THE PSYCHOLOGY OF OBJECT AND PATTERN RECOGNITION: A BRIEF INTRODUCTION

KEVIN BREWER
ELIZABETH WILLIAMS

ISBN: 978-1-904542-16-2

Orsett Psychological Services,
PO Box 179,
Grays,
Essex
RM16 3EW
UK

orsettpsychologicalservices@phonecoop.coop

Contents

Introduction

When information from the environment reaches the eye, it is a stream of light, shapes, and colours. Pattern or object recognition is the process by which the brain recognises the light, shapes and colours as particular objects or patterns. It is the process of "assigning meaning to the visual input by identifying the objects in the visual field" (Eysenck 1984). This ability combines perception, attention, and memory.

Pattern or object recognition involves a number of abilities:

i) the recognition of familiar patterns quickly and accurately;

ii) how the process works on unfamiliar objects;

iii) the accurate perception of objects at different angles and views;

iv) the ability to identify partly hidden objects;

v) the ease and automaticity of the process.

Pattern or object recognition usually occurs without conscious thought, and is instantaneous. Only in difficult viewing conditions, like darkness, does this process of pattern or object recognition need conscious attention and thought. The task for psychologists is to explain how pattern recognition occurs.

There are a number of theoretical explanations:

a) Template Matching Hypothesis;
b) Prototype Theories;
c) Feature Detection Theories.

Template Matching Hypothesis

The Template Matching Hypothesis (TMH) sees pattern or object recognition as occurring at the level of the whole object or pattern. The object is perceived as a whole, and a match is made from the templates (internal representations) in the memory (figure 1).

This is the process of template matching used by many computer recognition programmes, and is based on memory capacity. But most importantly, there must be a perfect match between the object seen and the template.

The individual will build up a store of templates for every object or pattern seen in their lives. The time taken in pattern recognition depends upon how close the pattern or object is to the template.

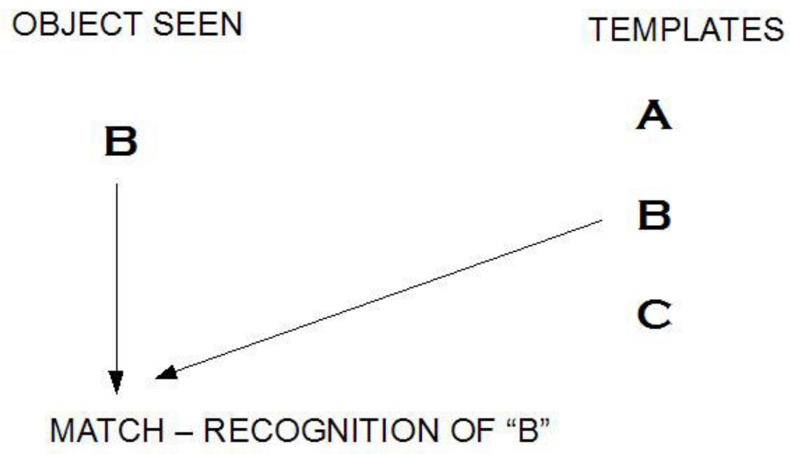

Figure 1 - Template matching process.

EVIDENCE FOR TEMPLATE MATCHING HYPOTHESIS

Shepherd and Matzler (1971; quoted in Solso 1991) used the time taken to recognise an object as their experimental measure (dependent variable). Participants were given a pair of pictures to look at, and asked if the pictures were different views of the same object (figure 2). The reaction time to answer was recorded.

The reaction time was affected by how much "mental rotation" was required to match the different views. For example, if the two pictures were of the same view of the object, the average reaction time was 1.1 seconds. But if one picture was at 180 degrees (ie: upside down), then the reaction time was 4.5 seconds on average.

The time taken to answer was based upon finding the correct template in the memory, and adjusting the rotation to fit the picture shown.

Thus common views of an object are recognised quicker than unusual views because the common view is the template view.

The TMH seems to be best for explaining the recognition of simple stimuli like letters and numbers.

