

THE ~~VISUAL~~ PERCEPTION OF FACIAL IMAGES

A thesis submitted for the degree of
Doctor of Philosophy in the Faculty
of Engineering, University of London

by

Marc Jean-Pierre Garneau

Imperial College, University of London

1973

"The face is what one goes by generally", Alice remarked in a thoughtful tone.

"That's just what I complain of", said Humpty Dumpty. "Your face is the same as everybody's; the two eyes - so (marking their places in the air with his thumb) - the nose in the middle, mouth under. It's always the same. Now if you had the two eyes on the same side of the nose for instance - or the mouth at the top - that would be some help".

Alice Through the Looking Glass
Lewis Carroll.

ABSTRACT

The research presented in this thesis is concerned with the relative importance of different facial features for recognition. Different types of stimuli are examined and these are subjected to different modes of visual presentation.

Facial features are classified into three categories: contours, discrete features and feature relations. Contours consist of the forehead-hairline boundary and the general facial outline. Discrete features include the eyes, nose and mouth. Finally, feature relations consist of the separations between discrete features and contours. Stimuli consist of full-frontal presentations of male adult caucasian faces. These appear in several formats: as monochrome slides presented on a television monitor, as Photofit composites or as line drawings. Stimuli presented by means of video equipment are altered by either visual noise, masking or tachistoscopic presentation.

The results from seven experiments are reported. These indicate that contours are the most important category of facial feature in recognition, especially those in the upper half of the face. Amongst the discrete features, the eyes are found to be more important than either the nose or mouth.

In judging the effects on similarity produced by altering feature relations, it is found that certain changes produce greater effects than others. Furthermore, it is suggested that changes in certain feature relations produce greater effects than changes in certain discrete features. However the effects produced by these changes vary according to the nature of the stimulus being tested, i.e. line drawing or Photofit face.

Finally, experimental results are briefly discussed in terms of their relevance to criminal identification.

CONTENTS

<u>Item</u>	<u>Page</u>
ABSTRACT	3
CONTENTS	4
ACKNOWLEDGEMENTS	7
CHAPTER I. INTRODUCTION	8
1.1 The human face as a stimulus	9
1.2 The nature of the stimulus	12
1.3 The author's field of study	13
1.4 Contribution of this thesis	15
1.5 Summary of later chapters	16
CHAPTER II. VISUAL PATTERN RECOGNITION	18
2.1 The evidence for simple perceptual elements	19
2.2 Neurophysiological evidence	25
2.3 Some models for pattern recognition	33
2.4 Summary	40
CHAPTER III. THE EXPERIMENTAL VARIABLES	42
3.1 General variables in a recognition task	43
3.1.1 Personal experience variables	43
3.1.2 Stimulus variables	46
3.1.3 Experimental variables	48
3.1.4 Summary	50
3.2 Facial attributes	51
3.3 The stimuli and the subjects	55
3.3.1 The stimulus faces	55
3.3.2 Photofit faces	55
3.3.3 The subjects	58
3.4 Preview of the experiments	58
CHAPTER IV. THE VIDEO EQUIPMENT	60
4.1 The general layout	61
4.2 Visual noise	64
4.3 The masking unit	68
4.4 The frame counter	68

CHAPTER V. NOISE AND MASKING IN RECOGNITION	70
5.1 The recognition of facial images in noise	71
5.1.1 Introduction	71
5.1.2 Results	74
5.1.3 Discussion	78
5.2 The first masking experiment	81
5.2.1 Introduction	81
5.2.2 Results	83
5.2.3 Discussion	85
5.3 The second masking experiment	86
5.3.1 Introduction	86
5.3.2 Results	90
5.3.3 Discussion	91
5.4 Summary of the first three experiments	92
CHAPTER VI. TWO EXPERIMENTS USING LINE DRAWINGS AND PHOTOFIT	95
6.1 An experiment using line drawings	96
6.1.1 Introduction	96
6.1.2 Results	101
6.1.3 Discussion	107
6.1.4 Summary	108
6.2 The first experiment using Photofit	110
6.2.1 Introduction	110
6.2.2 Results	113
6.2.3 Discussion	115
6.2.4 Summary	117
6.3 Summary of the chapter	118
CHAPTER VII. TWO MORE EXPERIMENTS USING PHOTOFIT	119
7.1 A test of similarity for Photofit components	120
7.1.1 Introduction	120
7.1.2 Results	121
7.1.3 Discussion	123
7.2 A test on the additivity of component effects	123
7.2.2 Results	128
7.3 Summary of the chapter	142
CHAPTER VIII. SUMMARY OF THE RESULTS	144

APPENDIX A. Circuit diagrams	152
APPENDIX B. Two experiments on facial expressions	156
APPENDIX C. A test case on biasing reference photographs	172
REFERENCES	174

ACKNOWLEDGEMENTS

The author wishes to express his thanks to his supervisor, Professor E.C. Cherry, for continued advice and encouragement throughout the course of this research, and also for procuring funds for the research.

The work was made possible through the generous assistance of the Athlone Fellowships Committee and the National Research Council of Canada who provided the author with a healthy living allowance covering a period of three years. Special thanks are also due to the Canadian Ministry of Defence for allowing the author to take three years leave from his commitments.

John Springate read many of the initial reports and Peter Beevor and Barry Stapley corrected the final manuscripts. The author is indebted to them for their untiring constructive criticism of his work. Special thanks go to Archie White and Barry Stapley for their patience in assisting the author with certain technical difficulties.

Many of the experiments reported in this thesis would not have been possible without the use of a Photofit kit which was generously lent to the author by the Police Research Division of the Home Office. The author owes a special debt of gratitude to Mr. Jacques Penry, the original designer of the Photofit kit.

Alfie Bass made possible two of the experiments reported in this thesis by kindly providing the author with a wide range of his facial expressions.

Finally, the author wishes to thank all those who acted as subjects for his experiments and especially his colleagues who provided initial criticisms.

CHAPTER I

INTRODUCTION

The introduction presents a short review of some of the fields of study in which the human face has been used as a stimulus. It also examines the nature and complexity of the stimulus with respect to the more commonly used perceptual stimuli. This is followed by a short summary of the objectives and extent of the author's field of study. Finally, the chapter ends with a preview of later chapters.

1.1 The human face as a stimulus

Most people would agree that the human face is one of the most common stimuli that we encounter in our daily life. Soon after birth, a child becomes familiar with his mother's face and associates it with security and reward. A little later comes the realisation that the visual environment is filled with a multitude of faces and initially this upsets the child (Ambrose; 1961, 1963). However, the child learns to live with this fact and begins to associate certain characteristics with each of these faces.

Later on in life he finds himself surrounded by many people, some of whom he will see more often than others. He learns to recognise them, not exclusively by their face, but rarely without it. He may subconsciously attribute a certain personality to them by virtue of some of their facial characteristics (Brunswik and Reiter, 1938). He is sensitive to mood and recognises facial expressions. He finds that he can recognise a great number of people after a single encounter and without any apparent effort.

Because of the significance of the human face as a social stimulus, it is not surprising that faces have been studied in several related disciplines. Faces have always been important to us. In sculpture and painting, they have often been the main theme. In literature, the characters have been described not only in terms of their personality but also in terms of their facial characteristics. The face has always been considered the most representative feature of man.

Probably the earliest systematic studies of the face were those dealing with the communication of mood by means of facial expressions. Although Darwin (1873) was not the first to study facial expressions, his work remains as a milestone in that field. He was preoccupied with whether expressions were learnt or were innate in man and with whether different races used common facial expressions to convey the same mood. Since the publication of his results, many research workers have concerned themselves with the role of facial expressions in non-verbal communication (see Vine (1970) for an extensive review).

Another field of research dealing with the human face has been concerned with its physiognomy. More specifically, it has long been held that certain details of the face convey information about personality, (e.g. the strong chin, the broad forehead). Probably, the first large

scale research into this field was that of Brunswik and Reiter (1938), using schematic drawings of faces. Their successors have used more realistic facial images and found that often people are consistent in attributing certain subjective qualities to certain faces (Fisher, 1971).

Yet another approach to the study of physiognomic perception has been that of Penry (1952). Basically, Penry maintained a relation between facial structure, personality and endocrine balance. To the author's knowledge, this work has not been followed up. Another unusual approach has been the study of facial asymmetry. Working on the assumption that the right side of the face reflected the outward public aspect of personality whilst the left half of the face expressed unfulfilled wishes and concealed tendencies, McCurdy (1949) postulated that facial asymmetry was related to neuroticism. Although he could not establish the desired relation, positive correlation was found in a later paper (Linzey et al., 1952).

Although most early research dealing with faces concerned itself primarily with the personality and mood conveyed by the face, recent research using faces has been more concerned with the non-social aspects of the face as a stimulus. Specifically, the face has been chosen as the typical "complex" stimulus. The concept of stimulus complexity will be discussed in the next section. For the moment, it is sufficient to look upon the human face as a complex geometric pattern. Research into pattern recognition has increased a great deal in recent years and the study of "simple" stimuli such as straight lines, angles, squares and triangles has been joined by research into more complex forms. Examples of research using the face as a "complex" stimulus will be given in the next section.

Another area of research in which faces have been used as stimuli deals with infant perception. Here, the face is used partly because of its possible social significance to the child but mostly because it is a common complex stimulus. Fantz (1962, 1963, 1966, 1967) wanted to disprove the idea that the newborn's environment was a "booming buzzing confusion". To do this, he used the face as one of many stimuli and found that infants showed preferential fixation of faces over other stimuli varying in colour, brightness and complexity. To him, this signified at least a capacity to discriminate. Other research in the same area concerned itself more with the social significance of the human face

and the possibility of 'imprinting' (Ambrose, 1961, 1963; Morgan and Ricciuti, 1969; Kagan et al., 1966; Stechler et al., 1966).

A more obvious subject for research into faces has been the field of police identification of criminals. This has led to the use of certain facial identification kits (Identikit, Photofit) which will be described in chapter III. The inadequacies of such kits have pointed strongly to the fact that we know very little about how a face is recognised or remembered. Indeed, the problem of face recognition has only been touched upon indirectly. For example Goldstein et al. (1971, 1972) have devised a man-machine interactive system which can recognise one individual face from a large group on the basis of a few feature descriptors. However, results of this sort, while emphasising the power of the computer, do not really bring us any closer to the problem of recognition. Other studies have compared recognition of positive and negative photographs of faces (Bradshaw and Wallace, 1971; Galper and Hochberg, 1971; Phillips, 1973) or recognition of upright and inverted faces (Bradshaw and Wallace, 1971; Hochberg and Galper, 1967; Scapinello and Yarmey, 1970; Yarmey, 1971; Yin, 1969, 1970) and concluded that inverted and negative faces are more difficult to recognise than upright and positive faces. This apparently obvious conclusion does have implications for pattern recognition and may tell us something about "inversion" and "negativity" but does not attack the problem of face recognition directly. One interesting sideline does deal with the problem of a perceptual deficiency known as prosopagnosia. Those afflicted cannot recognise others on the basis of the face alone. Other clues such as voice or physical gestures are required, (Hecren and Angelergues, 1962; Yin, 1970). Because the deficiency can be traced to a specific cerebral lesion, it has been postulated that a certain area of the brain specialises in faces.

Thus it becomes apparent from this brief introduction that the human face is not a new stimulus for study. But at the same time, the problem of face recognition has not been studied very extensively. Although we are extremely good at recognising faces, we know virtually nothing about the underlying mechanisms.

1.2 The nature of the stimulus

In the previous section, the human face, or more specifically its two-dimensional image projection, was said to be a complex geometrical pattern. This assertion is often made in the literature and is based on the fact that a facial image can be considered as the piecing together of a great number of straight lines, curves, angles, areas and shades. As Gibson (1967) put it: "The face presents a complete and complex array of stimulus properties, including brightness, contrast, shading, contour, shape, colour, texture, motion, transformation, features in the literal sense of the term, the pattern of features, and even the motion of these features that constitutes facial expressions." On the other hand, stimuli such as straight lines, curves and angles have always been considered as simple stimuli. Thus, it would appear that pattern complexity is dependent on the number and variety of simple forms that are combined to make up a complex stimulus. At best, this assumption is accepted for convenience since no serious attempt has ever been made to define a proper measure of complexity.

Attneave (1954) examined the nature of complexity in more detail and viewed perception as an information handling process. He realized that complex patterns are not simply a summation of simple patterns but that other concepts such as redundancy are involved. Redundancy in complex stimuli can best be described in terms of Gestalt concepts such as symmetry. Obviously, an optimum information handling process must be capable of extracting those features which are necessary for discrimination and this assumes a capacity to overlook *irrelevant* information. At the same time, it is not always necessary for the whole stimulus to be perceived before recognition occurs. Pasnak (1971) found that subjects respond to distinctive parts of complex random forms but to the whole contour of simple random forms. Thus, it may be that a complex stimulus can be perceived on the basis of a few distinctive simple features rather than as a complete Gestalt.

Another property of the human face which must be considered is that concerning its familiarity. Faces are not merely complex patterns. They are familiar patterns with high social significance and this fact must be considered in relation to stimulus complexity. Goldstein and Chance (1970) showed that different faces were better recognised than

different snowflakes or ink blots. All three classes of stimuli can be viewed as complex, but faces can be said to be more familiar to us. We have experience with the face class and therefore it might be reasonable to assume that we are capable of better discrimination within that class.

Yarmey (1971) compared recognition memory for familiar "public" faces, unfamiliar human faces, canine faces and architectural forms. He found that familiar faces were better recognised than the other three classes of complex patterns. Thus, the concept of familiarity as a variable applies within the face class itself. Whether architectural forms are of similar complexity to human faces is a difficult problem to resolve regardless of the results obtained by Yarmey. We presumably are capable of differentiating between faces on the basis of distinctive features but our daily experience may be such that we do not spend much time examining the subtle nuances between different architectural forms. Therefore, even if some measure of complexity could be attached to stimuli, recognition scores would still be dependent on our degree of familiarity with the stimuli. At the same time, familiarity does not necessarily imply ease of discrimination within a stimulus class. There must also be sufficient stimulus properties to serve as a basis for discrimination (Gibson, 1967) and a rationale for attending to these stimulus properties.

