects of the experimental manipulations. Secondly, the difference in the total exposure time between the trials that had the interface segment and the trials that did not have this segment might have been confounded with the effects of the interface manipulations. Therefore the second experiment was chiefly designed to see (a) whether the interface effects could be replicated with the timing problem minimized and (b) whether the location effects could be bigger when the subjects could not guess the correct response based on partial information (e.g., based on seeing the first part).

In order to overcome the familiarity problem, an object decision task was employed. The objects used in Experiment 2 were identical to those used in Experiment 1 while the non-objects were constructed by recombining the components of objects from Experiment 1. Thus the objects and non objects shared the same set of components but differed in the spatial relationships between the components. It was hoped that this would minimize the familiarity factor by preventing the subjects from making the response by processing a single

segment, since subjects would have to see all the segments before making the response.

The timing confound of Experiment 1 was minimized by a slight modification to the non interface condition. In this condition the presentation frame included an additional mask segment between the two part segments. This equated the total exposure duration of the trial with the trials in the interface conditions and the time between the initial segment and the final segment. In addition the 500 ms SOA was dropped in Experiment 2 because the addition of the mask segment seemed to perceptually disrupt the apparent continuous flow of segments.

Method

Subjects

Thirty six University of Massachusetts students (undergraduate and graduate) were recruited to participate in this experiment. They were all fluent speakers of English and had normal or corrected vision. They were paid or given class credits for participating in the experiment.

Stimuli

The stimuli consisted of parts of 30 common objects (animate and inanimate) and 30 non-objects. Both the objects and non-objects the stimulus were divided into parts, pieces

and interface as in Experiment 1. The object stimuli were the same as in Experiment 1 and had the same parts, pieces and interfaces. The non-objects were constructed by taking the parts of different objects (from Experiment I) and placing them together to form a composite whole (see Figure 6). Consequently, all the non-objects had two parts and one interface.

Design

The conditions for this experiment were basically the same as experiment 1.

(1) Presentation Type

There were 3 types of presentation. They were parts only, parts + interface and pieces + interface.

(2) Location

As in Experiment 1 there were 2 location conditions: the correct and incorrect location. The parts were presented either in the right spatial configuration or wrong spatial configuration with respect to each other.

(3) SOA

Only two SOA's (75 msec, 250 msec) were used in Experiment 2. The experiment had a mixed design with the presentation type and location as the within subject factors and the SOA as the between subject factor.

Procedure

Unlike Experiment 1 this experiment had only two blocks of trials. The trials were blocked by location such that every subject saw one correct location block and one incorrect location block. The order of presentation of the correct and incorrect blocks were counterbalanced across subjects. The 3 type conditions (parts only, parts + interface and pieces + interface) appeared in every block. The order of presentation of the stimuli was counterbalanced across blocks and randomized across subjects.

As in Experiment 1 a trial consisted of a masked sequential presentation on the Megatek monitor. Since this experiment involved an object decision task, the subjects were required to make the response by pressing a right hand key for an object and a left key for a non object. As in Experiment 1, both the response times and error rates were the measures of interest.

Results and Discussion

Experiment 2 was the second part of this study and was designed to see if clearer and more interpretable results could be obtained by eliminating the familiarity and total processing time problems of Experiment 1. The results reported here are primarily the significant effects. Furthermore non significant effects of theoretical interest are also discussed.

The response times and error rates in Experiment 2 were essentially comparable to the data from Experiment 1 (see Tables 1,2,3 and 4). However the results in Experiment 2 were just as contradictory and difficult to interpret as Experiment 1, reflecting the possibility that perhaps componential parsing and integration is a very complex but flexible process. Therefore it is possible that the experimental manipulations were not sufficiently sensitive to the complex and subtle aspects the componential stages of the recognition process.

The location manipulation was expected to test the validity of the assumption that the visual system is capable of extracting different types of relational information (between object parts) from the retinal image. Furthermore the concrete geometric relationship information which was available in the correct location condition (but not in the incorrect location condition) was hypothesized to enable the most efficient integration.

As predicted, there was a main effect of the location condition. Overall, the correct location condition had a faster response time than the incorrect location condition (1099 msec vs 1132 msec) $F(1,32)=7.26$, $p < .01$. This replicates the results of Experiment 1. Furthermore the error rates in Experiment 2 were consistent with the response time data in that the correct location condition had a lower error rate than the incorrect location condition (23% vs 25%),

$F(1,32)=5.319, p < .03$. Thus it appears that the subjects were somewhat quicker and more accurate in responding when the parts were seen in the correct spatial configuration than when they were in an incorrect configuration.

These findings are qualified however, by the fact that there was a location by SOA interaction $F(1,32)= 5.071, p < .03$. At the 75 msec SOA, the correct location trials had a faster response time than the incorrect location condition (941 vs 1002) while at the 250 msec SOA there was no significant difference in response times between the two location conditions (1257 vs 1262). This interaction is surprising because it indicates that perhaps at the 250 ms SOA the increased processing time proved beneficial, which is contrary to the findings in Experiment 1 where there were location effects even in the 500 ms SOA.