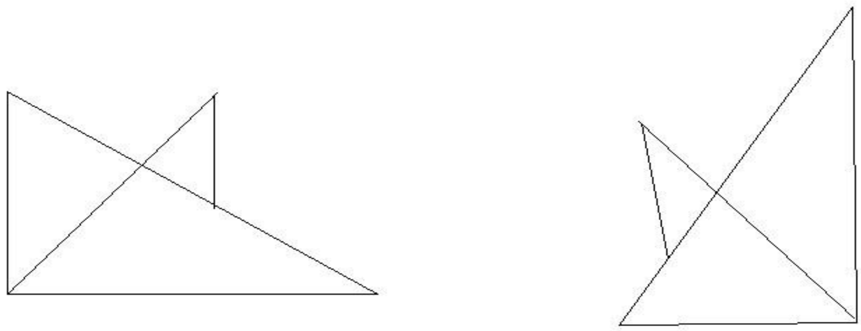

Figure 2 - Example of pair of pictures as used by Shepherd and Matzler.

PROBLEMS WITH TEMPLATE MATCHING HYPOTHESIS

The idea of template matching has a number of problems.

1. Memory capacity problems

If a perfect template match to the object or pattern is required, then this means a massive store of templates in the memory. For example, the "A" can be written in many different ways (eg: a, *a*, a, a), there would have to be a template for each variation, and this would mean a lot of unnecessary information stored (hundreds of templates of different types of "A")

Larsen and Bundesen (1992) asked participants to write "1" to "10" in their own handwriting, and then tested a machine recognition system for recognition of the numbers. Where the machine had one standard template for each number, recognition was poor. But when there were sixty different templates for each number, the machine recognised about 90% of the numbers. Applying such an idea to the human brain would mean an incredible number of templates.

The research here is an example of computational modelling, which is used in cognitive psychology. It involves the use of computers to understand human cognitive processes.

It would also be time-consuming and a waste of energy for the brain to search all the templates every time we looked at an object or pattern.

2. Unfamiliar patterns

The TMH cannot explain how we recognise unfamiliar patterns or objects never seen before, and for which we have no templates.

Also the TMH has difficulty explaining objects or patterns that are different to the original template - for example, in rotation, size, or colour.

3. Conflicting interpretations

Two individuals can see the same object or pattern in different ways. For example, with the Necker cube illusion (figure 3), this can be seen as a cube inwards or outwards. How does this happen? Are there different templates for each version of the cube? The TMH finds it hard to explain such an illusion.

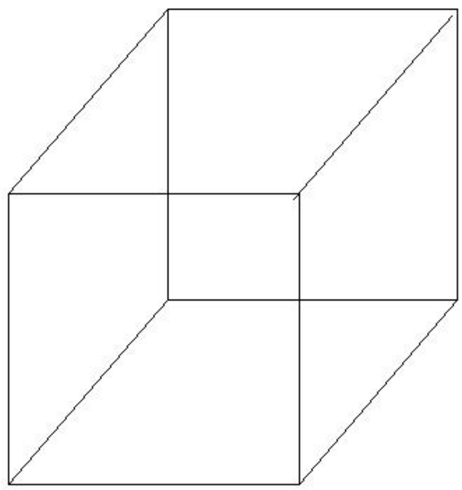

Figure 3 - The Necker Cube illusion.

MODIFICATIONS TO TEMPLATE MATCHING HYPOTHESIS

Eysenck and Keane (2000) have suggested that the match between what is seen and the template need not be perfect, but quite similar a match. For example, the letter "a" is "normalised" (adjusted to a standard) before searching for the "a" template.

The context of the pattern or object becomes important also. Research has shown that individuals are quicker at recognising simple pictures if they are presented in a relevant context (eg: pictures of food after the pictures of a restaurant), then pictures presented out of context.

While Larsen and Bundesen (1996) have suggested that features of the object or pattern are analysed first, then combined, and compared to the templates. In their computer models, thirty-seven templates per item were still needed for accurate recognition. This theory can be seen as an attempt to combine the TMH and Feature Detection Theories.

Prototype Theories

Eysenck (1984) defines prototypes as "abstract forms representing the basic elements of a set of stimuli". Whereas a template is an exact representation of an object or pattern, prototypes are general categories. For example, the recognition of an aeroplane is based upon the prototypes of a long tube with two wings attached.

Also prototypes allow the sharing of general features among objects and patterns, and requires less memory space.

Pattern or object recognition is the process of matching the stimulus with the correct prototypes. These theories can explain the recognition of objects never seen before, which can be compared to similar prototypes. This process is more flexible than template-matching.

One version of the prototype theories is the "geon theory" or "recognition-by-component" (Biederman 1987). Information is stored as "geons" (3D geometrical icons or shapes), and recognition is by the combination of "geons". Thus the emphasis is upon the components that make-up an object or pattern, and are shared between them; eg: "cylinder" shapes.