It can be seen that faces are not only complex stimuli in terms of geometry but are also familiar stimuli. For that reason, it is dangerous to look upon them merely in terms of complexity. Finally, it appears that faces are not difficult to discriminate. In that sense it can be argued that we are really dealing with simple stimuli.

1.3 The author's field of study

In the first section of this chapter, research using the human face as a stimulus was described. At the same time, it was concluded that very little research had dealt specifically with face recognition. Comparative studies involving negative and inverted faces were mentioned as well as research using faces as one of many other classes of stimuli. However, the question of how we recognise a face from a larger group of faces or of the relative importance of different facial features has not been answered. Any progress in those directions could be of considerable

use to the Police (Connolly and McKellar, 1963) and to pattern recognition research.

To begin, it is essential that the process of recognition be clearly defined in the context of the author's research. Using the terminology of Dember (1963), recognition is defined as follows: "The subject must make multiple similarity judgments, comparing the stimulus with some standard on a large number of attributes. When explicit standards are provided with which to make the comparison, the task is referred to as recognition."

The purpose of the present research is to examine the basis on which face recognition occurs. This implies a classification of facial features in terms of their relative importance. It further assumes that a specific set of facial features has been defined. Finally, it concerns, not the problem of recognising a face from other stimuli of a different nature, but rather of recognising individual faces from a much larger group of faces.

Having defined what is meant by recognition, it becomes necessary to define the stimuli. In the present research, the majority of stimuli are adult, male, European faces shown on a black and white television monitor. The faces are presented on slides and each is in the full-frontal position. Any details below the chin boundary were also removed and the faces were "neutral" in expression. Furthermore, none of the faces have beards, mustaches, glasses, scars or unusual hairstyles. This last constraint cannot be defined objectively but is based on personal judgement in order to satisfy a rough criterion of homogeneity. Obviously, the stimuli are lacking in certain properties: three dimensionality, movement, expressiveness, colour and different viewing positions. Nevertheless, a great deal can be learnt from such basic facial images before proceeding to a higher level of experimentation. The faces described above were those of outside students. In addition, some of the faces were prepared from the Photofit kit which will be described in chapter III. These also appeared as slides. Finally, line drawings prepared from projected negatives were also used. These were presented on white sheets of paper. In all cases subjects had never before seen the stimuli.

In the definition of recognition, standards against which to compare the stimuli were mentioned. In the present case the standards were photographs of the stimuli presented on a reference board.

Since a face presented at high resolution for an unlimited period of time will almost certainly be recognised correctly, it becomes necessary to establish certain constraints if we wish to obtain any error data. In the present research blanking of certain portions of the face, visual noise and limited exposure durations were all used as constraints. These tools will be described in chapter IV which deals with the video equipment used for this research.

Another, more subjective approach to recognition has also been used in the present research. It concerns the use of a similarity criterion. This means that subjects were shown faces differing in one or more ways and asked to evaluate them on a similarity basis with respect to a standard reference. Data of this sort must be carefully interpreted but also has merits when used in the proper context.

Most people have their own homespun ideas about how they recognise others. Some say that the mouth is most important, others the eyes, still others the overall shape. Many do not make the important distinction between face and people recognition. In the latter case, recognition occurs not simply by means of the face but by the use of many other attributes, e.g. the voice, the physical shape of the body, the clothes, the movements and a host of other factors. How important a role the face plays in such a case is completely unknown. Returning to the case of face recognition, we find that very little systematic research has been done to establish how it proceeds and whether certain features are more important than others. The present research attempts to do this.

1.4 Contribution of this thesis

The purpose of this section is to indicate precisely the original work that is reported in this thesis.

To the author's knowledge, none of the experiments reported in this thesis, except for those described in Appendix B, have ever been performed *before*. The common aim of each experiment is to arrive at a classification of facial features according to their importance for recognition within the class of faces. Three sets of facial attributes were defined for this purpose: contour attributes, discrete feature attributes and feature relation

attributes (section 3.2). The relative importance of each set of attributes and of members within each set is evaluated in the experiments reported in chapters V, VI and VII.

Stimulus faces are presented in three different forms depending on the particular attributes being investigated: as monochrome slides, as line drawings and as Photofit composites. The type of line drawing used (chapter VI) and the presentation of Photofit composites (chapter III) are completely new as methods of preparing the stimuli.

Video equipment developed by Stapley (1972) is used in the presentation of the monochrome slides. Several options available in the equipment are used in the presentation format; visual noise, electronic masking and frame counting (chapter IV). Although none of these techniques is new, their use in tests involved with face recognition is original.

Finally, the effect of changes in certain facial attributes is investigated according to a criterion of similarity. The possible additivity of these effects is also examined (chapter VII).

1.5 Preview of later chapters

Chapter II presents a general cross-section of the research which has dealt with pattern recognition. The study of simple geometrical forms such as lines, angles, squares and triangles is reviewed. This is followed by a description of some of the recent research in neurophysiology concerned with the operation of the visual pathways. Finally, the last section examines some of the perceptual models which have been postulated to explain the way in which we process visual information.

Chapter III is concerned with a description of the variables which influence recognition. First, a general discussion of these variables is presented and this is followed by a more specific exposition of the variables which are relevant to the author's experiments. The experimental stimuli are defined in terms of three sets of attributes, each to be examined in separate experiments.

Chapter IV describes the video equipment used in the experiments. The generation of visual noise, the use of masking and the purpose of the frame counter are explained. The general operation of the overall system is described from the preparation of slides to the final appearance of the stimulus in the subjects' testing cubicle.

Chapter V reports three experiments concerned with recognition of faces either partially masked or embedded in visual noise. The first experiment is essentially exploratory and helps in defining many of the areas to be investigated in later experiments.. The second and third experiments investigate recognition in conditions where different portions of the face are masked. Contours and discrete features are examined in particular.

Chapter VI describes two unrelated experiments. The first is concerned with the effects of changes in feature relation attributes evaluated on a similarity criterion. Line drawings are used for this purpose. The second experiment makes use of the Photofit kit and examines the importance of certain facial attributes by interchanging Photofit segments from different reference faces. Stimuli are presented for very short periods of time by means of the frame counter.

Chapter VII uses Photofit to investigate the possible additivity of the effects of changes in Photofit segments, once again on a similarity criterion. The first experiment of chapter VI is essentially replicated as part of the larger Photofit test and the results obtained in both tests, in the one case with line drawings, in the other with Photofit, are compared. Finally, a comparison is made between the effects of different types of facial attributes.

Chapter VIII gives a general summary of the separate conclusions to be reached in the previous chapters.

CHAPTER II

VISUAL PATTERN RECOGNITION

This chapter is concerned mainly with pattern recognition. Since facial images are regarded as complex patterns, it is necessary to examine our present knowledge in the field of pattern recognition. The first section reviews research dealing with simple stimuli such as straight lines, curves, angles and simple geometrical forms such as squares, triangles and circles. This is followed by a description of the neurophysiology of the visual system as it is known at present in sub-human species. Section three describes perceptual models which have been postulated on the basis of results from the two previous sections. This is followed by section four which is a summary of the chapter.

2.1 The evidence for simple perceptual elements

One of the classical debates of perceptual psychology has concerned the question of whether a stimulus is initially processed as a whole or broken down into smaller units before recognition takes place. The former explanation was advanced by the Gestalt school and the latter was proposed by Hebb in his cell assembly theory. Both theories are reviewed by Gibson (1967).

The breakdown of the stimulus into smaller units was based on the belief in the existence of specific property analysers in the visual processing system. These react to specific properties of the stimulus. It was assumed that the number of different kinds of property analysers was limited, whereas the number of discriminably different stimuli was potentially infinite (Weisstein, 1969). A great deal of research concentrated on establishing whether stimuli could be broken down into more primitive "building blocks" such as lines and angles. It was believed that these might be processed by specific cells sensitive to lines and angles before the actual recognition process took place. The existence of cells within the visual cortex of cats and monkeys which reacted to specific stimulus properties such as line size and orientation further strengthened the idea that a stimulus was not processed as a whole but as a composite of simpler elements. Neurophysiological findings will be reviewed in the next section.

In the search for basic perceptual elements, various methods have been used and these are reviewed in Weisstein (1969). One of them employed the method of stabilised images proposed by Ditchburn. This consisted in counteracting the natural saccadic movements of the eyeball in order to produce a stable image of the stimulus on the retina. The method is described by Pritchard et al. (1960) in one of the typical experiments using stabilised images. The basic assumption underlying this approach is that the retinal image will eventually disappear following saturation of the retinal receptors. Furthermore, if disappearance occurs by fragmentation of the image, then those parts which disappear as units can be regarded as basic perceptual elements processed by specific cells in the cortex. Pritchard et al. noted the following sequence of events resulting from stimulation of the retina by a stable image. First, the target remained present for a length of time which depended on

its geometrical complexity. This was followed by a disappearance of the target, leaving only the more dimly illuminated background field. This field was then replaced by an intense black patch and finally, the image reappeared.

Using different simple images, they obtained the following results: 'meaningful' targets were seen longer than 'meaningless' ones; straight lines tended to act as units in appearing and disappearing; parallel lines acted in a correlated manner; jagged images were less stable than rounded ones; a good gestalt figure was more likely to act as a complete figure than a poor one; no evidence was found to support the idea that angles or corners might be perceptual elements; unclosed figures exhibited closure tendencies and finally, images tended to be abstracted into three dimensions.

Results were in support of some of the traditional concepts of Gestalt theory such as wholeness, goodness, closure, proximity and similarity. At the same time, they provided strong evidence for the cell assembly theory. It was concluded that straight lines and short curves were basic perceptual elements.

In a slightly modified replication of the same experiment, Eagle and Klein (1962) found results in contradiction with those of their predecessors. For example, they found that angles were more stable than lines. An angle is composed of two line segments terminating at their point of intersection. The two segments are portions of longer lines which might for example be two sides of a triangle. Results in agreement with those of Pritchard et al. were also found: jagged lines were less stable than curved lines and 'meaningless' forms were less stable than 'meaningful' ones.

One of the strongest criticisms to be levelled at the method of stabilised images concerned the reappearance of the image after its initial fading. This finding was considered as incompatible with the operation of the retinal receptors. Campbell and Robson (1961) suggested that the image was not really stable and gave several possible explanations: slippage of the contact lens fixed upon the eye surface, changes in accommodation, pupil area variations, vignetting of the optical pathways or changes in intraocular tension. All of these could produce small residual changes in the position, size and illumination of the retinal

image. To bypass these doubts, Campbell & Robson developed a technique for observing the two entoptic shadows of the retinal blood capillaries and the macular pigment as stabilised images. It was found that both shadows disappeared and never reappeared spontaneously.

It therefore appeared that fragmentation and reappearance of the image were due to faults in the stabilising apparatus causing a non-stable image rather than to specific neural activity.

In an effort to justify the use of stabilised images, Evans (1965) studied images of various shapes using a stabilisation device capable of covering thirty degrees of the visual field. He found very large and consistent differences between the percentages of disappearance of most targets and this was found not to be a function of length of line or boundary. The presence of corners and intersections in a target increased the percentage of disappearance. Jagged angular forms disappeared more than rounded, topologically similar ones. When acute angles were present, the disappearance rate was radically raised. About ten percent of the time when a target disappeared, it disappeared as a complete unit, though this figure varied somewhat from target to target and was very high in the case of the circle. A mean of about fifteen percent of all disappearances were 'patterned' or 'structured' and these proportions held true for all subjects.

These results appeared to validate the use of stabilised image techniques since the consistency of results argued against such mishaps as lens slippage which would supposedly be random in nature. In a more specific paper, Evans (1967) achieved what he described as perfect image stability by the use of prolonged after-images. His results suggested the presence of recognition units mediating the perception of straight lines which varied in size from approximately thirty to ninety minutes of arc in the human fovea and perifovea. These units appeared to be arrayed independently of orientation and were thought to be primarily represented in the cortex.

McFarland (1968), performing a replication of the Pritchard et al. experiment, found not only lines but angles acting as units, though to a lesser extent. Using the method of luminous targets in a dark room, he obtained similar results. Finally, he tested which angles were more resistant to disappearance in stable images and found that 90° angles

were most resistant. He used this to explain why spiky forms tended to disappear more easily than forms which had sharp corners or right angles in them. In a previous experiment (McFarland, 1965), he had tested sequential part presentation of simple forms and concluded that perception was a sequential process although subjectively simultaneous. His method consisted of presenting the parts of a stimulus (the lines or angles making up the stimulus) in sequence, with intervals between the presentation of each part. Time thresholds at which subjects no longer felt that all parts had been presented simultaneously were measured as well as times when subjects no longer felt that all parts were joined. He concluded that two sequential processes occurred. First, the analysis in which individual parts give rise to responses and second, the integration performed on all of these responses.

An original approach to the study of perceptual elements was that used by Marsh (1967). Straight lines, curves, and angles (two sides of a triangle) were generated on a television monitor and embedded in visual noise (visual noise will be described in chapter IV). Subjects were given a detection task and the detection threshold was measured as a signal-to-noise ratio. Marsh found that vertical straight lines could be detected more easily than similar but slightly curved lines. Increasing the length of the straight line improved its detectability, until visual angles of between one and two degrees were reached, when the improvement became much less marked. The effect was virtually independent of line orientation. Broken lines were less detectable than continuous lines of the same total length. A symmetrically bent line (an angle) was no more detectable than one of its straight segments when the angle between them exceeded about 10° . The orientation of a straight line made little difference to its detectability, although on average, there was a slight reduction in detection threshold for the horizontal and vertical positions. Two identical parallel lines proved no more detectable than one line alone, if they were separated by a viewing angle of more than five minutes.

An experimental technique to gain ground in recent years involves the measurement of eye fixations. Richards and Kaufman (1969a, 1969b) used this method to establish whether certain features of stimuli were preferentially fixated. Their findings indicated that contours of small closed or partially closed figures such as circles, squares and

angles were not fixated but rather a position within the figures analogous to the centre of gravity. The apparent centres of gravity for closed and partially closed figures were found to be about the same. Richards & Kaufman concluded that if the eye was oriented towards the centre of gravity of a small figure, then the corners or boundaries might be processed simultaneously with optimal accuracy during the fixation pauses. However, this conclusion did not seem to apply in the case of larger figures. A test performed on figures of greater dimension found that fixations tended to occur at discontinuities, corners and along contour lines. In particular, angles were fixated more than lines. It appeared therefore that fixations depended very much on the size of the stimulus but that they did occur in a specific manner.