An error analysis also revealed a location by SOA $F(1,32)=23.43, p < .0001$ interaction. At the 75 msec SOA, the correct location had a lower error rate than the incorrect location conditions (29% vs 35%) where as in the 250 msec SOA the correct location condition had a slightly higher error rate than the incorrect location condition (17% vs 15%). Since the 75 and 250 ms SOA show a different pattern of results it is difficult to attribute the observed error rates to any particular factor.

The location manipulation was expected to test the validity of the idea that more than one type of relation

information is extracted from the retinal image. The main effect of the location manipulation and the location by SOA interaction for both the error rates and response times is consistent with either: (1) the idea that redundancy helps: three types of relationship information seem to enable better integration than two and/or (2) the idea that the concrete geometric relationship afford more efficient integration. However, the interaction with SOA indicated that this information was of little benefit at the 250 ms SOA.

If the concrete geometric relation is a crucial unit of information in the integration process, then the interface conditions should enable more efficient recognition than the non-interface conditions. As predicted, there was a main effect of the interface in the response times. The non-interface condition had a longer response time than the interface (parts + interface and the pieces + interface) conditions (1152 vs 1091 vs 1103) $F(1,32)= 12.63, p <.001$.

This apparently replicates the results of Experiment 1. While the interface effect indicates that the concrete geometric relation information may afford a more efficient integration than the other two types of relation information, the interpretation is complicated by a significant SOA by location by type interaction $F(1,32)=4.066, p <.02$. At the 75 msec SOA, there seems to be no differences in the overall response times between the non interface and interface conditions (see Table 3) while at the 250 msec SOA both the

interface conditions are faster than the non interface condition. Furthermore at the 250 msec SOA the interface conditions in both the correct and incorrect location are faster than the non interface condition (see table 3) while this is not the case at the 75 ms SOA.

The error results in this experiment however, are much more complicated than the response time results. Overall, the parts + interface condition (which was expected to enable the most efficient integration) had the highest error rate (31%). On the other hand, there was no significant difference between the parts only and the pieces + interface condition. This was reflected in a main effect of the interface manipulation $F(1,32)=26.80$, $p < .001$. Furthermore, this appears to be consistent in the 75 and 250 ms SOA as well as in the correct and incorrect location trials of the two SOA's which can be seen in the lack of an interaction between the interface, location and SOA manipulation $F(1,32)=.95$. There appears to be no apparent explanation for this anomaly. The possibility that the interface segment might be have an interfering effect is not a satisfying explanation because the pieces + interface condition (which also has the interface segment) does not seem to show this effect.

Further analysis were done to see if there were any object non-object differences that could help in interpreting or resolving some of the ambiguous findings in Experiment 2.

Overall the objects had a faster response time than the non-objects (1079 vs 1152), $F(1,32)=22.87$, $p < .0001$. This is not very surprising given that we are familiar with objects but not non-objects. This could also be the consequence of a response bias, because of the tendency to say 'yes' faster than to say 'no'.

There was an object/ non object by location interaction $F(1,32)=10.186$, $p < .003$. For the objects, the correct location trials were 56 ms faster than the incorrect location trials (1051 vs 1107), whereas in the difference for non-objects was only 11 ms (1147 vs 1158). This finding can in fact be considered as supporting the notion that perhaps we have stored representations of the componential information that are accessed during the integration stage. One explanation could be that since the objects presumably have stored representations, the disruption of spatial configuration of parts should affect the integration of object more than the integration of non objects (which lack representations of relation information).

This interpretation however becomes slightly complicated if one looks at the response times for objects and non-objects separately in the interface and non interface condition. In both cases there seems to be a significant interface effect. For objects the response times for the parts only, parts + interface and pieces + interface are (1119 vs 1058 vs 1059). For the non-objects they are (1147 vs 1124 vs 1147). However,

this interaction is marginally significant $F(1,32)=2.8$, $p < .06$. The response times for the objects are straightforward indicating that viewing the interface facilitated the integration of the components. In contrast the response times for the non-objects are rather unusual because the interface seems to have facilitated the integration of components in the whole part condition and not the pieces + interface condition. It could be the case that this is merely a redundancy effect since more information is available in the parts + interface condition than the pieces + interface condition.

The error data can be interpreted as basically reflecting a speed accuracy trade off. Overall, the objects had a higher error rate than the non objects (27% vs 21%), $F(1,32)=6.71$, $p < .01$. There was an interaction of objects /non-object by location by SOA $F(1,32)=6.58$, $p < .003$.

At the 75 msec SOA, there was no difference in the error rates between objects and non objects for the correct location (28 vs 29). On the other hand objects in the incorrect location had a higher error rate than the non-objects (40% vs 31%).

The 250 ms condition had essentially the same pattern of errors as the 75 ms condition. For the correct location trials the objects had a lower error rate than the non objects (16 vs 18) while the pattern reversed for the incorrect location trials. These results seem to support the idea that the spatial configuration parts is more salient for objects

than the non objects indicating that perhaps we have stored representations of the information about relationship between parts for objects. On the other hand, it is possible that it is easier to decide that something in an unusual configuration is a non-object.