This theory is based upon the observations of how people describe objects. They divide objects into components, like "blocks" and "funnels", known as volumetric concepts (figure 4).

MUG = CYLINDER + SEMI-CIRCLE

SUITCASE = SQUARE + SEMI-CIRCLE

PAINT TIN = CYLINDER + SEMI-CIRCLE

Figure 4 - Examples of volumetric concepts in object recognition.

This theory has been tested experimentally by the use of "recognisable" and "non-recognisable" versions of objects. The time to recognise objects depends upon the information available in "degraded" versions of drawings of objects (Biederman 1987).

In figure 5, drawing B can be recognised as quickly as A because there is enough information about the "geons", but not in drawing C.

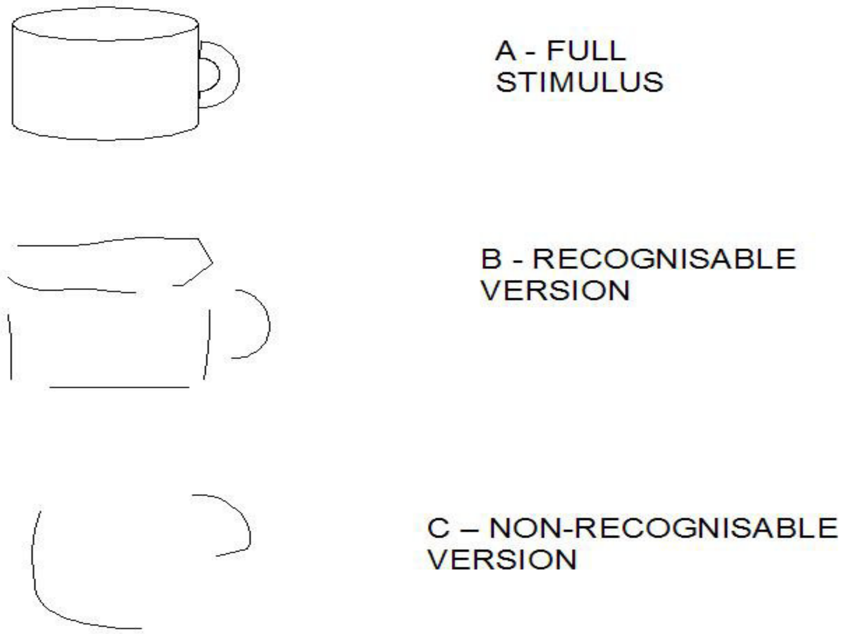

Figure 5 - Drawings as used in experiments by Biederman (1987).

Biederman believed that all recognition is based upon thirty-six different "geons". But there is no research evidence to confirm this number.

Prototype theories focus upon the perception of components of the stimulus and then the combination of those components. So it is a middle point between the TMH and the Feature Detection theories.

PROBLEMS WITH PROTOTYPE THEORIES

There are two main problems with the prototype theories for explaining pattern recognition:

i) What are the properties or components of objects and patterns shared in the prototypes?

In other words, the exact nature of prototypes is not specified.

ii) It ignores the context of pattern recognition, and assumes the same process occurs in all cases.

RECENT MODIFICATIONS

McClelland and Rumelhart (1985) proposed the connectionist model of cognition which has been used in computer modelling of the brain. Cognition occurs as a result of a pattern of activity among different processing units (or nodes) (figure 6). The particular pattern for recognition of an object or pattern is stored in the brain or computer. The pattern of response of processing units is similar to a prototype. Within the response, there is also the opportunity to store specific and general information.

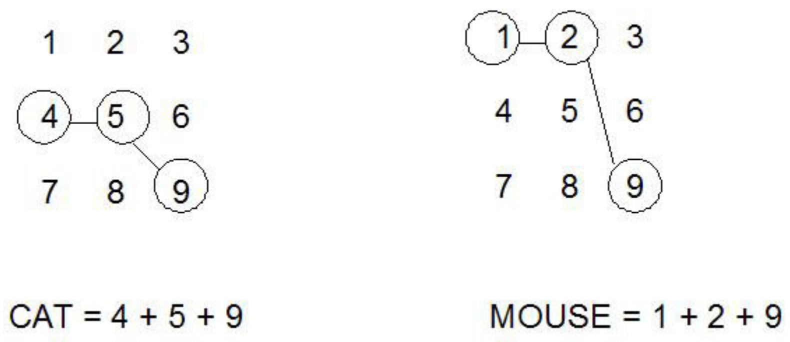

Figure 6 - Examples of pattern of nodes activated in connectionist model of cognition.