Noton and Stark (1971) made recordings of eye movements and concluded that the internal representation, or memory, of an object was a 'piecemeal affair': an assemblage of features, or more strictly, of memory traces of features. To them, recognition occurred when the internal representation was matched serially with the object, feature by feature. The features of an object were the parts of it (such as the angles and curves of a drawing) that yielded the most information. They postulated that the memory traces recording the features were assembled into a complete internal representation by being connected to other memory traces which recorded the shifts of attention required to pass from feature to feature, either with eye movements or internal shifts of attention. These attention shifts were supposed to connect the features in a preferred order, forming a feature ring and resulting in a scan path, which was usually followed when verifying the features during recognition.

The results discussed so far have all been concerned with human vision. However, a great deal of this research was motivated by earlier studies dealing with sub-human species. The work of Lashley (1938) on rats stands out in particular. In fact, his methods have long been used as standards in research on pattern recognition. One of his aims was to discover specific stimulus properties to which rats were reacting. To use his own words: "Whenever the attempt has been made to discover to what characteristics of the stimulus pattern the rat is reacting, by eliminating parts of the pattern or by testing for equivalence with new patterns, the adequate stimulus has been found to be some rather simple perceptual relation within the pattern..... The range of these general

properties revealed by equivalence tests seems quite limited. Presence or absence of a figure, direction, relative distance, relative brightness or size, and geometrically simple proportions probably exhaust the range. Yet by using these and by isolating part-figures, the animal is able to identify configurations of almost any degree of irregularity and numerical complexity of elements. Perhaps man employs no greater range of abstractions in his visual perceptions."

It is clear that Lashley believed in the existence of specific stimulus properties which could be abstracted and in the capacity of the rat to attend to those portions of a figure which were relevant to the perceptual task.

Another classic study involving animals was that of Sutherland (1957) using the octopus. His results from discrimination tests indicated that the octopus could more easily discriminate vertical and horizontal rectangles than oblique rectangles. Furthermore, a square was more readily discriminated from a triangle than a diamond from a triangle. Sutherland concluded from these findings that stimuli were partly classified in terms of horizontal and vertical dimensions, at least in the case of the octopus.

The results discussed in this section have all stressed the existence both in man and in animals of property analysers which react to specific stimulus characteristics. Other methods discussed in Weisstein (1969) and Dodwell (1971) have also reached similar conclusions. It appears that certain elements such as straight lines and curves are detected by specific analysers and that their characteristics such as size and orientation are analysed. A pattern is therefore not perceived as a pattern but as a composite of simpler elements with particular characteristics.

Many models have been proposed on the basis of these findings and some of these will be discussed in section 3. However, many of the models have also been postulated in the light of recent neurophysiological findings and these will now be discussed.

2.2 Neurophysiological evidence

In the previous section, some approaches and results to the problem of pattern recognition were examined. In this section, neurophysiological data will be discussed in the light of recent research dealing with sub-human species. Although the psychologist and the neurophysiologist may attack the problem of pattern recognition in a completely different manner, their conclusions are often very similar. While psychologists are preoccupied with enumerating a possible list of properties which can explain recognition, neurophysiologists are looking for specific pattern 'analysers' in the brain (Weisstein, 1969). The research described in this section will be presented as far as possible in chronological order. It must be remembered, however, that a variety of animals have been examined and that any generalisation of the results must be viewed cautiously. Above all, it must be kept in mind that none of the research concerns the human visual system. For a general review of the research to be presented, see Dodwell (1971) and Stent (1972).

Probably the first to specify the problem of pattern recognition was Lashley (1942). He believed that if the recognition of a pattern was to be independent of the particular set of receptors stimulated on a given occasion, then there had to be some general property of the nervous system which would account for the fact. The question he posed established the aims of much of the research that followed.

Before discussing particular findings, it is useful to map out briefly the neural pathway connecting the retina to the visual cortex. This pathway is of course slightly different for each species but can be described as a structure with the following general components. The retinal receptors (rods and cones) are the first cells to be activated in the pathway leading to the cortex. When an image falls upon them, certain photo-chemical processes occur which cause electrical activity to be transferred by chemical means across the junctions (synapses) between the retinal cells and the bipolar cells which constitute the next stage in the pathway. Bipolar cells then react with a third group of cells known as ganglion cells. There are fewer of these cells than there are retinal cells indicating the convergence of neuronal activity.

Apart from this direct vertical transmission of information towards the cortex, there is also horizontal transmission of activity at

the bipolar and ganglion levels. This task is accomplished by the amacrine and horizontal cells. It has become clear that the original information at the retinal receptors has undergone considerable transformation because of the complicated horizontal cross-connections and converging operations which occur between the retina and the ganglion cells.

From the ganglion level, electrical activity is relayed along the optic nerve fibres and through the optic chiasma (intersection of nerve fibres from both eyes) to the lateral geniculate body (one for each side of the cortex). At this point, information from both eyes has been received after having been separated into two parts at the chiasma: information from the right half of the visual field goes to the left geniculate body and information from the left half of the visual field goes to the right geniculate body. The amount of crossover at the chiasma varies between species.

From the lateral geniculate body, activity is finally relayed to the visual cortex which can be described in terms of successive layers of cells with different properties.

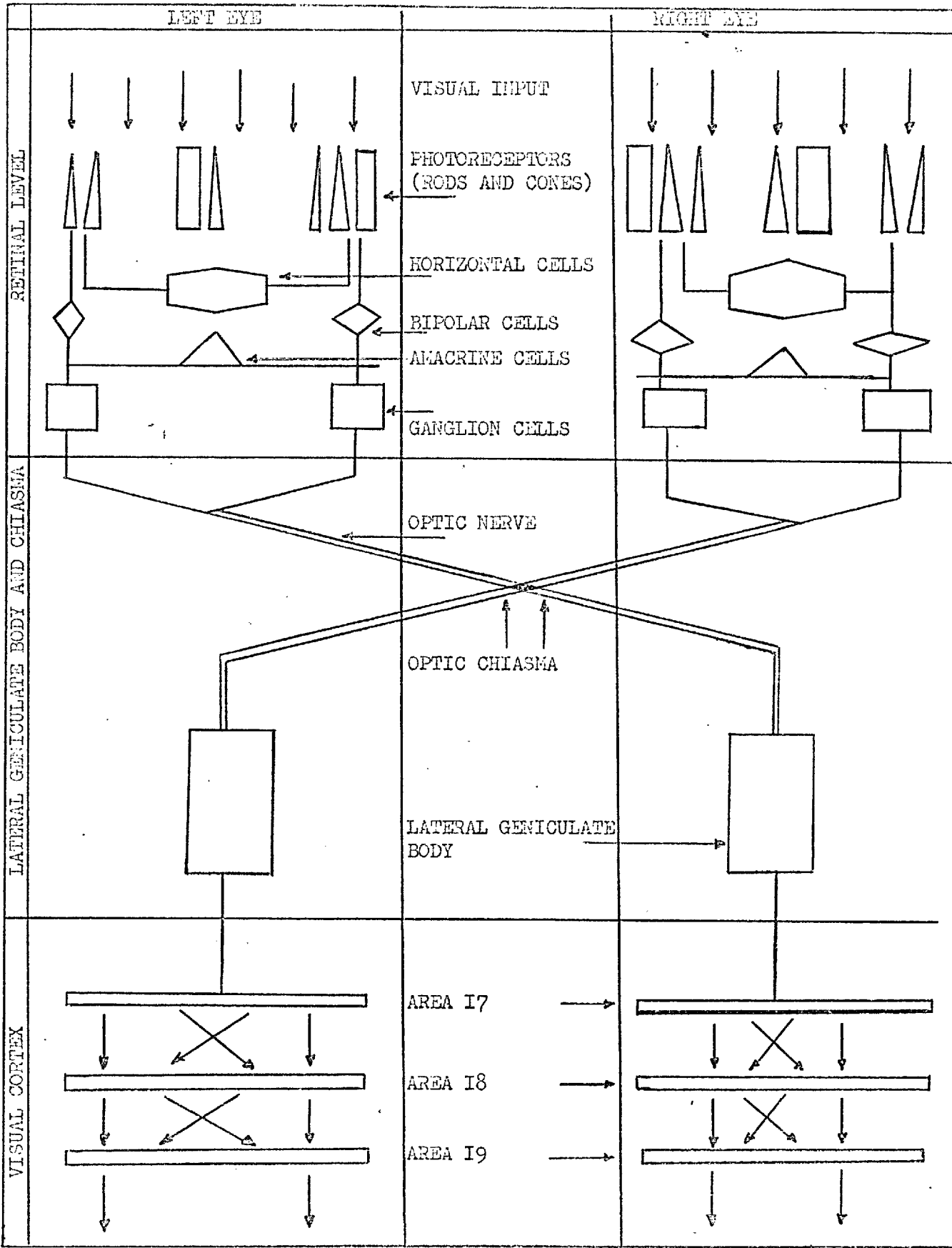
This very brief sketch of the visual pathway will now be described in greater detail in the accounts that follow. Figure 2.1 illustrates the basic pathway described.

It was mentioned that ganglion cells represented the last processing level before the transfer of neural activity to the cortex via the lateral geniculate body. For this reason and because of the relative ease with which measurements could be made of the electrical activity within this type of cell, ganglion cells were first chosen for systematic examination.

Kuffler (1953) in his research with cats discovered that the impulse traffic carried by the ganglion cells to the brain was not raw sensory data but an abstracted version of the visual input (Stent, 1972). At the time, he was investigating the possible existence of receptive fields in ganglion cells. A receptive field is that area of the retina whose interaction with incident light influences the impulse activity of individual ganglion cells.

At the beginning, Kuffler found that even in the dark, ganglion cells produce impulses at a fairly steady rate and that illuminating the

Figure 2.I General layout of the human visual pathway.



entire retina with diffuse light does not have any dramatic effect on the impulse rate. This finding suggested paradoxically that light does not affect the output activity of the retina. Kuffler then, however, projected a tiny spot of light into the cat's eye and moved the image of the spot over various areas of the retina. In this way, he found that the impulse activity of an individual ganglion cell does change when the light spot falls on a small circular territory surrounding the retinal position corresponding to the ganglion cell. That territory is the cell's receptive field.

On mapping the receptive fields of many individual ganglion cells, Kuffler discovered that every field can be subdivided into two concentric regions; an 'on' region, in which incident light increases the impulse rate of the ganglion cell and an 'off' region, in which incident light decreases the impulse rate. Furthermore, he found that the structure of the receptive fields divides ganglion cells into two classes: on-centre cells, whose receptive field consists of a circular central 'on' region and a surrounding circular 'off' region, and off-centre cells, whose receptive field consists of a circular central 'off' region and a surrounding circular 'on' region. In both types of cells, the net impulse activity arising from partial illumination of the receptive field is the result of an algebraic summation: two spots of light shining on different points of the 'on' region give rise to a more vigorous response than either spot alone, whereas one spot shining on the 'on' and the other on the 'off' region give rise to a weaker response than either spot alone. Uniform illumination of the entire receptive field gives rise to virtually no response because of the mutual cancellation of the antagonistic responses from 'on' and 'off' regions.

On the basis of these results, it could be concluded that the function of retinal ganglion cells is not to report the intensity of light registered at the retina but rather the degree of light and dark contrast that exists between the two concentric regions of the receptive fields. Such contrast information is of course essential to the recognition of forms within the animal's environment. Thus, it can be seen that the ganglion cells perform an abstraction of the visual information in a specific manner.

In a more detailed study, this time using frogs, Lettvin et al. (1959) approached their study with the following argument: "Suppose that the nervous apparatus in the eye is itself devoted to detecting certain patterns of light and their changes, corresponding to particular relations in the visible world. If this should be the case, the laws found by using small spots of light on the retina may be true and yet, in a sense, be misleading. Consider, for example, a bright spot appearing in a receptive field. Its actual and sensible properties include not only intensity, but the shape of its edge, its size, curvature, contrast, etc.

"We decided then how we ought to work. First, we should find a way of recording from single myelinated and unmyelinated fibres in the intact optic nerve. Second, we should present the frog with as wide a range of visible stimuli as we could, not only spots of light but things he would be disposed to eat, other things from which he would flee, sundry geometrical figures, stationery and moving about, etc. From the variety of stimuli, we should then try to discover what common features were abstracted by whatever groups of fibres we could find in the optic nerve."

The authors reached the following conclusion from their work: that the information at the retina was abstracted in terms of the following properties, namely information about local sharp edges and contrast, about the curvature of edge of dark objects, about the movement of edges and about the local dimmings produced by movement or rapid general darkening.

It appeared therefore that the visual system was responsive to specific stimulus properties and that these properties were detected following certain transformations in the original visual input.

In a study of the horseshoe crab (*Limulus*), Ratliff (1961) described the processes by which contrast information was picked up as a result of excitatory and inhibitory influences over the interconnections of the retina. The *Limulus* has the advantage of possessing a fairly simple visual system and was indeed used as far back as 1932 when Hartline used it to postulate the concept of receptive fields. Ratliff described how information at the retina was selected from the immense detail in the 'temporal and spatial pattern of illumination' of a stimulus and enhanced at the expense of less significant information before being transmitted to the central nervous system.

Ratliff concluded: "Among the most significant features of a pattern of illumination are the loci of transitions from one intensity to another and from one colour to another. Indeed, if only these contours are represented, as in a line drawing or cartoon, much of the significant information is retained."