There was also a hint of an interaction between object/non-object by location by component type in the response time data (see table 3) $F(1,32)=2.9$, $p < .06$. This interaction seems to be stronger in the error analysis $F(1,32)=6.58$, $p < .002$. It is interesting to note that for the non-objects the interface condition in the incorrect location condition is faster than the correct location condition. It is not clear why this is so.

In conclusion, it appears that there were no obvious differences between the objects and non objects to enable any substantive conclusions about componential processing. However, the fact that disrupting the spatial configuration of parts had a greater effect on the objects than the non objects suggests that objects may indeed have stored memory representations for relation information which are sensitive to disruptions. Overall, Experiment 2 did not provide unequivocal evidence for the assumptions of the componential approach. Basically the data from this experiment was similar to the data from Experiment 1 in that there is no clear supporting evidence that componential information is extracted and used during the recognition process.

The componential approach seems to provide a fairly logical and plausible account of the recognition process. As mentioned earlier it is possible that componential processing is very complex and subtle process that can be demonstrated only by very sensitive and subtle experimental manipulations. Although the study did not provide strong evidence in favor of the componential approach, the results suggest definite directions for further research.

CHAPTER 5

GENERAL DISCUSSION

Recent trends in the theoretical and experimental research on object recognition show an emphasis on the importance of componential information in the identification process (Biederman 1987, Hummel and Biederman 1991). The principal idea underlying this viewpoint is that the early stages of object recognition involve the parsing and integration of componential information from an object.

It can be argued that there is nothing inherently distinct about the basic premise of the componential approach since most current theories of object identification propose an initial parsing and integration of the object into smaller units of information. The critical difference between the theories however is in their definition of the salient units of information and the representation of this information. As described in the introductory chapter, the componential approach assumes that information from an object can be differentiated into several levels of increasing complexity and furthermore, the salient units of information at an intermediate level are the components and their relationships (spatial and other).

The hypotheses and predictions in this study were formulated within the componential framework while alternative predictions were based on the template approach. The template approach that was used as the basis for the alternative

hypothesis and predictions was an extreme version of the view. It was felt that the template approach would serve as an adequate conceptual contrast since the primary goal of this study was to explore some of the assumptions of the componential approach and the experimental manipulations were not designed to be sensitive to subtle differences between alternative theories that adopt a similar parsing and integration approach. This section provides a discussion of the findings from this study in relation to the assumptions about the parsing and integration principle, the different types of componential information and the time course of componential processing.

The Parsing and Integration Principle

In order to investigate the assumption of parsing and integration of componential information, the paradigm involved a sequential presentation of the stimulus segments, since a parsing and integration theory of object recognition predict a successful integration of componential segments that are displayed in quick succession. The findings from the study did not provide unequivocal evidence in support of this assumption.

The response times and error rates across both experiments (which had different tasks) were comparable. The response time data appeared to indicate that the subjects were fairly fast. However the error rates were higher than

expected, particularly in the naming task. It would appear that a speed accuracy trade off is at best a superficial explanation because a masked sequential presentation is not an approximation of the normal viewing conditions (especially situations where all the whole object is seen). Thus the presentation method might have been a closer approximation of more difficult viewing conditions where parts of objects are occluded or there is insufficient lighting etc (see introductory chapter pages 14-17 for a more detailed discussion), therefore that performance in this task may have been a better reflection of difficult viewing conditions. It is thus unlikely that accuracy would have improved if speed were not emphasized.

Comparing a simultaneous presentation with a sequential presentation would probably be a more sensitive test of the parsing and integration assumption. In a simultaneous presentation where parts are presented in an incorrect configuration, if performance were as poor as the analogous condition of a sequential presentation, it can be considered as strong evidence for the idea that the object recognition process involves the parsing and integration of parts and their relationships.

The Componential Representations

To explore the assumption that at an intermediate level the salient units of information consist of object components

(whole meaningful parts) and the relationships between them (spatial and other), four types of componential information were hypothesized (see Chapter 1 for a detailed description).

Parts Of Objects

One of the theoretical assumptions in this study was that the visual system parses the retinal image into whole meaningful components and not just arbitrary segments. A possible criterion for the wholeness of a part could be the amount of contour that describes it. Therefore components that have the entire contour (parts) can be considered whole meaningful components while components that have a missing contour (pieces) can be viewed as non-whole segments.

The experimental manipulation that tested this assumption involved presenting two different types of segments called 'parts' (Figure 1) and 'pieces' (Figure 3) where the parts were segments with a complete contour and pieces were segments with a missing contour. If our system parses an object into whole components, then integrating complete parts should be easier than integrating arbitrary segments. Alternatively (according to the whole template) if the system parses an object into arbitrary parts then there should be no difference in the efficiency of integration afforded by the two experimental conditions. The data from both the experiments indicated that there were no significant differences between the two conditions.