Feature Detection Theories

This approach to pattern recognition focuses upon the individual features of the object or pattern, and these are combined to produce the full object or pattern seen (figure 7). There are stages in this process (figure 8).

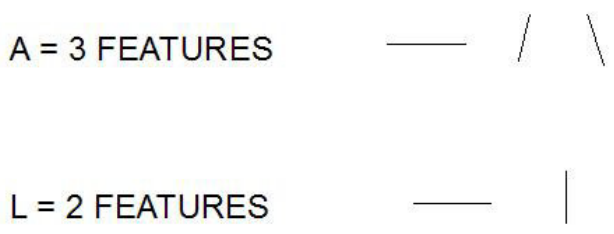

Figure 7 - Number of features of two letters.

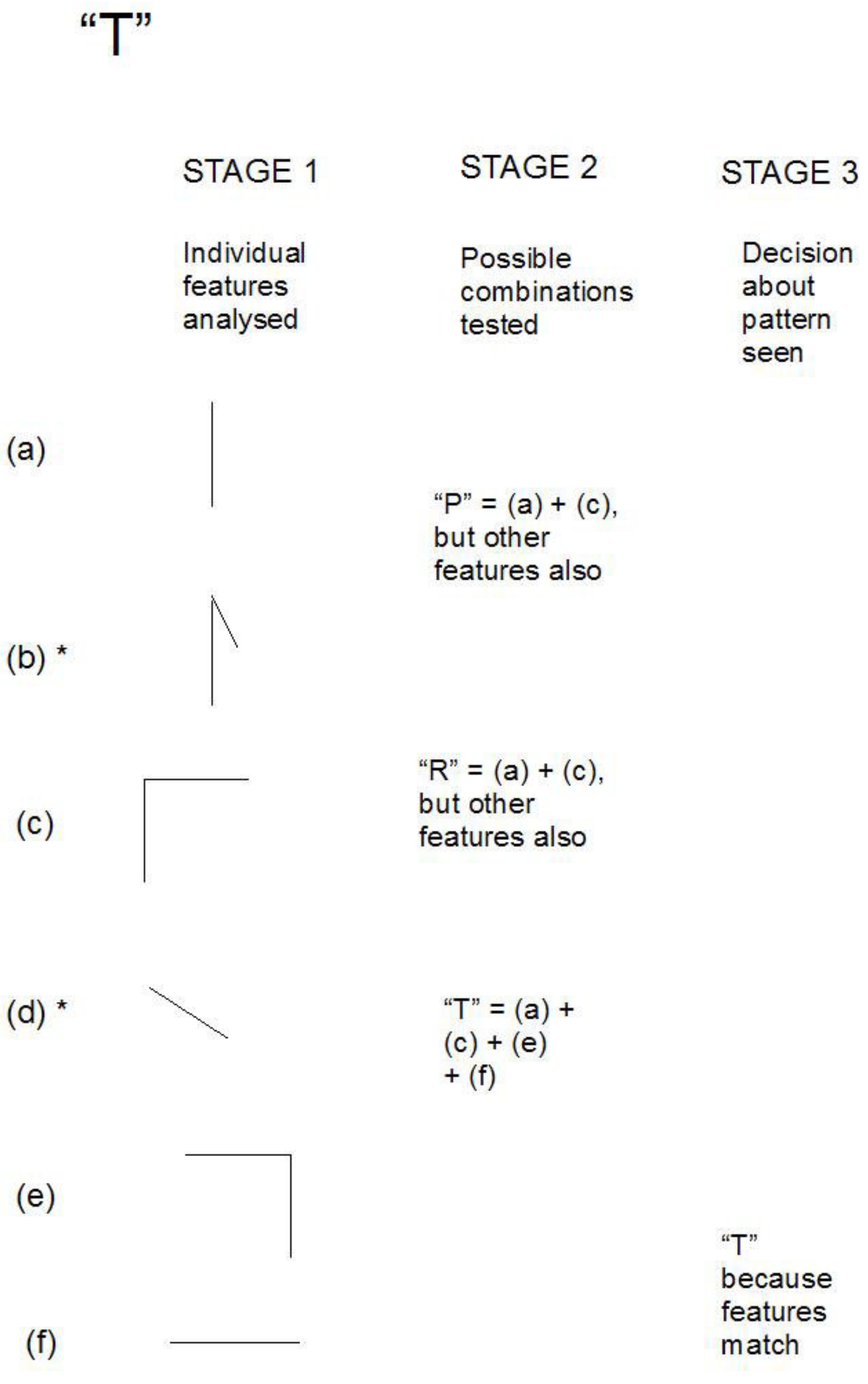

Figure 8 - Stages in feature detection of letter "T".