Probably the best known neurophysiological studies of the visual system are those of Hubel & Wiesel (1962, 1963, 1968). Their work concentrated largely on cells within the visual cortex of cats and monkeys. They also examined cells within the lateral geniculate body and found that these had receptive fields similar to those of ganglion cells, but probably of a more selective nature. Receptive fields of cells in the cortex were, however, of a rather different nature. Their findings in the case of the cat can be summarised as follows (Hubel and Wiesel, 1962):

1. The visual cortex was studied by recording extracellularly from single cells. Spots of light of various shapes, stationery and moving, were used.
2. Receptive fields of cortical cells varied widely in their organisation. They tended to fall into two categories, termed 'simple' and 'complex'.
3. There were several types of simple receptive fields, differing in the spatial distribution of excitatory and inhibitory ('on' and 'off') regions stimulated at the retina. Summation occurred within either type of region; when the two opposing regions were illuminated together their effects tended to cancel. There was generally little or no response to stimulation of the entire receptive field with diffuse light. The most effective stimulus configurations, dictated by the spatial arrangements of excitatory and inhibitory regions, were long narrow rectangles of light (slits), straight-line borders between areas of different brightness (edges), and dark rectangular bars against a light background. For maximum response the shape, position and orientation of these stimuli were critical. The orientation of the receptive-field axis (i.e. that of the optimum stimulus) varied from cell to cell; it could be vertical, horizontal or oblique. No particular orientation seemed to predominate.
4. Receptive fields were termed complex when the response to light could not be predicted from the arrangements of excitatory and inhibitory regions. Such regions could not generally be established; when they

could, the laws of summation and mutual antagonism did not apply. The stimuli that were most effective in activating cells with simple fields -- slits, edges and dark bars -- were also the most effective for cells with complex fields. The orientation of a stimulus for optimum response was critical, just as with simple fields. Complex fields, however, differed from simple fields in that a stimulus was effective wherever it was placed in the field, provided that the orientation was appropriate.

5. Complex fields were generally larger than simple ones.

6. The cortex was found to be divisible into discrete columns; within each column the cells all had the same receptive-field axis orientation. The columns appeared to extend from surface to white matter. Within a given column, one found various types of simple and complex fields; these were situated in the same general retinal region, and usually overlapped, although they differed slightly in their exact location.

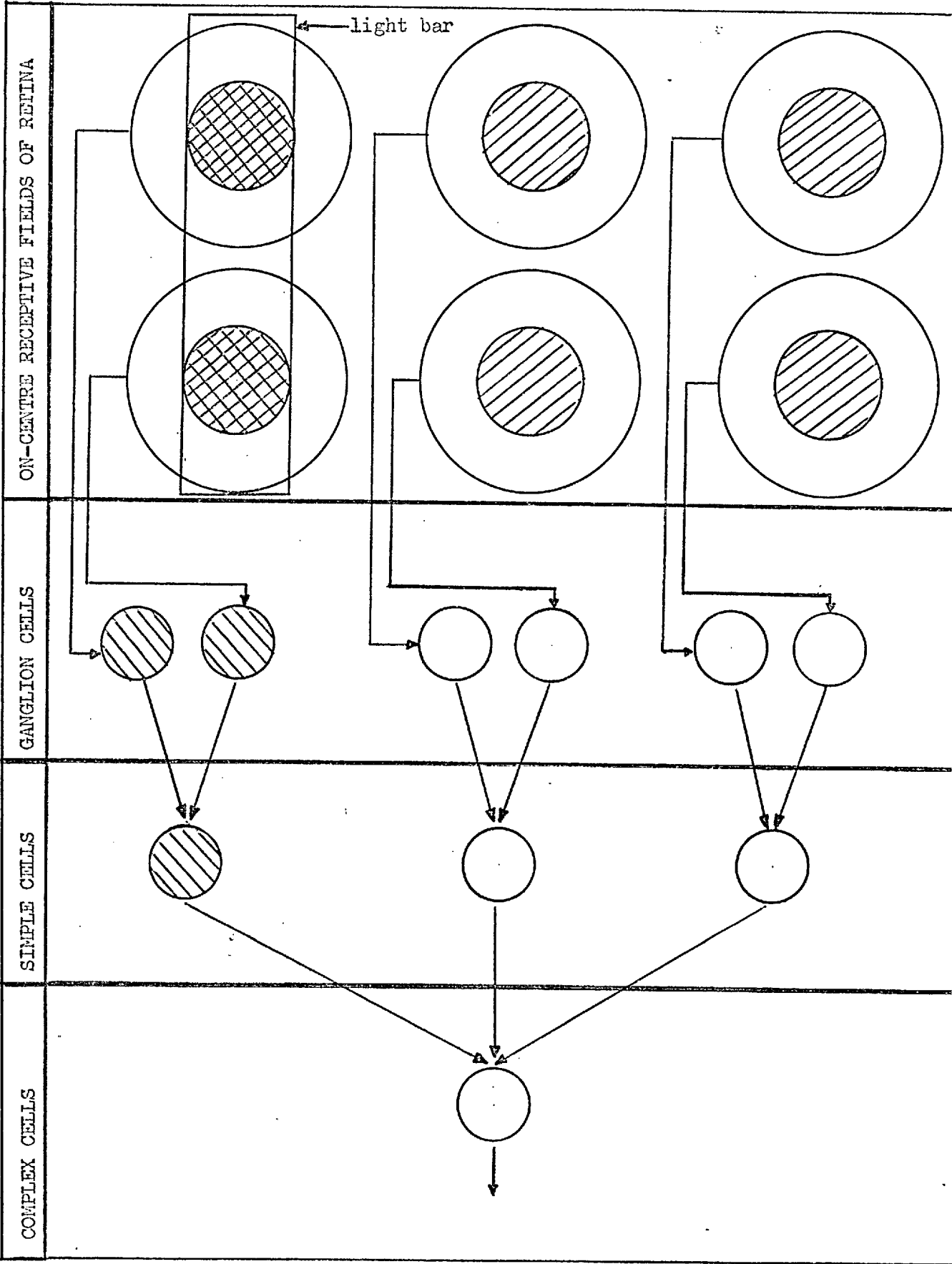
7. The authors suggested that columns containing cells with receptive fields of the same axis orientation were functional units, in which cells with simple fields represented an early stage of organisation, possibly receiving their inputs directly from lateral geniculate cells, and cells with complex fields were of higher order, receiving projections from a number of cells with simple fields within the same column.

In later work, they were able to identify cells in the visual cortex whose optimal stimuli reflected even higher levels of abstraction than straight lines, such as straight-line ends and corners (hypercomplex cells).

From this brief review of the recent research in neurophysiology, it is evident that the visual input is considerably modified on its way along the neural pathway. Furthermore, specific cells react to specific stimulus properties and the specificity increases as information penetrates deeper into the cortex. Figure 2.2 presents a simple illustration of this phenomenon.

From these findings, it becomes difficult to accept the Gestalt viewpoint as it stands at present. While it is true that the interpretation of the neurophysiological evidence presented so far is not without criticism, it nevertheless remains as the best ground on which to base theories of pattern recognition. Finally, although none of the present findings apply to human vision, it seems reasonable to assume that the more general conclusions may be applicable with perhaps some modification.

Figure 2.2 A schematic representation of the abstraction and convergence of visual information (a light bar) in the visual system of the cat.



Having presented both psychological and neurophysiological research pertinent to pattern recognition, the third section will now deal with models which have been proposed in the light of research described in the two previous sections.

2.3 Some models for pattern recognition

In the last two sections, evidence was presented for the existence of property analysers concerned with the processing of visual information. In this section, a few models of pattern recognition will be described. None of these can be discussed in depth since each would warrant a chapter by itself.

In recent years an enormous number of models have been proposed to account for the process of recognition and some of the more important ones have been reviewed by Hake (1957). Earlier models suffered from a lack of experimental evidence and were often too general to account for a specific phenomenon. Later models have, on the other hand, been too specific and inflexible and could not account for more general phenomena. Furthermore, neurophysiological evidence to explain recognition is still negligible. Only the initial stages in the analysis of the visual image have been studied and we are left to speculate about the neuronal processes by which the analysed image is recognised. One possible explanation was proposed in the first section of this chapter (Noton and Stark, 1971). Basically this model assumed a serial, feature by feature, matching process of the input image with some internal representation built up on past experience. The internal representation was described as a 'feature ring' in which feature memory traces were connected by other memory traces which recorded shifts of attention from one feature to the next in the original scanning of the stimulus. This was based on evidence from eye movement recordings indicating that each person will scan a particular stimulus in the same manner everytime, except for a few isolated deviations. Features were described as 'those parts of a stimulus yielding the most information i.e. corners and contours in line drawings'.

This model is an example of the speculation which must go into the building of models to explain recognition. As a matching-type model, it is not new although it uses a different experimental approach.

Furthermore, it does not go as far as it should in explaining certain questions. For example, is the input image matched with every single internal representation that we have of our environment? If so, is the process exhaustive or self-terminating (Bradshaw and Wallace, 1971)? How does recognition occur when only part of the stimulus is perceived? Pasnak (1971) showed that recognition for complex stimulus forms often occurred on the basis of part-perception which would indicate a modified feature-ring structure.

Most models of recognition have drawbacks and it is always easier to criticise them than to propose alternative solutions. The purpose of this section is not so much to offer criticism as to describe some of the interesting attempts which have been made in the construction of models.

Before describing the first model it would be useful to describe what is meant by this term. Dodwell (1971) described a model in the following manner: "Usually, the process of constructing a model goes through the following stages: (1) a specification of the to-be-explained behaviour is made; (2) a system of 'elements' or 'units' is devised which have certain well-specified properties and relations with one another; (3) these elements are put together into a structure that has properties which in given respects mirror the real system that is to be modelled. The behaviour of the system is then predicted on the basis of deductions made from the model and its properties. Such explanatory devices can be distinguished from theories, which generally may be of wider scope but also more vague in their predictions."

One of the first models to be partially derived from neurophysiology was that of Deutch (1955). Before stating his model, he began by describing a set of established facts. For example, animals can recognise shapes independently of their size, orientation and their position in the visual field. Mirror images appear alike and are occasionally confused by both rats and humans. The octopus and the rat find it difficult to distinguish between the square and the circle. All the above facts apply even after removal of most of the cortex.

Having stated these facts, Deutch then described his model. It assumes the existence of a two-dimensional array of units (cells), which has information arriving on it from retinal light receptors. Each group of cells is joined to a particular retinal element in such a way that

neighbouring retinal elements also excite neighbouring groups of cells on this two-dimensional array. It is assumed that only the excitation from contours is passed on from the retina. Thus when a ball is focused on the retina, a circle will appear on the two-dimensional array.

Each unit on the array can be excited by a contour falling on the region of the retina to which it is joined. When such excitation from the retina arrives, each unit transmits a pulse down a final common cable. It also excites its neighbour. The neighbour will pass it on in turn. Therefore when a contour is projected on the two-dimensional array, two things occur: a signal consisting of one pulse is passed down the final common cable by each cell on which that contour lies. Thus a measure of the total number of cells stimulated is passed down. Second, the contour excites all the cells which lie next to it on the two-dimensional array. These pass the excitation on to their neighbours but not down the final common cable. The assumption is made that a cell passes on its excitation at right angles to the contour of which it happens to be a component. Thus the excitation produced by a straight line will be a line equal and parallel to it and gradually moving away. In the case of a convex arc, it will also be an arc moving away 'parallel' to it, but gradually decreasing as it draws away, coming to a point and then fanning out again, this time concave.

As such lateral excitation from a point in a contour advances, another signal will be sent down the final common cable as soon as the excitation coincides with another point on a contour imposed on the two-dimensional array. This signal down the final common cable will at each moment give a measure of the number of points thus brought into coincidence. For example, if the figure is a rectangle, the first set of pulses sent down the final common cable will be a measure of the total number of cells excited by the contour. The second set of pulses down the final common cable will be sent when the excitation produced by the longer sides actually coincides with these sides.

As these two sides are parallel to each other, the excitation and the contours will coincide simultaneously and send a number of pulses proportionate to the length of these sides down the final common cable all at the same instant in time. Some time later the disturbance from the two shorter lines will reach each of these lines and another message will be sent down the final common cable, indicating that a smaller number of cells

have been fired off this time but also all simultaneously. Thus the message for a rectangle will consist of three sharp volleys, the proportions of the last two being governed by the ratio of the longer to the shorter sides. The message generated by a square will be two sharp volleys, both equal in magnitude. As the two pairs of sides are equidistant, the advancing disturbances will hit the sides simultaneously. As they are parallel, the disturbance generated and the contours will overlap all along their length at the same instant in time, causing the cells representing the contour to fire down the final common path together. A circle would also generate two sharp volleys.

To account for the recognition of a shape regardless of its size, the following operations might occur. First the measure of the total number of cells involved in signalling the length of the contour would pass down the final common cable. It could then act on a section of the cable through which the whole message would soon pass on its way to the final detector units. It would throw this section of the final common cable into a state of refractoriness proportional to the magnitude of this measure. It would thus alter its threshold of response to the following signal to differing degrees. Only a very small proportion of a signal following a large pulse would survive. A large proportion of the signal, after a small pulse had set the 'resistance', would pass. Thus if both these signals were derived from the same shape, let us say at different distances, they would both give rise to a signal, the height of whose pulses was physically identical. This mechanism would explain recognition regardless of size.

The author then followed his description by a set of predictions which could be tested. He also pointed to inadequacies of the model, namely that it could only handle one figure at a time and that in fact humans were very capable of distinguishing between squares and circles.

Furthermore, it could be said that Deutch has not really explained recognition but rather the initial coding which takes place before recognition. The model is nevertheless ingenious for its simplicity and has inspired many later models which also make use of spatial and temporal properties. Finally, Deutch believed that his model was based on feasible neurophysiological processes. A few years later, Sutherland (1957) proposed a model to explain discrimination of

orientation and shape in the octopus. The model also used the concept of the two-dimensional array of cells but this time analysed the image in terms of its horizontal and vertical components. As a model specific to the octopus, it appeared to be satisfactory but like Deutch's model, it lacked generality.

A more comprehensive model was proposed by Dodwell (1964). This time, however, the author went as far as suggesting a mechanism for recognition. After describing a system for coding visual information about shape, he showed how the output from this coding system could be coupled to a classifier, or recogniser, in such a way that shape equivalence was preserved. To him, a major weakness of previous coding-type theories of shape recognition was remedied in his system by the use of a built-in memory in the recogniser which became self-organising in terms of stimulus regularities in the environment. Finally, the author outlined the possibility of connecting the recognition system to a response system. Essentially, the initial coding system transformed spatial information into a temporal analogue, once again by the use of two dimensional arrays of cells. The temporal analogue consisted of sequences of pulses of different amplitudes and delays which all travelled down a single transmission line. The model produced the same output for a stimulus regardless of its position in the visual field, its size or its orientation.