While this lack of a difference may appear to support the predictions of a whole template approach, a possible alternative is that the manipulation might have lacked sufficient sensitivity to pick up any differences. Since the 'pieces' were segments that had only a slight portion of the contour missing, it is possible that the visual system may have had sufficient information to recover the whole part and thus masked any differences between parts and pieces.

Comparing integration of whole segments with integration of arbitrary halves of the same object may be a more sensitive test of this assumption. Some possible arbitrary segments are vertical halves of objects whose parts are horizontally connected or horizontal halves of objects whose parts are vertically connected.

An underlying implication of the idea of representations of whole meaningful parts of objects is that these representations are highly visual in nature. In this study since the components were assumed to be visual in nature the manipulations were not designed to distinguish between two alternatives like a visual or abstract representation. There are however other studies that claim to do so.

Biederman and Cooper (1992) reported a series of repetition priming experiments where they manipulated the size and different forms of the same object. They obtained an advantage of identical image over different exemplar, suggesting that the priming effects were visual. In another

set of experiments , Biederman et al (1991) presented subjects with complementary images formed by deleting every other edge. There was no significant difference in priming by the same image and the complementary one. They were also named faster than different exemplars. They reasoned that our memory representations accessed during object identification is highly visual in nature because if representations of object components are abstract or non-visual then there should have been no differences between the identical prime and a different exemplar.

It is possible that our memory representations of components have both a visual and abstract component. Several researchers (Rosch et al 1976 and Tversky 1984) have investigated the idea that object parts are categorized on the basis of their function. A functional component can be considered an abstract or non-visual attribute.

Relationships Between Object Parts

A second type of componential information consists of the relationships between parts of objects. The experimental manipulations were designed not only to demonstrate the importance of relationship information between parts of objects but also provide some evidence for how this relational information might be represented in memory. Three types of relational information were outlined and explored in the study (refer to chapter 1). They were concrete geometric relational

information, abstract spatial relation information and potential relational information.

The concrete geometric information is conceptualized as the area defined by the contour around the interface between two parts. The information that is extracted from the interface is considered to be the coordinates of the contour around this interface which is represented in memory as an encoding of these coordinates in a spatiotopic map.

The abstract spatial relation is propositional information about where one part is in relation to the other for example to the right or left or top or back etc. The relational information that is extracted from the retinal image are the above described propositions and two possible memory representations are considered. One possibility is a multiple viewer centered propositional representation of the abstract spatial relational information. For example, in the case of a cup the memory representation would be the handle goes to the left or right or front or back of the cylindrical body. A second possibility is a single object centered propositional memory representation. An example of the abstract object centered spatial relationship between the parts of a cup would be 'the handle goes to the side of the cylindrical body'.

The potential relational information is conceptualized as a very general representation that includes all types of relational information (both structural and functional), that

can enable the viewer to integrate the parts of an object. In the case of objects the set of potential relations is assumed to be constrained by various sources of information that are available, like the physical structure of the parts or knowledge about the parts themselves (functional knowledge) or whether the central axis of the parts is in a horizontal or vertical orientation which would determine whether the parts are in a horizontal or vertical configuration with other parts. The information that is extracted is proposed to be propositional in nature and similarly represented in memory.

Two manipulations were designed to test the three hypothesized relationship information. One involved presenting the object components either in the correct spatial relation or an incorrect spatial relation. When presented with the components in quick succession in the correct spatial location, the system was expected to extract all three types of spatial relational information while, when presented with the information in the incorrect spatial location the system was expected to afford only the abstract spatial and potential relational information. Therefore the condition with the correct spatial location of parts was predicted to result in the more efficient recognition than the condition with the incorrect location condition.

Results from both Experiment 1 and 2 supported this prediction. Overall the correct location trials showed a faster integration than the incorrect location trials.

However as predicted there was a location by SOA interaction only in Experiment 2 but not in Experiment 1. This result needs to be viewed with caution because the error rates in both experiments were rather high although there was a main effect of location in both the experiments with the errors being higher in the incorrect location.

The location effect does seem to indicate that the visual system is capable of extracting more than one type of relation information. A crucial aspect of the results is the relatively small difference between the incorrect and correct location trials in Experiment 1 and 2. This is perhaps the most compelling evidence that there is actually more than one type of information about the relationship between parts of objects and that the concrete geometric information plays a fairly minor role at least in a sequential presentation. If there were only a single type of relationship information (concrete geometric) performance in the incorrect location trials would have been much worse than the correct location trials.

The second manipulation involved either presenting the components and interface (Figure 2) between them or only presenting the components without the interface (Figure 1). The interface segments were expected to enable direct access to the memory representation of the concrete geometric information. Thus the condition with the interface segment was expected to enable more efficient recognition than the

condition without the interface segment. This was an additional test of the significance of concrete geometric relation. The findings with respect to the interface manipulation were however inconclusive.

First let us consider the data from Experiment 1. The facilitatory effects of the interface suggested that perhaps there are different types of relational information and that the concrete geometric information is highly significant in the integration of componential information. However this effect could have been merely a reflection of extra exposure time available in the two interface conditions.