There are a number of principles for selecting a set of features for capital letters:

i) the features should be present in some members of the set and not others;

ii) the features should remain unchanged with variations in size or perspective;

iii) features should yield unique pattern for each letter;

iv) there should be a small number of features used.

EVIDENCE FOR FEATURE DETECTION THEORIES

1. Letter Detection Task

Neisser (1967) showed that participants were slower in finding a single letter hidden among other letters when there were similar features in the letters. For example, it is hard to find "Z" among "M", "X", and "E" than among "C", "G", and "O". The first group of letters have similar features, primarily straight lines which make pattern recognition more difficult (and thus slower) (figure 9).

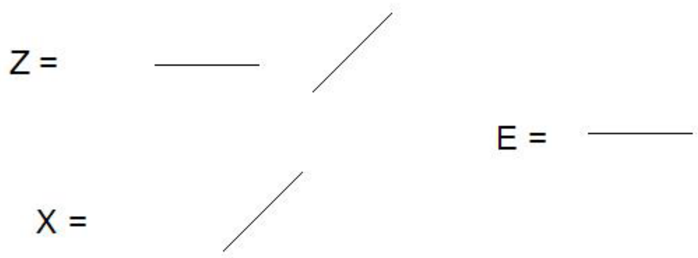

Figure 9 - Features shared by "Z", "E", and "X".

Likewise, it is harder (and takes longer) to distinguish "P" and "R" (similar features) than "G" and "W".

2. Physiological Studies

Extensive work over 20 years by Hubel and Wiesel (1979) using cats, primarily, have identified cells in the visual cortex that respond to particular features of a stimulus only. The cats were paralysed by anaesthetic, but remained conscious. The researchers used minute micro-electrodes to measure the electrical activity of individual brain cells at the back of the surface of the brain.

Lines of different angles and orientations were shown on a screen in front of the cat's eyes. Painstakingly, the researchers measured the response of individual cells, and built up a picture of how cells in the visual cortex work.

Hubel and Wiesel (1959) identified three types of cells in the visual cortex:

i) "simple cells" - these cells respond to particular features of the line only (eg: horizontal), and in particular locations of the visual field;

ii) "complex cells" - these cells respond to particular orientations also, and receive information from the simple cells;

iii) "hypercomplex cells" - these cells are also sensitive to the length of the line, and receive information from the complex cells.

The information from each cell is processed in an upward direction (ie: from simple to hypercomplex). Working downwards through the cortex, the researchers found that the cells were stacked in "ocular dominance columns" (figure 10).

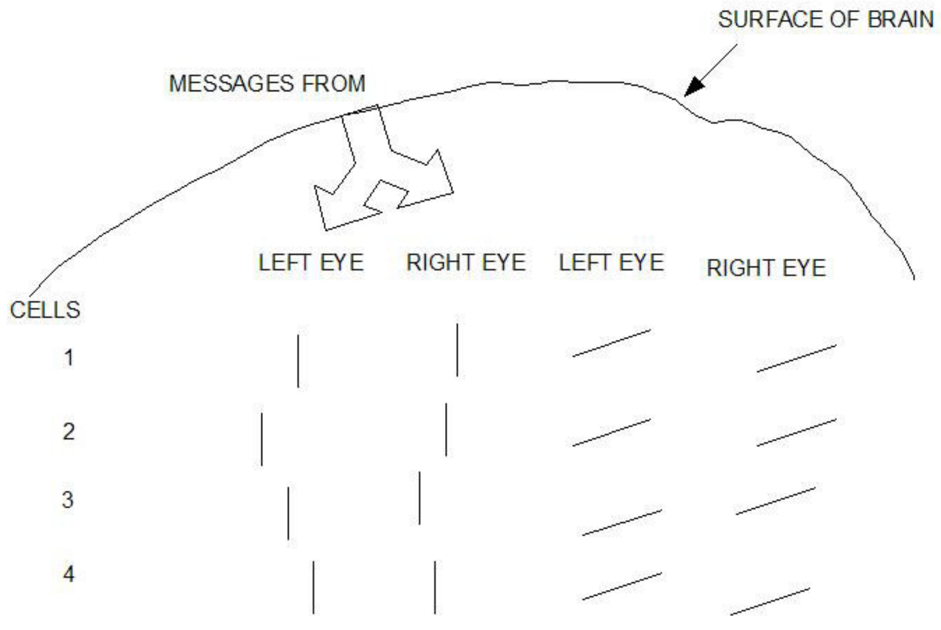

Figure 10 - Example of cells stacked in "ocular dominance columns".