The coded information was then fed into the recogniser. Since the information was in the form of pulses of different amplitudes and delays, the author postulated a conversion of this analogue information into a digital equivalent. The particular analogue-to-digital converter required that the recogniser have a number of inputs, each one being sensitive to a particular pulse amplitude range. A unit impulse would fire at a particular input when a pulse arriving at that input was of the right amplitude. The inputs were connected to counters which registered either the initial arrival of unit impulses from the inputs or the temporal patterns of impulses from the inputs. This was achieved by feeding each input into a series of delay units, and then to the counters. In this manner, 'equivalent stimuli' could be defined as inputs from the receptors which gave rise to the activation of the same set of counters in the recogniser.

The coupled system described so far has no 'memory'. Its performance is independent of its previous states. To show how the system could be modified to incorporate the concept of learning, Dodwell proposed the notion of a conditional probability computer. Consider the following hierarchical classification of counters: counters A and B are connected to a counter AB, counters A and C are connected to a counter AC, counters B and C are connected to a counter BC. Furthermore, counters AB, BC and AC connect to a counter ABC. Simultaneous inputs at, for instance, A and B will activate the counters A, B and AB. Simultaneous inputs at A, B and C will activate all counters. Suppose that each counter has a limited memory, such that each time it counts, its state is altered, but gradually tends to revert to its initial condition. Suppose, further, that in this system two counters in approximately the same state will count if either one of them is activated. For example, if B has almost always occurred at the same time as A, then if A is activated (but not B) the B and AB counters will be activated. By choosing a suitable criterion for 'same state', such a system computes conditional probabilities at a specified level. The probability that B occurs, given A, determines whether or not the B and AB counters register when A occurs alone). Dodwell summed up his model as follows: "It is clear that the system described will recognise patterns, conjunctions of patterns and conjunctions of patterns occurring in rapid succession. Such conjunctions of patterns must lead to generalisation or inference by the system and this is all that is required for the coupled system of coder and recogniser to exhibit the type of perceptual learning required."

Dodwell went on to describe a possible response system connected to the recognition system. Overall, his model of pattern recognition represented a serious attempt to explain not only the coding of visual information but also the actual recognition process. The system borrowed from Uttley's (1954) classification system of signals in the nervous system and appeared to conform with the existing knowledge in the field of neurophysiology. Dodwell pointed out the importance of studying how stimuli were recognised after the initial coding process. As such, his model although very hypothetical, is a distinct contribution.

A number of models of pattern recognition have resulted from attempts to build machines which can recognise. Some of these models do

offer analogies between the machines for which they are proposed and the possible recognition mechanisms in the visual system. One model which did not specifically attempt to do this but which contained many ideas later used by psychologists was that of Grimsdale et al. (1959). Another, which in fact was very intent on making analogies, was that of Kabrisky et al. (1968) which will now be discussed briefly.

The model begins with the following assertion: "The principal, and certainly most striking, feature of the retina - geniculate body - cortex transmission channel, at least in primates and man, is that it maps the retina homeomorphically on to the functionally flat cortex." In other words, the visual input is essentially preserved from any recognition processing until it reaches the cortex. It is in the cortex, that recognition takes place.

The cortex is composed of successive layers which have been assigned area numbers. For example, the first layer receiving projections from the lateral geniculate body has been termed area 17. The layer beneath it is area 18 and so on. Kabrisky et al. were concerned in mapping the projections from one layer to the next. To do this, they first made an analogy with optical systems. One of the computations that can be made by an optical system is to produce the two-dimensional Fourier transform of any two-dimensional transparency. It produces this transform because of the interference and reinforcement of various portions of electro-magnetic fields propagating along pathways of different length within the system of lenses. They proposed that in the case of a neurone model of such a system, it would probably be necessary to maintain 'consistent relative' propagation times in axonal pathways and to provide sufficient overall synchronization of activity to allow the possibility of post-synaptic excitatory potential and post-synaptic inhibitory potential interaction in individual neurones to provide the equivalent of reinforcement and interference found in the optical system.

Kabrisky et al. then postulated that the cortex did in fact compute the Fourier transforms of inputs by means of cross-correlation computations resulting from the mapping between the different cortical areas. In other words, the projections from area 17 to area 18 provided the initial cross-correlation computations in the build-up of the Fourier transform. This hypothesis was tested by means of a digital computer investigation

and was confirmed according to the results. It was concluded that an N by M pattern characterisation had a unique and invertible N by M Fourier series transform and that only one half of the transform was required in order to describe a real valued pattern uniquely. Furthermore, a recognisable reconstruction of most patterns could be made by inverting only the lower frequency portion of the transform. Finally, by simply correlating the lowest spatial frequencies of transforms, a start could be made in the actual recognition process.

Whether or not the visual cortex does in fact compute Fourier transforms cannot be ascertained until neurophysiologists have made accurate maps of the projections between cortical areas and actually tested whether cross-correlation does occur in such a way as to arrive at Fourier transforms. The model proposed by Kabrisky et al. is very appealing but at the same time, perhaps very far removed from the real situation. It does, however, indicate the trend in model building towards more precise and mathematical formulations.

The models described in this section represent only a small cross-section of the work being done in this field. Most of them have used the findings presented in the first two sections of this chapter. They have also borrowed from the fields of computer theory and machine pattern recognition. It is difficult to evaluate how successful any of these models are. Most of them have introduced interesting new concepts, but until more is known about the neurophysiology of the visual system, they will remain highly speculative.

2.4 Summary

In section 1, experimental evidence was presented to suggest the existence of specific visual analysing units in the brain whose function was to record the presence of particular stimulus properties. Results suggested that basic perceptual elements such as lines, curves and angles were analysed individually in terms of characteristics such as size and orientation. Different experimental approaches were presented and these resulted in findings which did not always agree, e.g. whether angles were basic perceptual elements. However, the general conclusion was that stimulus information was broken down in terms of specific properties before being processed for recognition. This breakdown was referred to in section

3 as the initial coding of the information. Cell assembly theories were seen as more plausible than the Gestalt theory.

Section 2 reviewed the latest results obtained in neurophysiological studies of the visual system of sub-human species. Receptive fields at different levels in the visual pathway were described and the evidence suggested once again that stimulus information was being processed at the different levels by cells responsive to particular properties. Cells in the visual cortex were found to react to increasingly specific stimulus characteristics.

Section 3 presented a few recent models of pattern recognition. Most of these made use of a variety of disciplines in their formulation: computer theory, neurophysiology, psychology, optics and machine pattern recognition. Although each model could be faulted in some respect, it represented a distinct and interesting approach to the problem. Basically, the visual system was divided into three parts: the initial coding system, the recogniser and the response system.

In conclusion, it was stated that despite the originality of such models, they would remain highly speculative until more was known about the neurophysiology of the visual system, particularly in man.

It should be added that these models were not presented with a view to backing the author's own results but simply because they were considered as representative of the work being carried out in the field of pattern recognition.

CHAPTER III

THE EXPERIMENTAL VARIABLES

Section 1 gives a general description of the variables which may affect recognition. These are broadly classified as personal experience variables, stimulus variables and experimental variables. Some examples of the effects of these variables are presented.

Section 2 describes the stimulus variables which have been defined in the present research for the case of facial images. Section 3 describes the stimuli and subjects used in the author's research and Section 4 gives a general preview of the experiments to be described in chapters V, VI and VII.

3.1 General variables in a recognition task

The purpose of this section is to present a broad classification of the variables which affect us in recognition. It is essential to understand these variables in order to interpret recognition data correctly. As a general statement, it may be said that every individual is affected in a different manner by the variables which will now be described. These can be classified into three categories: variables related to the individual's personal experience and state, variables which depend on the stimulus configuration and variables which are introduced by the experimenter in a recognition task. These will now be discussed in greater detail.

3.1.1 Personal experience variables

In section 2 of chapter 1, the concept of familiarity was presented as a variable which affected recognition. It was pointed out that faces were generally easier to recognise than snowflakes or inkblots (Goldstein and Chance, 1970) partly because we were more familiar with them. As a variable, familiarity might appear to be vague but it should be recalled that it was introduced to counteract the general assumption that geometrically complex configurations are more difficult to recognise than simple ones. Because of this, it must be regarded as an important general variable. For example, it was mentioned in section 1.1 that inverted and negative photographs of faces are more difficult to recognise than upright and positive ones. Obviously there are physiological reasons for this, but at the same time, the simple fact that we are not familiar with such representations in our daily life is often forgotten. In other words, the approach to such configurations becomes too restricted and the learning aspect of our behaviour is neglected.

In the present context, it must be pointed out that the concept of familiarity is only defined for the processes of recognition and identification and not for detection and discrimination. It is assumed that a stimulus which has not been seen before may still easily be detected in one case and discriminated from another stimulus in another case.

In an experiment designed to test the recognition of human faces, canine faces and architectural forms, Scapinello and Yarmey (1970) found

that stimuli presented initially for seven successive inspection trials were significantly better recognised than were stimuli inspected only once. Chance et al. (1967) investigated the recognition by school children of their classmates' faces when certain portions of these were covered by masks. They found that the length of acquaintance between the subject and the pictured child did not affect performance but that the 'degree of friendship' between them did. In the same experiment, they also tested the significance of the children's ages. Strictly speaking, age is not an independent variable affecting recognition since it is often implicit in the concept of familiarity. Chance et al. found a positive relationship between age and accuracy of identification. Brooks and Goldstein (1963) made a similar finding in a recognition task involving inverted photographs of faces. The age range was from three to fourteen years old.

In a test designed to measure incidental recall of photographed human faces, Messick and Damarin (1964) found three variables which influenced the subjects' performances. First, subjects who had been found to be field-dependent in an embedded-figures test recalled more faces correctly than did those who were field-independent. Subjects who were narrow categorisers on the Pettigrew Category-Width Scale had better recall than those who were broad categorisers. Subjects who thought the photographed persons were relatively young did better than those who thought they were older. The authors claimed that these three stylistic consistencies were mutually independent. These three 'perceiver-variables' are perhaps very specialised aspects of perception, but they indicate nevertheless the variety of effects which may alter each person's capacity to recognise.

In a broad study of perceiver variables, Baker (1967) found a wide range of variables which influenced the recognition of Identikit pictures of faces; field-dependence, degree of extraversion, general motivation and interest, ability in discrimination tasks, verbal description tendencies and sex. He emphasised the importance of considering these variables when interpreting data from experiments. In the case of the subject's sex, women have often been found superior to men in the recognition of faces. Witryol and Kaess (1957) found clear evidence of this in their research. In a breakdown of the results, they found that males were superior at recognising male faces and females were superior at

recognising female faces as well as being superior overall. Age and intelligence were not found to be variables (the subjects were adults). Female superiority was attributed to women's greater social dependence and motivation.

One of the few studies to investigate recognition of faces between different races has been that of Malpass and Kravitz (1969). Black and white faces were tested in both a predominantly white and a predominantly black American University. Overall, white faces were found to be more discriminable than black faces and subjects were better at recognising their own race. Questionnaire data suggested differential experience with persons of the other race only in the case of black subjects at the white University.

A variable which may be important in recognition concerns the saliency of features of the stimulus. In other words, it refers to the importance, or saliency, that we attach to particular stimulus features. In a study using simple schematic faces of the same type as those used by Brunswik and Reiter (1938), Friedman et al. (1971) found that the effect of feature saliency was greater on response bias than on accuracy. This means that we tend to concentrate on certain features that we think important, when we are performing a recognition task even though these features may not be as useful as others in assisting our recognition. This type of bias will also vary with the stimulus configuration in question and can therefore be partially classified as a stimulus variable.

This brief review of some of the variables which are classified as personal experience variables suggests the importance of considering certain aspects of the subject's background before experimenting with him. At the same time, care must be taken in establishing sensible variables in the first place. Women, for example, might be superior to men at recognising cosmetics, but it is highly doubtful this would influence the outcome of a test designed to measure the recognition of faces. Therefore, a 'recognition of cosmetics' variable is not really needed and can be replaced by the more general variable of familiarity. However, race, sex, familiarity and sometimes age may often be important factors that affect the outcome of recognition experiments.

3.1.2 Stimulus variables

One of the least defined stimulus variables is that which relates to class homogeneity. The homogeneity of a stimulus class depends on the number of features which can serve as a basis for discrimination between different stimuli in that class. Like the concept of familiarity, it remains of a qualitative nature. Probably the best known perceptual theory to use the concept of homogeneity in its formulation is that of Gibson (1967). Her theory of learning is one of differentiation and is completely stimulus-oriented. In other words, learning and development occur through selective differentiation of stimulus properties. This is done by attending selectively to those features of a stimulus which help to differentiate it from another. For example, a child may at first recognise only its mother's face. With time, however, it learns that there are many faces within its environment and that each of these has distinguishing features. In the process, it has realised that faces have certain features in common which make them different from, say toys, but that they also have features which help distinguish one from another. In other words, the child first learns to recognise classes of stimuli and this is followed by recognition of members within that class. To Gibson, stimulus variables such as class homogeneity establish how easy or difficult it is to recognise a particular stimulus. This perhaps presents an incomplete picture of the range of variables which may in fact be responsible. For example, Gibson does not consider the variables mentioned in the previous subsection which certainly do have an effect on recognition even though they have nothing to do with the stimuli themselves.

Another stimulus variable is that of geometrical complexity. Intuitively, a stimulus is said to increase in geometrical complexity if it is composed of more and more lines, angles and curves. Whether a complex stimulus as defined above is more difficult to recognise than a simple stimulus has been the subject for considerable debate. On the one hand it is held that a complex stimulus has more features which can be used to distinguish it from another equally complex stimulus and that recognition should therefore be easier than in the case of simple stimuli. On the other hand, a simple stimulus is considered to be composed of a

small number of 'building blocks' and should therefore be easier to recognise than a complex stimulus. The confusion between the two assertions is probably due to the fact that different types of stimuli are used to illustrate each argument and that consequently different variables come into play. For example, the former argument would maintain that facial images, although geometrically complex, are easy to recognise. This neglects the variable of familiarity which certainly affects our ease of recognition in the present case. The latter argument would maintain that randomly drawn complex forms which are relatively homogeneous are more difficult to recognise than the equivalent simple forms. Once again, familiarity or rather the lack of it may account for this result. In both cases, class homogeneity may be a relevant factor.