If the concrete geometric information is a salient unit of relationship information, then there should have been a bigger interface effect in the incorrect location trials because in the correct location trials the interface segment is merely an additional source of concrete geometric information while in the incorrect trials it is crucial in resolving the location ambiguity.

However, there were smaller interface effects in the incorrect location trials than the correct trials. This suggests that if the presence of more than one type of relationship information results in a cumulative contribution towards integration then the effects of the interface segment would presumably be small. However, it is again somewhat puzzling that the interface effects were greater in the 75 msec SOA than in the 250 msec SOA in Experiment 1. If

additional exposure time were the reason then there should have been a consistent increase of the interface effects with increases in SOA across the two experiments. Nevertheless there seems to be very little consistency in the interface effects to make strong claims.

The data from Experiment 2 show a similar pattern for the interface manipulation. There is a main effect of the interface manipulation, suggesting that concrete geometric information is important. These findings are, however, complicated by the presence of various higher order interactions and high error rates in one of the interface conditions. In the 75 ms SOA, it appears that the interface conditions are not significantly faster than the non interface condition. The pattern is different at the 250 ms SOA where the interface conditions are significantly faster than the non interface condition. Furthermore, at the 250 ms SOA the interface conditions for both the correct and incorrect locations are faster than the same for the non interface conditions. The 75 ms SOA trials however do not show this pattern. The finding that the interface had an effect in the 250 ms SOA but not the 75 ms SOA is somewhat surprising.

There was also an unexpected result in the error analysis of Experiment 2 where there was a much higher error rate in the parts + interface condition (for 75 and 250 ms SOA) compared to the other two conditions. There is no obvious explanation for this result.

It is possible that the high error rate in the parts + interface condition in Experiment 2 is a result of task demands. Experiment 2 involved an object decision task where both the objects and non objects shared the same set of parts. The only difference between the two types of stimuli was in what parts were combined. If it were the case that the visual system relied on the perceived completeness of the segments to make the decision then, the interface segment could have caused interference in the parts + interface condition (where all the information was available in the two segments that contained the parts). This suggests that perhaps there is a threshold for redundancy of information where too much information might actually interfere rather than facilitate the recognition process.

Time Course of Componential Processing

A final manipulation involved the timing. This was expected to give some idea about the time course of componential processing. However it was not possible to make any conclusive inferences from the results.

The basic assumption was that componential information is extracted early in processing, and therefore that effects of the manipulations designed to tap this information should be at the very short SOA (assuming that small SOA's are an approximation of the early stages of the time course of processing). A longer SOA on the other hand can be considered

to stretch out normal processing time. Therefore performance in the 500 ms SOA was expected to show no effects of the experimental manipulations (designed to investigate the relationships between parts) because there was presumably sufficient time for the subjects to use problem solving strategies to resolve relation ambiguities.

However, contrary to the prediction, the facilitatory effects of the relation information was the greatest at the 500 ms SOA in Experiment 1. Furthermore the facilitatory effects of the interface manipulation were inconsistent between the two experiments. It is possible that the facilitatory effects were an artifact of the increased exposure duration in Experiment 1 (in this condition) and/or the redundancy of information. It appears that the timing manipulation was not really sensitive to the time course of componential processing in object recognition.

Conclusions

Overall the findings from the study are largely inconclusive as far as the assumptions of the componential theory are concerned. The main finding is the location effect in Experiment 1 and 2 indicating that redundancy in relational information facilitates integration of componential information and that concrete geometric information may be important. However the small location effects in both experiments suggest that in a sequential presentation the

concrete geometric information may not be particularly important.

The results of this study indicate that componential processing may be a lot more complex and subtle than was outlined in the theoretical assumptions of the study. The experimental manipulations may have lacked sufficient sensitivity to uncover the subtle effects. Despite the inconclusive results, the findings from this study have provided definite directions for further research in this area. One possibility is to try and use a more sensitive design to demonstrate that the representations of parts of objects consist of information about the whole meaningful parts and not merely arbitrary segments. Another possibility is to see if it is possible to use an experimental design that would be able to provide an unequivocal evidence for the use of information about relationships between parts of objects in the recognition process.

75 Msec

	Parts Only	Parts + Interface	Pieces + Interface	Mean
Correct Location	797	711	794	754
Incorrect Location	818	780	780	806
Mean	808	745	787	

250 MSEC

	Parts Only	Parts + Interface	Pieces + Interface	Mean
Correct Location	888	770	794	817
Incorrect Location	888	869	838	863
Mean	886	820	816	

500 Msec

	Parts Only	Parts + Interface	Pieces + Interface	Mean
Correct Location	917	735	700	784
Incorrect Location	971	742	772	828
Mean	944	738	736	

Table 1. Response Times By SOA and Presentation Condition in Experiment 1.