Hubel and Wiesel's work is with cats. The question of how appropriate this is to apply to humans is important. The question of animal ethics can be used as evaluation of this research.

PROBLEMS WITH FEATURE DETECTION THEORIES

1. Role of context

Concentration upon the individual features of a stimulus ignores variables, like context and expectations. The same features can produce different patterns, and different features can produce the same pattern depending on the context.

Palmer (1975) showed participants drawings of ordinary objects for a brief period of time. Recognition of the drawings was aided by the prior presentation of the appropriate context. For example, participants found it easier to recognise a pan after a picture of a kitchen rather than after a picture of a gym. Feature Detection Theories would argue that recognition time would be the same in both situations.

Also it is easier to detect a line if hidden in a 3D drawing than in an incoherent pattern. This is known as the "object-superiority effect" (Weisstein and Harris 1974). Feature Detection should work the same irrelevant of the context.

2. Role of expectations

Pattern recognition depends upon the expectations of what the pattern is likely to be. This is an example of "perceptual set" - "a perceptual bias or predisposition or readiness to perceive particular features of a stimuli" (Allport 1955).

In a famous experiment, Bruner and Minturn (1955) showed participants a list of numbers or letters quickly using a T-scope (tachistascope). Then the participants were shown an ambiguous figure (13). This was perceived as 13 or B depending on the list shown beforehand (figure 11).

In other words, those individuals seeing a list of numbers perceived this pattern as 13. According to Feature Detection Theories, pattern recognition would be the same irrelevant of the expectations.

11

12

A C 13

Figure 11 - Ambiguous figure as used by Bruner and Minturn.

3. How the features are combined

Different features will be important in different situations, and past knowledge helps the individual to know when certain features are important and others not (figure 12).