One study which attempts to obtain some measure of the variable of geometrical complexity is that of Vitz and Todd (1971).

They propose a model to explain how humans perceive straight-line two and three dimensional geometric figures such as triangles, trapezoids and cubes. The major characteristics of the model are as follows: a) a figure is represented as a hierarchy of perceptual elements (lines, angles and areas), b) the process of perceptually organising a figure is portrayed as a stochastic sampling of the elements, c) the total number of trials required to sample the elements, corrected by a measure of redundancy due to symmetry, is proposed as a measure of the figure's complexity. The model is found to be successful in judging the complexity of a wide variety of two-dimensional figures. With some additional rationale, it also predicts the perceived three dimensionality of figures such as Necker cubes.

In this model, Vitz & Todd are partly successful with their predictions because they have considered the important effect of redundancy. However, such a measure is not easy to apply to the majority of visual stimuli and indeed very few stimuli are composed simply of straight lines and angles. Nevertheless, the model presents a distinct contribution because it recognises the importance of redundancy unlike some of the pattern recognition models reviewed in chapter two. It would, after all, seem reasonable to suppose that our brain does take advantage of perceptual redundancy in processing information (Attneave, 1954).

Another model, in some ways similar to the previous, is that of Tversky and Krantz (1969). This time the authors are not measuring complexity but dissimilarity. The stimuli consist of schematic faces composed of three features: an elliptic facial contour, a pair of circular eyes and a line mouth. Each of these features is presented in one of two forms: the elliptic contour is drawn with its major axis either horizontal or vertical, the eyes are either blank or filled in as black dots and the mouth is either a straight line or an upturned curve as in the smiling expression. The variation of these three binary features produces a total of eight schematic faces. Tversky & Krantz used these to obtain direct and comparative judgements of dissimilarity. The resulting data was used to test the hypothesis that the overall dissimilarity between faces can be decomposed into three additive components, one for each feature. The hypothesis was strongly confirmed by both the direct and comparative judgements.

The concept of redundancy does not of course apply in the present case. The model is successful and does indicate a certain breakdown of effects to the feature level. However, in the opinion of the author, this type of model would be incapable of dealing with stimuli composed of a few more features especially if these varied in many ways. It is even possible that certain pairs of stimuli would be judged as more similar than others even though more individual features were different between them than in the case of other pairs. In other words, certain feature effects might be negative overall.

It can be seen from this brief review of stimulus variables that very little is known about the effects of variations between stimuli in the same class. Geometrical complexity and class homogeneity were discussed and it was pointed out that these variables would probably remain more qualitative than quantitative except in the very simplest cases, at least until more was known about the neurophysiological aspects of the recognition process.

3.1.3 Experimental variables

An important class of variables which affect the outcome of an experimental recognition task are those which are intentionally or unintentionally imposed by the experimenter. Often their effects cannot be

separated from those of variables discussed previously. For example, a face presented in negative form is not only transformed experimentally but also less familiar and a face which is partially masked is probably of a different order of geometrical complexity.

Laughery et al. (1971) investigated the effects of stimulus exposure duration and target position in a face recognition task. The experiment tested the ability of subjects to recognise a target face (previously seen) in a test series of one hundred and fifty faces. They found that the longer the original exposure to the target and the earlier the target appeared in the test series, the greater was the probability of recognition. They used both monochrome and colour photographs but found that this had no effect on the outcome.

Wallace et al. (1970) measured the effect of delay between initial presentation and testing in a face recognition test. They found that a forty-five second delay was optimum in their test but suggested that this would probably vary, depending on the size of the testing sample. The finding is interesting because it shows that a delay is more effective in some cases than a no-delay situation between initial presentation and testing. Wallace et al. suggested that the effect of the delay, which is called 'reminiscence', was to permit 'digestion and consolidation' of the information before testing.

A well known experimental variable is the effect of the number of alternatives on a recognition test. Fraisse and Blancheteau (1962), in a study of the recognition thresholds for nonsense syllables, found that these increased in proportion to the logarithm of the number of alternatives. The range of syllables extended from two to fifteen. Such relations have been found for other classes of stimuli.

In an interesting set of exploratory tests dealing with face recognition, Harmon (1971) found that the lighting conditions used in presenting face drawings affected recognition. Rear lighting of the drawings produced softer images which were preferred to front lighting and recognition was to some degree positively correlated with this preference. Harmon also tested recognition with computer generated faces and suggested that blurring of the image could sometimes improve recognition of faces not previously well-known.

Another exploratory study designed to measure the effect on recognition of prior knowledge of facial expressions was that of Galper and Hochberg (1971). The authors used photographs of smiling and non-smiling faces and suggested that the difference between these two expressions was in itself information useful to recognition. In more general terms, this implied that prior knowledge of a person's range of facial expressions did in fact make recognition of that person easier.

It becomes clear that the experimenter has a substantial range of variables which he can use to design his experiment. In some cases, however, variables are also introduced unintentionally by the experimenter. In a classic study of this problem, Rosenthal (1963) gave examples to illustrate the existence of experimenter bias. For instance, he found that the experimenter's expectation of the results and motivation could contaminate the outcome of certain tests. Furthermore, he indicated that this bias was not only introduced by verbal commands but also by non-verbal behaviour. He stressed the importance of uniform and carefully planned approaches to the experimental situation.

Experimental variables are usually classified as either random or fixed, depending on the way in which they are chosen (Edwards, 1968). For example, the procedure in choosing a group of subjects for an experiment is often random within certain specified limitations such as sex, age group or perhaps I.Q. level. For that reason, the 'subject' variable will usually be considered as a 'random effect' and the 'sex' or 'age group' variable will be considered a 'fixed effect'. In the latter case, the experimenter is capable of specifying the 'levels' of his variable whilst in the former case he is not. Experiments based on the use of both random and fixed effects are referred to as 'mixed models' and will be encountered in chapter V.

3.1.4 Summary

The purpose of this section was to present a broad classification of the variables which interact in a recognition task. Three general categories were identified; variables which depend on the perceiver's general experience and state, variables which are dependent on properties of the stimuli and variables which are introduced experimentally. Examples were given of research specifically designed to test the effect of these variables.

3.2 Facial attributes

In chapter two, patterns were described in terms of simple elements such as straight lines, curves and angles. These were said to be processed initially by a coding system sensitive to particular stimulus properties. In the previous section, a model was proposed by Vitz and Todd (1971) to explain the recognition of simple geometric forms. The components of each form were accurately defined and theoretical predictions were made of the complexity of each form. Obviously, the problem of defining stimulus properties in the case of facial images becomes more difficult. It should also be stressed that any definition of stimulus features remains arbitrary.

One interesting approach to the problem used the concept of feature descriptors (Goldstein et al., 1971). Examples of these are 'hair coverage' 'eyebrow separation' and 'ear length'. Goldstein et al. selected a group of 255 faces and asked a group of subjects to evaluate each of these faces using a set of 34 descriptors. A descriptor was rated on a numerical scale which usually varied from one to five. For example, in the case of 'ear length', a score of one was given for short ears, three was given for medium-length ears and five was given for long ears. It was found that the original set of 34 descriptors contained descriptors which were either inconsistently judged, i.e. the values assigned by subjects to the same feature of a particular face were not well distributed, or trivial, i.e. hair parting was considered 'a dead giveaway'. Therefore, the authors reduced the original sample down to a final group of 22 descriptors. Each of the 255 faces was evaluated according to the 22 descriptors. The authors then proceeded to test recognition by means of the descriptor values. In other words, a new group of subjects was given a sheet containing the 22 descriptor values particular to one of the 255 faces. Their task was to identify the face whose descriptor values they had at their disposal. All 255 faces were available to them. They proceeded in the following manner: first they found the most extreme valued descriptor in the list. Using this descriptor, they were able to immediately eliminate a large number of faces which didn't fit the extreme value. Having done this, they found the second most extreme valued descriptor in the list. This helped

in eliminating a large portion of the remaining faces. This process continued until the correct face was isolated. On average, seven descriptors were needed to identify one face from the sample of 255 faces.

In a later study (Goldstein et al., 1972), the authors made use of a man-machine interactive system to perform the same task. The system takes advantage of human superiority in detecting and describing important features and of machine superiority in making decisions based on accurate knowledge of the population statistics of the stored descriptors. In extensive tests, it was found that the population containing the 'face to be recognised' was narrowed down to less than four percent in 99 percent of all trials.

In choosing the final set of 22 descriptors, Goldstein et al. adopted a criterion which would provide 'distinctive and relatively independent measures which could be judged reliably'. In other words, they were concerned with finding a set of descriptors which would uniquely specify a face at least within a population of 255. Descriptors which were considered either ambiguous or difficult to judge reliably were eliminated. The Goldstein et al. model is undeniably successful in what it does and it presents interesting possibilities for the computer identification of faces. However, it does not explain recognition nor does it pretend to do so. Rather it demonstrates that certain feature characteristics are judged reliably and consistently and that faces can be recognised on the basis of a few feature descriptors.

In the author's research, a more general approach is used in specifying facial features. These are split up into three categories: discrete features, contour features and feature relations. The general term used to indicate a feature in any of the three classes will be the word attribute. The three classes of attributes will now be described.

a) Discrete feature attributes.

These are the mouth, the nose and the eyes which include the eyebrows.

b) Contour attributes.

These consist of the forehead-hairline contour and the head outline contour which includes the ears.

c) Feature relation attributes.

These include the following separations: the mouth to chin tip separation, the forehead width, the eye to eye separation, the eyes to nose separation, the eyes to mouth separation and the mouth to nose separation. Separations are measured from the points illustrated in figure 3.1.

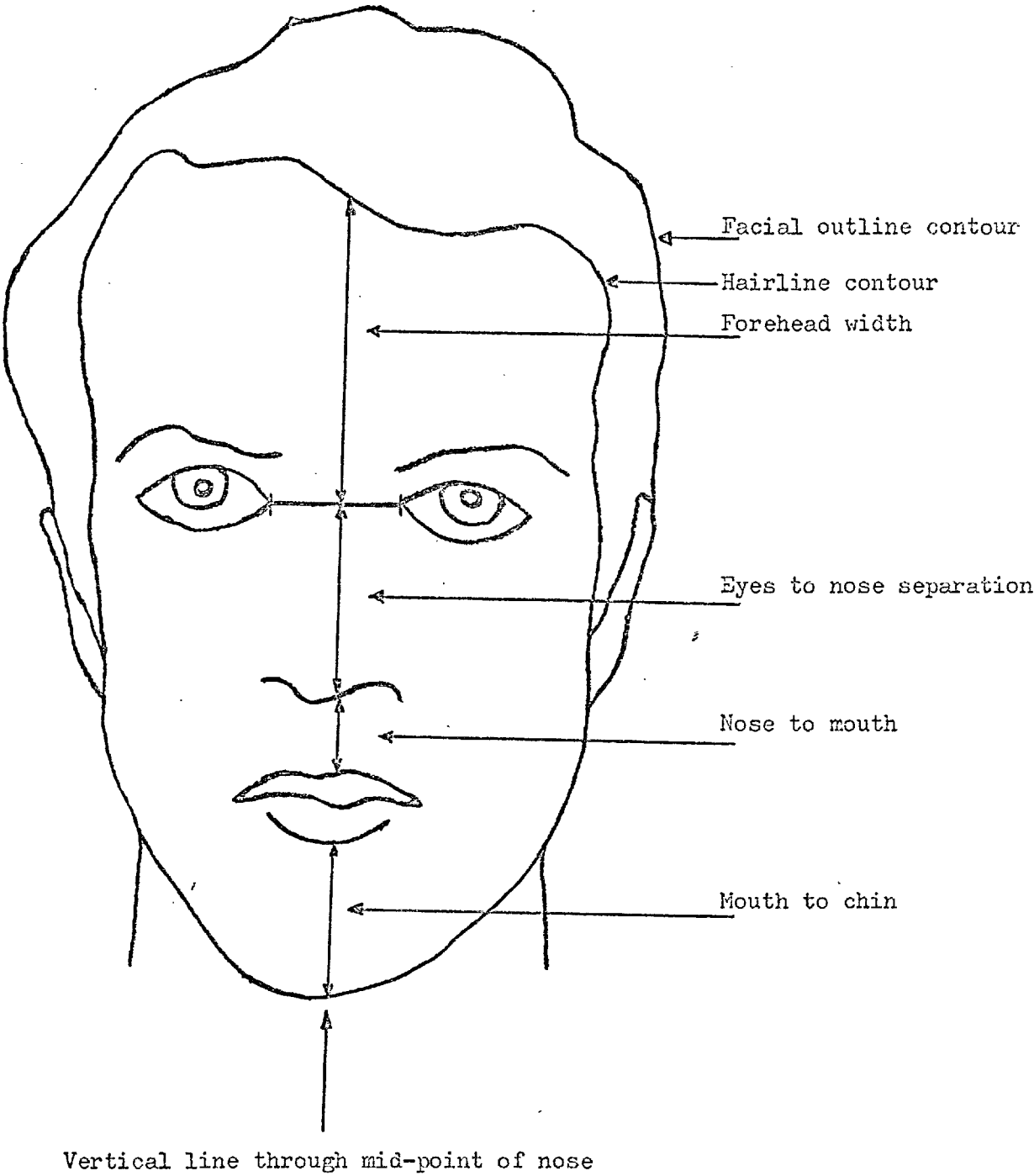
It should be remembered that the author's research deals only with full-frontal face presentations. The range of attributes chosen is not implied to represent a complete set. Rather, these attributes were chosen in the belief that they represented those which would be important in the pattern recognition of facial images. Texture and brightness attributes were not defined and an effort was made to keep them constant throughout the range of experimental stimuli.