75 Msec

	Parts only	Parts + Interface	Pieces + Interface	Mean
Correct Location	15%	19%	17%	17%
Incorrect Location	30%	30%	30%	30%
Mean	23%	25%	24%	

250 Msec

	Parts only	Parts + Interface	Pieces + Interface	Mean
Correct Location	17%	14%	18%	16%
Incorrect Location	23%	20%	27%	23%
Mean	20%	17%	23%	

500 Msec

	Parts Only	Parts + Interface	Pieces + Interface	Mean
Correct Location	21%	22%	24%	22%
Incorrect Location	29%	25%	25%	26%
Mean	25%	24%	25%	

Table 2. Error Rates by SOA and Presentation Condition in Experiment 1.

Overall

	Parts Only	Parts + Interface	Pieces + Interface	Mean
Correct Location	1119	1058	1059	1079
Incorrect Location	1119	1097	1113	1132
Mean	1153	1078	1086	

75 Msec

	Parts Only	Parts + Interface	Pieces + Interface	Mean
Correct Location	936	935	910	927
Correct Location	1008	1007	1032	1015
Mean	972	971	971	

250 Msec

	Parts only	Parts + Interface	Pieces + Interface	Mean
Correct Location	1296	1223	1251	1257
Incorrect Location	1365	1240	1261	1289
Mean	1330	1232	1256	

Table 3. Response Times by SOA and Presentation Condition in Experiment 2.

Overall

	Parts Only	Parts + Interface	Pieces + Interface	Mean
Correct Location	19%	29%	20%	23%
Incorrect Location	21%	32%	22%	25%
Mean	20%	31%	21%	

75 MSEC

	Part Only	Part + Interface	Pieces + Interface	Mean
Correct Location	29%	29%	28%	29%
Incorrect Location	32%	37%	37%	36%
Mean	30%	33%	32%	

250 MSEC

	Parts only	Parts + Interface	Pieces + Interface	Mean
Correct Location	10%	29%	12%	17%
Incorrect Location	9%	27%	7%	15%
Mean	10%	28%	10%	

Table 4. Error Rates by SOA and Presentation Condition in Experiment 2.

Objects

	Parts Only	Parts + Interface	Pieces + Interface	Mean
Correct Location	1076	1025	1090	1051
Incorrect Location	1163	1090	1066	1107
Mean	1120	1058	1059	

Non-Objects

	Parts Only	Part + Interface	Pieces + Interface	Mean
Correct Location	1164	1144	1133	1147
Incorrect Location	1208	1104	1161	1158
Mean	1186	1124	1147	

Table 5. Response Times by Object/Non-Object and Presentation Condition in Experiment 2.

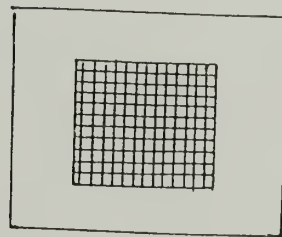
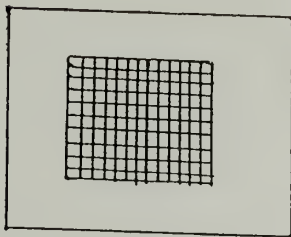
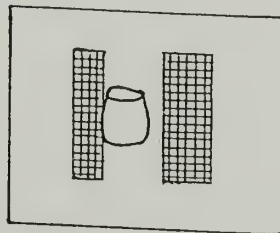
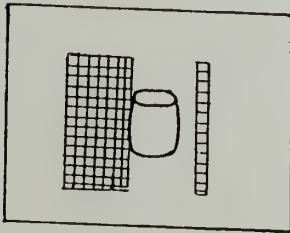
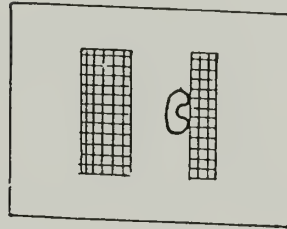
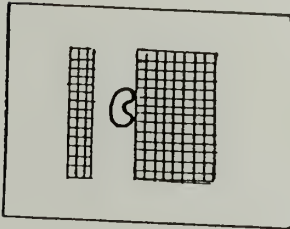
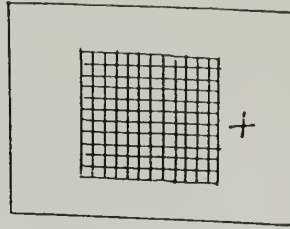
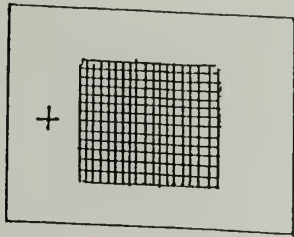
Objects

	Parts only	Parts + Interface	Pieces + Interface	Mean
Correct Location	21%	27%	18%	22%
Incorrect Location	24%	43%	26%	31%
Mean	23%	35%	22%	

Non-Objects

	Parts only	Parts + Interface	Pieces + Interface	Mean
Correct Location	17%	31%	22%	23%
Incorrect Location	17%	21%	19%	19%
Mean	17%	26%	20%	