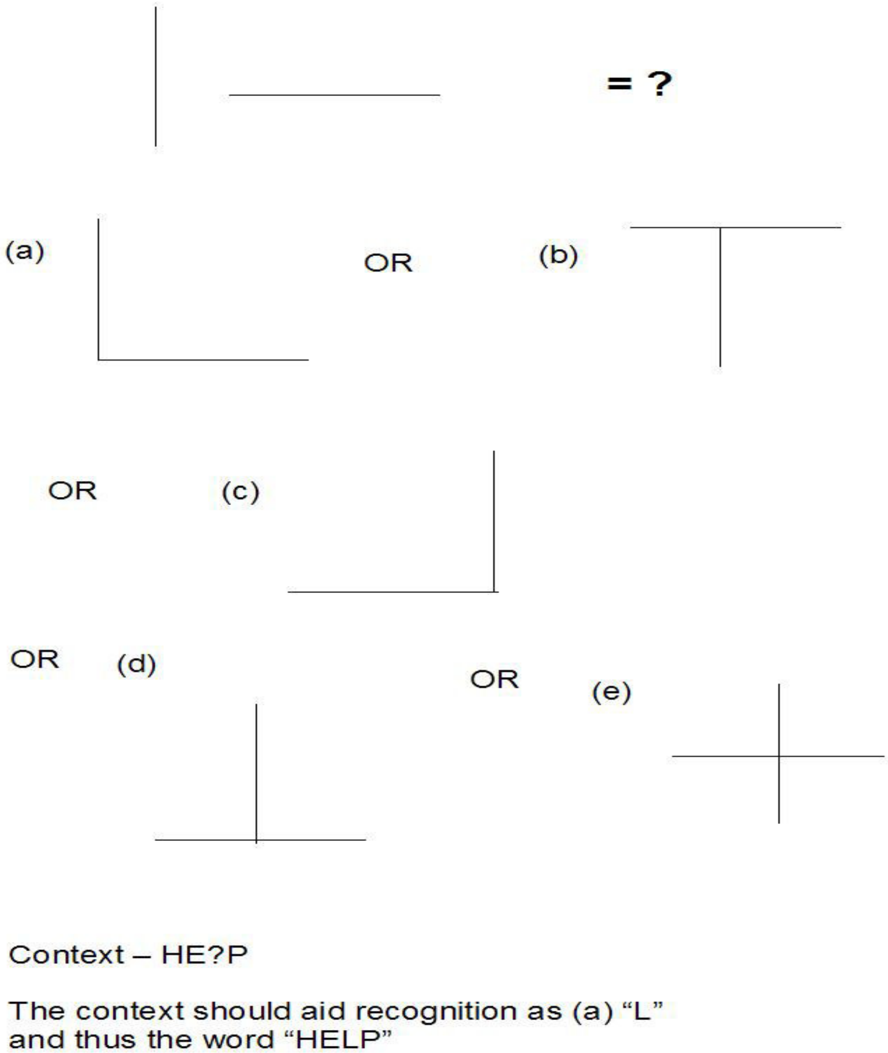

Figure 12 - Importance of context in pattern recognition.

Where the features are ambiguous or insufficient information available for recognition, stored knowledge must guide the recognition process

4. Gestalt approach

The Gestalt approach to pattern recognition sees the process based upon the whole stimuli not the individual parts. This approach takes account of the relationship between the features in perception of the whole; ie: "L" = more than but relationship between them as whole.

Conclusion

The Feature Detection Theories are classed as "bottom-up" models because they use the stimuli only. Whereas "top-down" models make use of extra information (like context or past knowledge) as well as the stimulus in pattern recognition.

The Template Matching Hypothesis and Feature Detection Theories are in opposition. The evidence that supports one can be used as criticism of the other.

Table 1 compares the main strength and weakness of each approach to pattern or object recognition.

	STRENGTH	WEAKNESS
TEMPLATE MATCHING HYPOTHESIS	exact match between object and template	large amount of memory space needed for all templates
PROTOTYPE THEORIES	flexible: focuses on general characteristics	what is exact nature of shared properties between prototypes
FEATURE DETECTION THEORIES	explains pattern recognition by features and stages	ignores context and different perception of same features in different situations

Table 1 - Main strength and weakness of three approaches to pattern or object recognition.

References

Allport, G (1955) Becoming - Basic Considerations for a Psychology of Personality, New Haven, CT: Yale University Press

Biederman, I (1987) Recognition-by-components: a theory of human image understanding, Psychological Review, 94, 115-147

Bruner, J & Minturn, A (1955) Perceptual identification and perceptual organisation, Journal of General Psychology, 53, 21-28

Eysenck, M (1984) A Handbook of Cognitive Psychology, London: Lawrence Erlbaum Associates

Eysenck, M & Keane, M (2000) Cognitive Psychology: A Student's Handbook (4th ed), Hove: Psychology Press

Hubel, D & Wiesel, T (1959) Receptive fields of single neurons in the cat's striate cortex, Journal of Physiology, 195, 215-243

Hubel, D & Wiesel, T (1979) Brain mechanisms of vision, Scientific American, 249, 150-162

Larsen, A & Bundesen, C (1992) The efficiency of holistic template matching in recognition of unconstrained handwritten digits, Psychological Research, 54, 187-193

Larsen, A & Bundesen, C (1996) A template-matching pandemonium recognises unconstrained handwritten characters with high accuracy, Memory and Cognition, 24, 136-143

McClelland, J & Rumelhart, D (1985) Distributed memory and the response of general and specific information, Journal of Experimental Psychology: General, 114, 159-188

Neisser, U (1967) Cognitive Psychology, New York: Appleton-Century-Crofts

Palmer, S (1975) The effects of contextual scenes on identification of objects, Memory and Cognition, 3, 519-526

Solso, R.L (1991) Cognitive Psychology, London: Allyn & Bacon

Weisstein, N & Harris, C (1974) Visual detection of line segments: an object-superiority effect, Science, 186, 752-755

Reflective Essay

ROLE OF FEATURE DETECTION APPROACHES IN PERCEPTION AND RECOGNITION OF OBJECTS

Feature detection theories emerged around the 1970s as one set of theories to explain how human beings perceive and recognise objects. Greene (1990) defined feature detection:

> ..the basis of feature detection is that objects are recognised by extracting cues about features from patterns of sensory inputs.

Investigating this kind of theory can be problematic, because as human beings, we are constantly assessing our environment; a process so natural it becomes hard to take it apart for analysis. One role of this kind of theory is to offer a model to explain how objects are processed, from perception as rays of light, to recognition as whole objects in the brain. Feature detection, from the definition above, attempts to explain object recognition in terms of an elimination process of individual features.