In the research which follows, certain assumptions are made. To begin, it is assumed that certain attributes are more 'important' than others. Importance is defined in the following manner: If attribute A belongs to a less homogeneous class than attribute B, then it is considered more important. As an example, if the nose belongs to a less homogeneous class than the mouth, i.e. noses vary more than mouths, then noses are more important than mouths. It is assumed that the attribute belonging to the less homogeneous class will provide a higher contribution to the recognition process.

The second assumption is that the importance that subjects tend to give to particular attributes does not necessarily match the importance that attributes have according to the definition above. In other words, subjective evaluation of importance does not necessarily correspond with defined importance. In the above definition, attribute importance is exclusively dependent on the stimuli and not on the subjects.

Finally, it is assumed that if attribute A is more important than attribute B, then recognition scores will on average be lower when attribute A is masked than when attribute B is masked, all other attributes being visible.

Figure 3.I A line drawing to illustrate the contour attributes and the separations used in the measurement of the feature relation attributes.



3.3 The stimuli and the subjects

3.3.1 The stimulus faces

The majority of the stimulus faces used in this research were presented on black and white positive transparencies. The only exception to this consisted of line drawings which will be described in chapter six. The projection of the stimulus slides onto the presentation television monitor will be discussed in chapter four which deals with the video equipment.

Two types of faces were used in the preparation of the slides. One type consisted of faces prepared from the Photofit kit. This category will be described in the next subsection. The other type consisted of faces photographed by the author. These faces were chosen to satisfy a particular criterion; faces were of male European adults. None of the faces had beards, glasses, moustaches, scars or unusual hairstyles. All faces were photographed in the full-frontal position with no detail showing below the chin contour. Black spotting ink was used to ensure the latter requirement. When photographed, faces were 'neutral' in expression. Finally, none of the faces were known beforehand to any of the experimental subjects.

Decisions concerning the 'neutral' expression and the homogeneity of the stimulus sample were made by the author. The final sample used in the tests was chosen from a larger original population. Faces were those of students at the City University, Imperial College and the Royal Dental Hospital. Front lighting techniques and 35 mm. film were used in all cases. Positive transparencies were prepared by contact printing and these were mounted in slide holders before projection on to the test monitor.

3.3.2 Photofit faces

Photofit is a kit originally developed by Penry (1971) in order to assist Police in suspect identification. The kit is composed of five basic features, each in the form of a photographic segment. These include the chin, eyes, nose, mouth and forehead-hair segments. The kit also provides accessories such as beards and glasses but these

will not be described since they were not used in the author's research. Each feature is mounted in a special holder when a composite face is prepared. There are many variations of each feature. At present, the kit is composed of 169 forehead-hair segments, 82 eye-pair segments, 70 nose segments, 86 mouth segments and 64 chin segments. New variations can easily be added to the kit, especially in the case of the forehead-hair segment as hair fashions change rapidly.

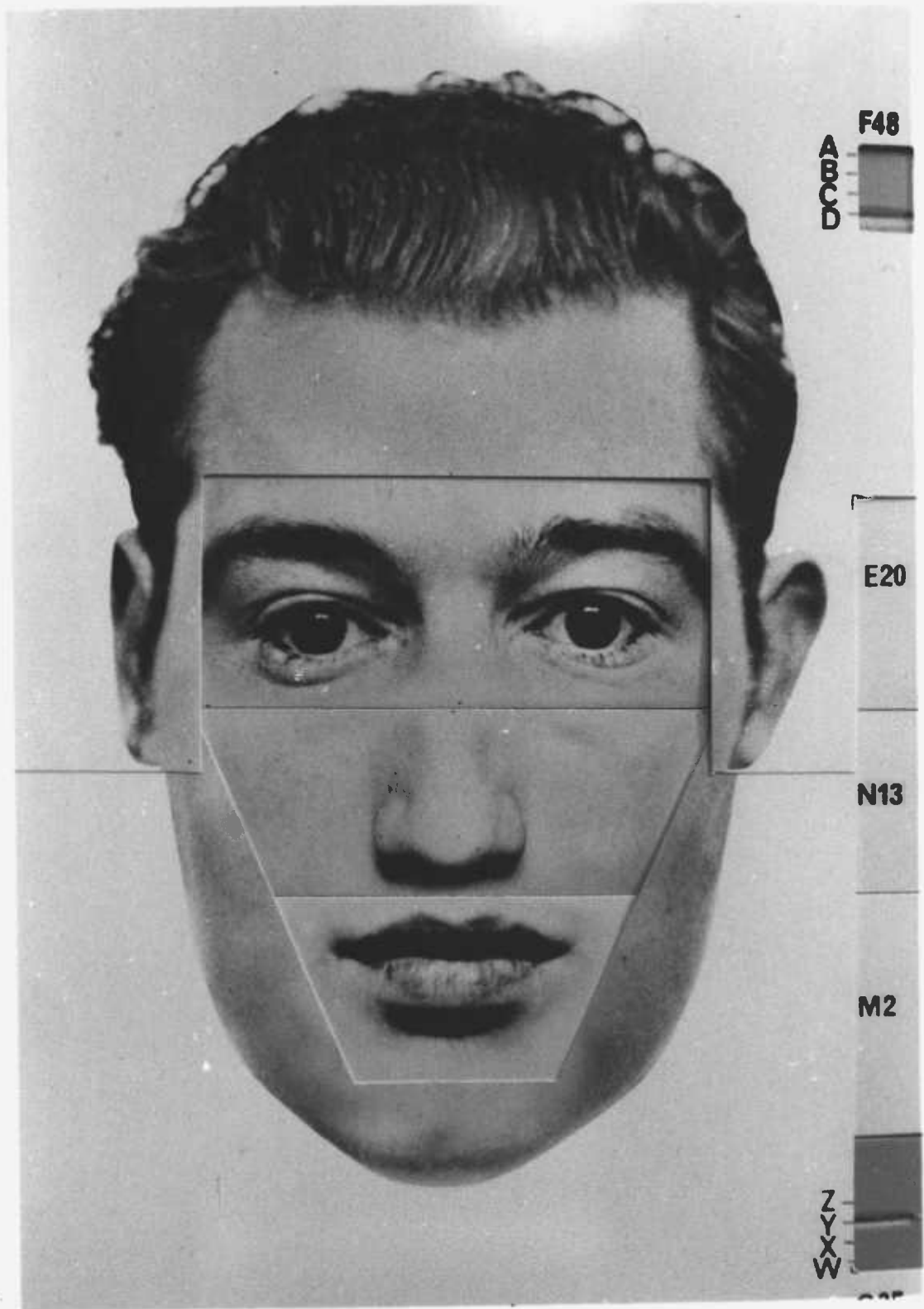
Photofit differs from Identikit in that Photofit faces look more realistic than Identikit faces which are prepared from line drawings. On the other hand, a more realistic face may entail certain drawbacks. For instance, a Photofit face leaves less to the imagination in that it appears to be more specifically determined. No actual research has been undertaken to compare the two kits although a progress report commissioned by the Home Office (1971) indicates that Photofit is being well received by the Police in England.

In the author's case, the Photofit kit has definite advantages for experimental purposes. First, photographs of Photofit faces are virtually the same as ordinary faces except that the boundaries between the different segments are visible. Secondly, a large number of different faces can be made up by interchanging the different variations for each feature. Many subjects reported that the Photofit faces had a 'criminal air' about them. Subjectively, this impression may be consistent but it does not have any effect on the outcome of the author's work since Photofit and ordinary faces were not mixed in the recognition experiments.

A sample of Photofit faces was prepared for experimentation. These were photographed in the same manner as described above. In all tests, subjects made use of a reference board containing the photographs of the faces to be recognised. When Photofit faces were tested, the reference faces were also in Photofit form and these always remained with subjects throughout the entire test.

Identikit faces have been used previously in research: Baker (1967), Bradshaw and Wallace (1971) and Phillips (1973). However, to the author's knowledge, Photofit has never been used. Figure 3.2 illustrates a typical Photofit face.

Figure 3.2 An example of a Photofit face.



3.3.3 The subjects

The majority of subjects used in the author's experiments were either undergraduate or postgraduate students ranging in age from about 18 to 28 years old. Almost all of them were studying in engineering. A very small number were studying in the arts field.

A small percentage of subjects (about 5%) were technicians or secretaries.

The majority of subjects were male. The ratio of male to female was usually 20 to 1. In one case, however, the ratio was 5 to 1. A test of significance was made for this case and no difference was found between sexes.

All students, secretaries and technicians were caucasian with the majority being of British origin.

Subjects who normally wore glasses were asked to wear them for the experiments. All subjects were fluent in the English language and had no difficulty in understanding the experimental procedure.

Subjects were given a standard fee for their participation. No bonuses or incentives were offered on the basis of scores. Subjects were found by advertising or through personal contact.

3.4 Preview of the experiments

The first three experiments are contained in chapter V. The first deals with the recognition of facial images in visual noise. Visual noise will be described in the next chapter. Subjects were given four successive testing sessions with the same stimuli in order to investigate their improvement in progressively noisier visual conditions. The experiment is designed to measure certain variables such as the effect of exposure duration and interval between sessions. The confusions arising from 'incorrect recognition' are discussed.

Experiment 2 investigates recognition under masking conditions. Two types of masking are used and two different portions of the face are masked. In top masking, only the portion below the eyebrow line is visible and in bottom masking only the portion above the tip of the nose is visible.

Experiment 3 also investigates the effect of masking. This time, three different portions of the face are tested for masking: a

horizontal strip containing the eyes and similar strips containing the nose and the mouth.

Experiments 4 and 5 are contained in chapter VI.

Experiment 4 uses line drawings and investigates the effect of changes in the feature relation attributes. Facial images whose feature relations have been altered from a standard format are compared with a reference standard on a similarity criterion.

Experiment 5 makes use of the Photofit kit. Reference faces are prepared and given to subjects. This is followed by the presentation of 80 slides some of which contain features from more than one reference face. The experiment is designed to measure which attributes are used by subjects in a situation where the stimuli are present for only a very short period of time. Facial images are presented for either 40 or 120 milliseconds.

Experiments 6 and 7 are described in chapter VII. Experiment 6 involves the rating of different variations of a Photofit feature segment with respect to a reference segment. Experiment 7 involves the rating of different Photofit faces with respect to a reference on a similarity criterion. The different Photofit faces are prepared from segments rated in experiment 6. A comparison is made of the effects of feature segments when rated individually (experiment 6) and when rated as portions of an entire face (experiment 7). The effects of feature relation variations, using Photofit, are also tested and compared with the results of experiment 4.

CHAPTER IV

THE VIDEO EQUIPMENT

This chapter describes the video equipment used in the author's experiments. Section one presents a general explanation of the function of each of the units involved in the overall system. In section two, a description is given of the visual noise generating unit. Section three illustrates the use of the masking unit and section four explains the function of the frame counter.

4.1 The general layout

The purpose of this section is to describe the general layout of the video equipment used in projecting a stimulus slide on to a test monitor. Several intermediate units exist throughout the system whose purpose is to modify the original slide information in some prescribed manner. The units and their functions are explained. These consist of the noise generating unit, the masking unit and the frame counter unit. A more detailed explanation of each of these units will be given in subsequent sections. The general layout of the equipment is illustrated in figure 4.1.

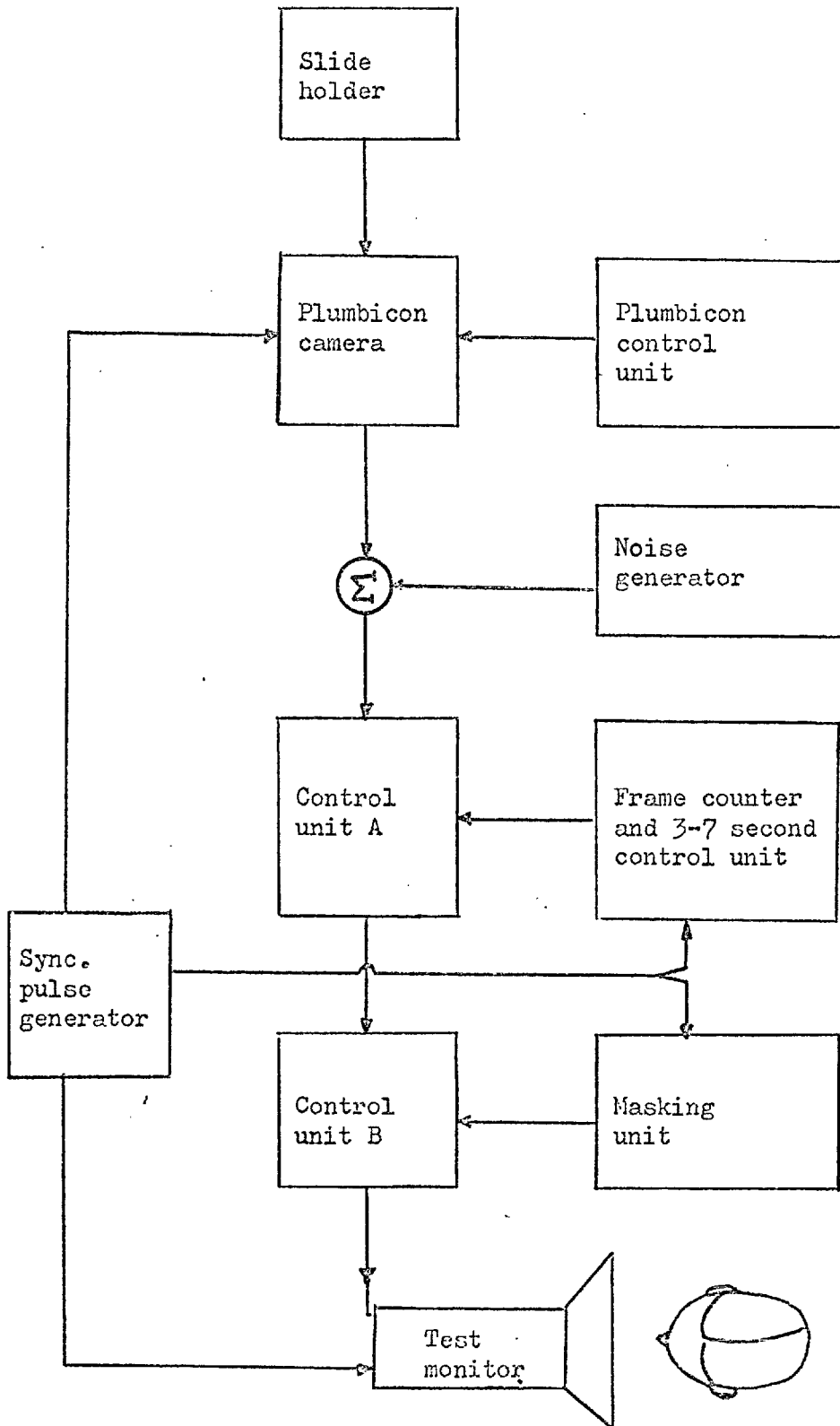
The overall system is seen as a versatile and compact unit capable of many operations commonly performed in experiments dealing with visual perception. The frame counter embodies the principle of the tachistoscope with the restriction that its shortest exposure time is 40 milliseconds, i.e. one television frame. The noise and masking units are seen as useful tools in modifying the stimulus information in ways which might prove to be fruitful. The system has the option of making use of a videotaperecorder which facilitates the experimenter's task and assists in standardising the testing format. Control facilities are available throughout the system to alter the brightness and contrast of the visual image.