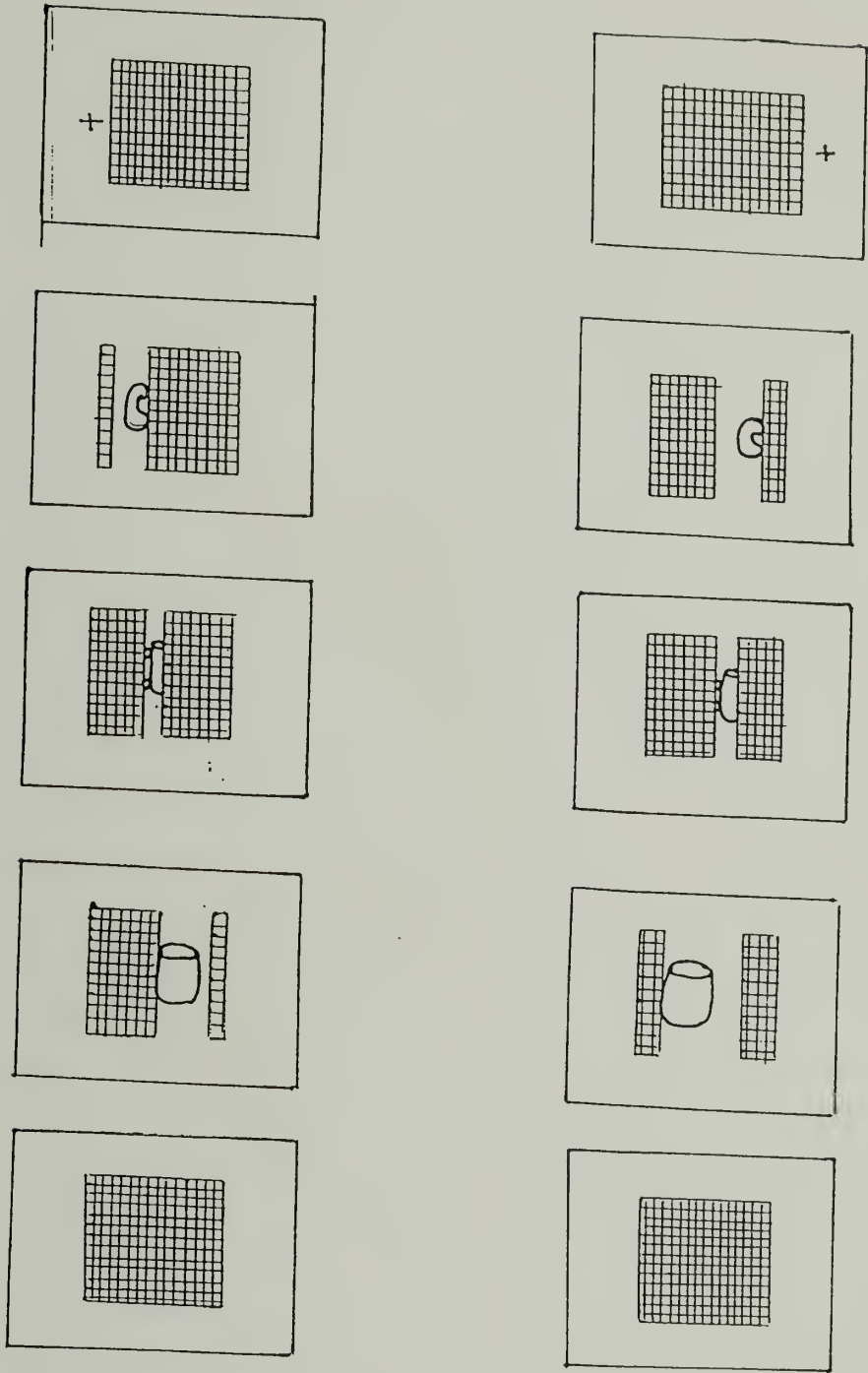
Table 6. Error Rates by Object/Non-Object and Presentation Condition in Experiment 2.