One theory of feature detection, "Pandemonium", will be examined to explain and elaborate upon these kinds of theory and their role. Several others, which incorporate the role of feature detection as part of the process of object recognition will be considered in an attempt to evaluate the theory. The crucial point of context and a brief look at perception theories that totally deny the role of feature detection can be used as comparison.

It would seem pertinent at this stage to separate perception and recognition. In perceiving the world, humans take in events and objects, via sensory organs. Sensory organs are responsible for receiving sensory inputs, which are integrated and passed on to the brain. These sensory inputs provide the "cues" in the definition above.

In this case, the sensory inputs are light, received by the eye. The rays of light and shade pass on to the retina and stimulate layers of cells to transmit the information to the visual cortex of the brain for recognition. Recognition occurs as each object is checked against those held as mental representations in the brain.

These ideas form the basic assumptions of the feature detection theory - that matching of perceived objects and stored objects occurs, based upon assimilated evidence from sensory cues.

The theory is easier to explain in terms of an example. "Pandemonium" was originally a computer programme, designed by Selfridge (1959), to recognise the individual dots as features in Morse code. Lindsay and Norman (1972) reworked the idea to form the model of feature detection in object recognition.

Letters of the alphabet (recognisable objects already stored in the brain) are matched, in sequence, according to the features of each letter. In each line of the sequence "demons" "shout" if their feature is contained within the letter perceived, from the "perception demon" to the "decision demon". At each shouting level, features are matched, in progression downwards from lines and angles to patterns. At the final elimination the "decision demon" decides and the letter is recognised.

As a theory, this would seem to support and echo the physiological account. Rays of light passing through different layers of cell and neurons reach the visual cortex in the brain as researched, famously, by Hubel and Wiesel (1959), where "feature detection cells" were specifically noted. However, cells do exist, whilst the "demons" are just clever metaphors to explain these "bottom up" theories. In this way, Lindsey and Norman's theory can be supported by physiological findings and potentially tested as such. Both can also be described in terms of a hierarchical "bottom up" method, as information perceived is passed from lower to higher levels.

"Pandemonium" shows information as passing through line, angle, and pattern "demons" on up to "decision demon", the drive upward with the information data.

The conventional reverse, known as "top down" processing, relies upon concept driven processes, moving downward from previous stored information to match objects. Sensory information (ie: perceiving) alone is not enough to actually recognise objects. But feature detection theories always assume that matching will take place - it is whether the matching occurs "top down" or "bottom up".

A theory such as Neisser's cyclic model (1976) incorporates both "top down" and "bottom up" processing, in his "analysis by synthesis" approach. Feature detection is included as part of this cyclical process, where perception occurs at the analysis stage and matching (recognition), at the synthesis, "top down" stage in an ongoing, active cycle. A perceptual model is hence formed and tested to become a perception, recognised. This seems to bridge the gap left in relying upon feature detection alone to recognise whole objects -

the perceptual model idea could account for the leap from analysis of features to recognition of an object.

It also offers the idea of active integration with objects in the context of the environment whereas feature detection appears to ignore the crucial issue of context and focuses singularly upon the object.

Another criticism of both feature detection and the cyclical model is that perception is indirect - it goes through a process, from perception to recognition - although this appears to be automatic. Treisman and Gelade's (1980) work offered the idea that attention was required to assimilate two or more features and hence the process would take even longer.

Gibson's (1979) theory of direct perception totally denied the idea of a processing stage, stressed that light stimulation alone was enough for interacting with the environment.

So, feature detection would appear to have a role in the explanation of object recognition, both as a lone theory in line with physiological research and integrated into cyclical theories, such as Neisser's.

However, criticisms linked to the gap between analysis and recognition, and contextual concerns would suggest that feature detection as a model offers a limited view of how object recognition occurs.

REFERENCES

Gibson, J.J (1979) The Ecological Approach to Visual Perception, Boston: Houghton Muffin

Greene, J (1990) Perception. In Roth, I (ed) Introduction to Psychology, Milton Keynes: Open University

Hubel, D.H & Wiesel, T.N (1959) Receptive fields of single neurons in the cat's striate cortex, Journal of Physiology, 148, 579-591

Lindsay, P.H & Norman, D.A (1972) Human Information Processing, New York: Academic Press

Neisser, U (1976) Cognition and Reality, San Francisco: WH Freeman

Selfridge, O.G (1959) Pandemonium: a paradigm for learning. In The Mechanisation of Thought Processes, London: HMSO

Treisman, A.M & Gelade, G (1980) A feature integration theory of selection, Cognitive Psychology, 12, 97-136