Correct Location

Incorrect Location

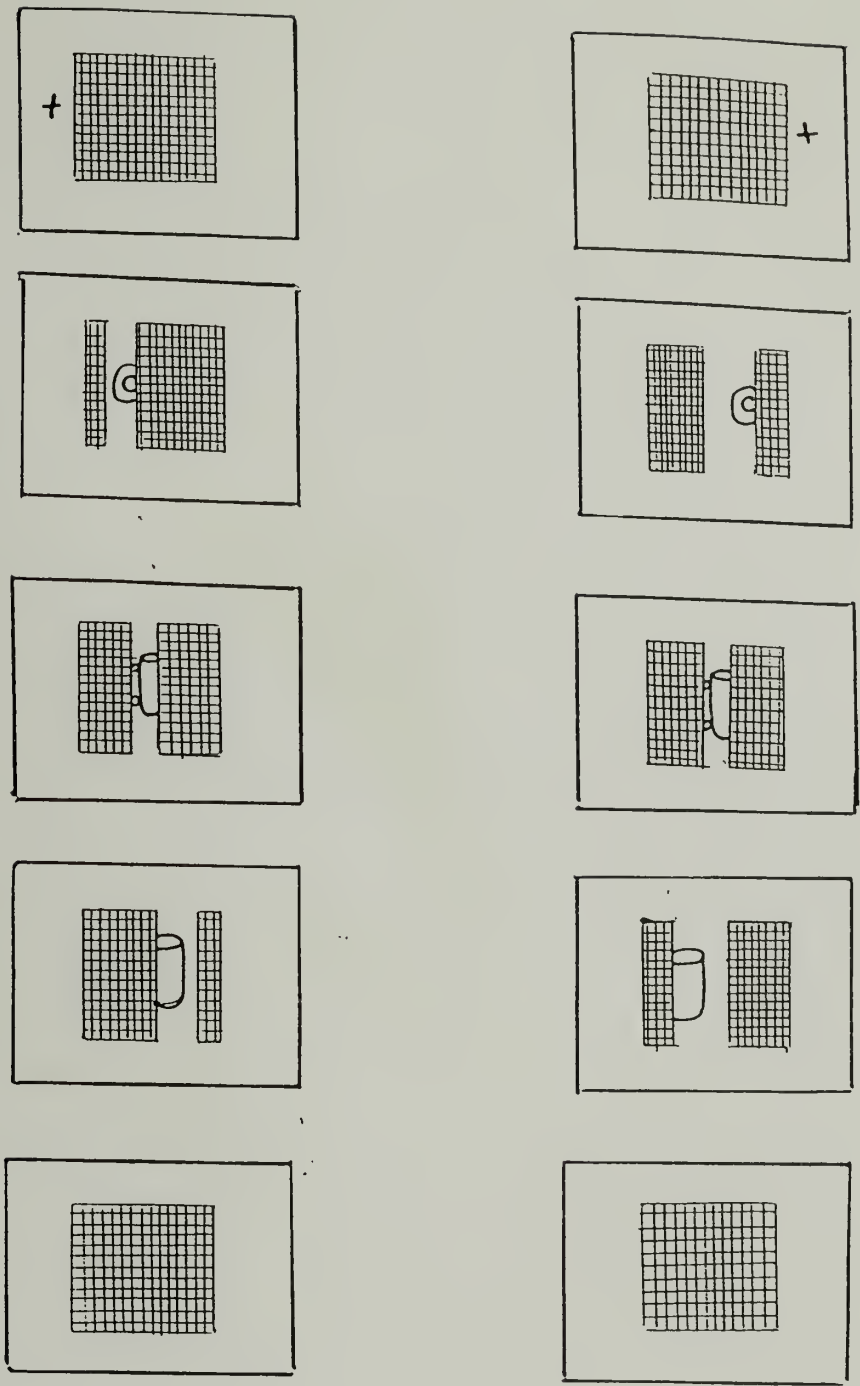
Figure 1. Parts Only in Correct versus Incorrect Location.



Correct Location

Incorrect Location

Figure 2. Parts + Interface in Correct versus Incorrect Location.



Correct Location

Incorrect Location

Figure 3. Pieces + Interface in Correct versus Incorrect Location.

BIBLIOGRAPHY

- Biederman, I. (1987). Recognition by components: A theory of human image understanding. *Psychological Review*, 94, 2, 115-147.
- Biederman, I. (1981). On the semantics of a glance at a scene. In M. Kubovy & J.R. Pomerantz (Eds.), *Perceptual Organization* (213-253). Hillsdale, NJ: Erlbaum.
- Biederman, I & Cooper, E. E. (1991). Priming contour-deleted images: Evidence for intermediate representations in visual object recognition. *Cognitive Psychology*, 23, 393-419.
- Biederman, I. & Cooper, E. E. (1992). Size invariance in visual object priming. *Journal of Experimental Psychology: Human Perception and Performance*, 18, 1, 121-133.
- Biederman, I., Mezzanotte, R.J., & Rabinowitz, J. C. (1982). Scene perception: Detecting and judging objects undergoing violation. *Cognitive Psychology*, 14, 143-177.
- Jolicoeur, P (1990). Orientation congruency effects on the identification of disoriented shapes. *Journal of Experimental Psychology : Human Perception and Performance*, 2, 351-364.
- Kroll, J. F., & Potter, M. C. (1984). Recognizing words, pictures and concepts: A comparison of lexical object and reality decisions. *Journal of Verbal Learning and Verbal Behavior*, 3, 39-66.
- Nelson, W.W., & Loftus, G.R. (1980). Functional visual field during picture viewing. *Journal of Experimental Psychology: Human Perception and Performance*, 6, 391-399.
- Pollatsek, A., Rayner, K. & Collins, W. (1984). Integrating pictorial information across eye movements. *Journal of Experimental Psychology: Human Perception and Performance*, 16, 199-210.
- Pollatsek, A., Rayner, K. & Henderson, J. M. (1990). Role of spatial location in integration of pictorial information across saccades. *Journal of Experimental Psychology: Human Perception and Performance*, 16, 199-210.
- Trehub, A (1991). *The Cognitive Brain*. A Bradford Book. MIT Press.