The majority of the equipment was originally designed and built by Stapley and a detailed description of most of the system can be found in his thesis (Stapley, 1972). Minor modifications were made to the visual noise generating system by the author and the use of the frame counter provides a new addition to Stapley's equipment.

A general description of each of the units will now be given, starting from the camera end and finishing with a description of the testing cubicle containing the monitor.

The preparation of the stimulus slides was described in section 3.3. These slides were individually mounted in a single-slide projector which was connected directly to a camera. A light bulb in the projector could be adjusted to provide different levels of slide illumination. A camera was used to scan the projected slide image. Plumbicon and Vidicon cameras were available for this function. A Plumbicon camera was chosen because of its faster response and lower

Figure 4.1 General layout of the video equipment



dark current (van Doorn, 1966). It proved to be very flexible over the range of light levels and contrasts of the slide transparencies. Furthermore, it was better suited to the scanning of still pictures than the Vidicon.

The Plumbicon was controlled by a unit with facilities to modify the gain, focus, black level and beam current of the output signal. It also provided for gamma correction and automatic gain control. The camera output was limited to 5MHz. by means of a linear phase LP filter. It was then fed into a gate which provided for the addition of visual noise. This noise, which resembles the snowy effect occurring on a detuned television, was produced by a random noise generator connected to attenuators. The band width of the noise was limited to 5MHz. and the attenuators were used to vary the level of the signal set at the generator. Thus, any ratio of camera signal to noise could be produced at the gate output.

This output could then be modified at control unit A which received control pulses from the frame counter. These timed the output with the result that any integral number of frames would appear at the monitor according to the setting made at the frame counter.

The resulting signal was then processed by control unit B which was regulated by the masking unit. The function of this unit was to mask any desired rectangular area of the final image produced at the monitor. Four switches were available on the unit whose settings defined the area to be masked.

The output from control unit B was then fed to the monitor for viewing by subjects. It could also be stored on a video taperecorder (Ampex 7003). The testing monitor could be adjusted for brightness and contrast and the level measured by means of a photometer.

The entire system was synchronized from one master unit (Philips EL8250/01).

The output at the monitor consisted of the image projection of the slide and this could be modified in the three ways described above: it could be embedded in visual noise, it could be partially masked and it could be adjusted for exposure duration. The subject viewing the monitor was seated in a cubicle measuring approximately ten feet by four feet by eight feet. The monitor was placed eight and a half feet from the subject's eyes, slightly below eye level. The

monitor had a seventeen inch screen (Rank Cintel) and subtended a vertical viewing angle of about seven degrees.

The walls of the cubicle consisted of white dust sheets and lighting was kept constant by means of a variac adjustment and the use of photometer.

Subjects were provided with a reference board which presented the photographs of all the faces which would appear on the monitor. The size of the photographs varied and will be given in the descriptions of the experiments. A procedure sheet was given to each subject in order to standardise the explanation of the experimental procedures. Subjects were also provided with a button panel containing buttons corresponding to the photographs on the reference board. Each of these were pressed according to the face the subject recognised on the monitor. In this way, the amount of interaction between subject and experimenter was minimised.

When seated in the testing cubicle, subjects could not see the experimenter and their only disturbance was the ambient noise of the laboratory which was kept to a minimum.

The experimenter could not see the subjects and he recorded their responses by viewing a panel of lights connected to the button panel. He also was able to follow the stimulus presentation by means of a reference monitor.

4.2 Visual noise

In the previous section, it was mentioned that visual noise could be added to the video signal at the output of the camera. The properties of this noise will now be described. Figure 4.2a illustrates the combination of the noise and video signal before processing at control unit A. A complete diagram can be found in Appendix I.

A General Radio (type 1390B) random noise generator was used as the noise source. The output was specified as white noise of gaussian amplitude distribution. The band width was limited to 5MHz. The frequency spectrum of the noise was tested prior to experimentation. The method of measurement is indirect and makes use of the fact that the noise power ($V_{R.M.S.}^2$) increases linearly with frequency (Pearson, 1965). The noise bandwidth was varied with 8 Butterworth LP filters and a measurement was made of the voltage (R.M.S.) in each case. Results indicated a flat

Figure 4.2a The noise-video gate (complete circuit in Appendix A)

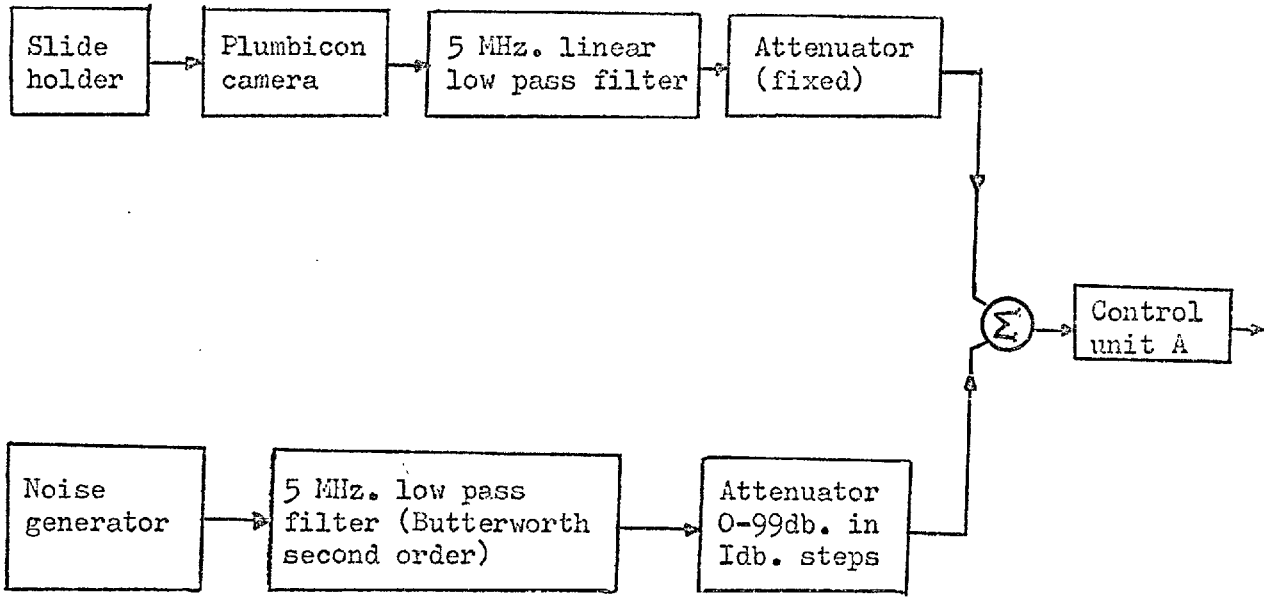
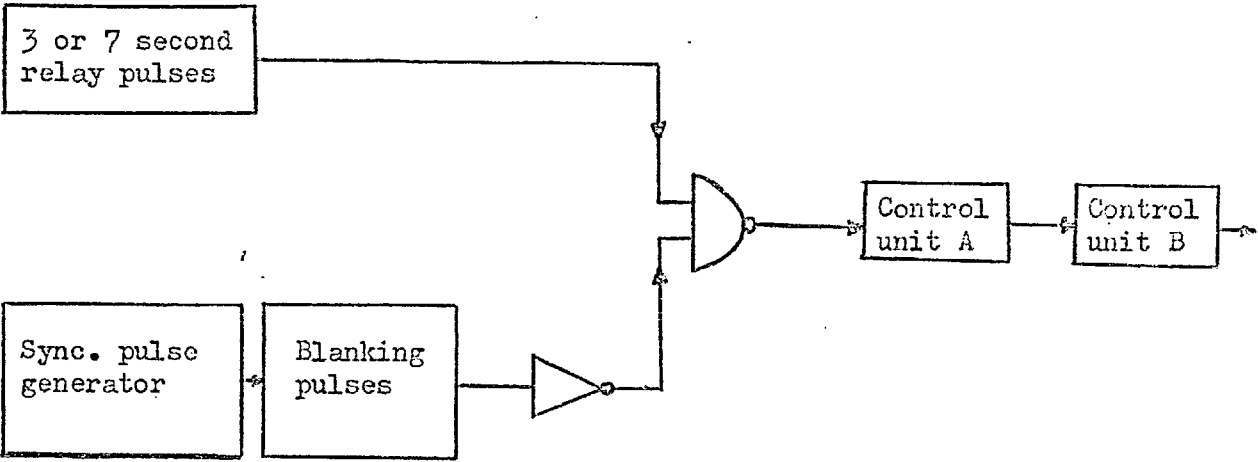


Figure 4.2b The 3-7 second timing gate (complete circuit in Appendix A)

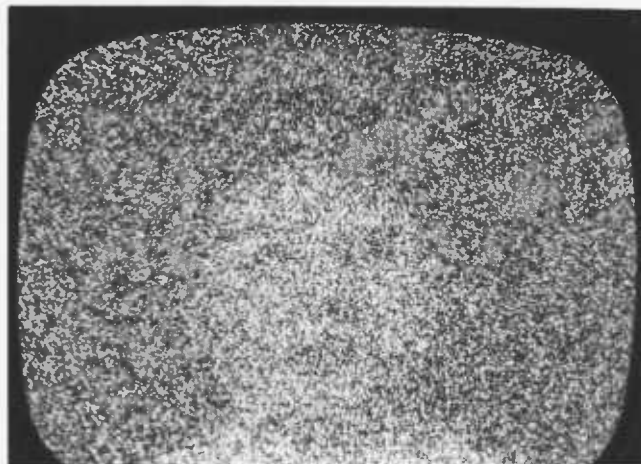
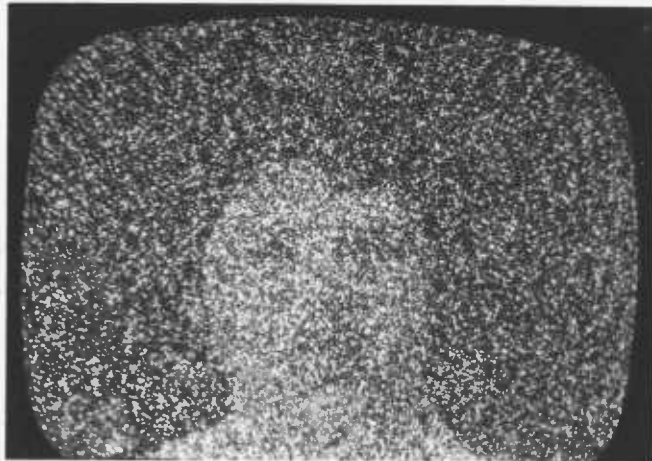
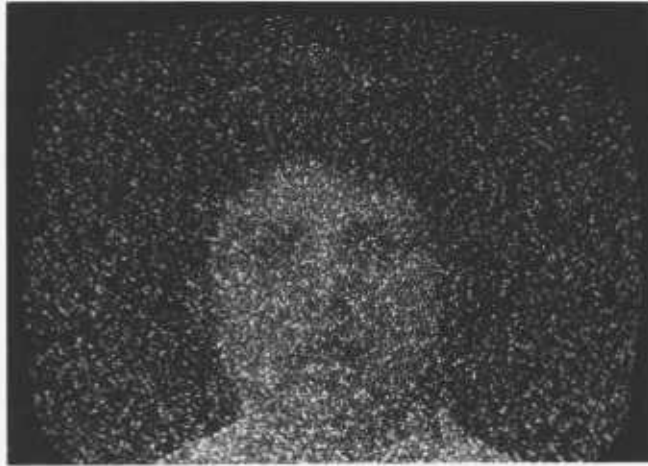


frequency spectrum. The results appear in Appendix I. The amplitude distribution was not tested prior to experimentation. Because of limitations in the amplifiers of the noise generator, the manufacturers warn that an assymetric distribution may occur instead of a truly gaussian one. This was found to be the case in a measurement made by Stapley (1972) just prior to the author's use of the generator. The output from the noise source was connected to attenuators (STC 74600) operating over a range of 99 decibels in steps of 1 decibel (db.). A Butterworth filter was also used to limit the bandwidth to 5MHz.

Figure 4.3 illustrates the use of visual noise at three different settings of the attenuators. The noise level at the source and the video signal producing the image were kept constant in all three cases. Their respective values were 0.2 volts R.M.S. and 1.2 volts peak. The latter value was controlled by the automatic gain control of the Plumbicon camera unit.

The use of visual noise is not new. McKay (1965) describes several areas in which it might prove useful: as a neutral surface for the study of the after-effects of visual movement, as a tool in sensory deprivation and as a stimulus producing subjective effects. In the later context, McKay found one particular effect, the 'maggot' effect, which led him to formulate certain hypotheses about the operation of the neural mechanisms in the visual cortex. Basically, he suggested that there were internal tendencies within the brain to organise visual information in a specific manner. Pearson (1965) in his research dealing with the transmission of visual images over noisy communication channels also made use of noise in his investigations. Visual noise was employed in the research of Marsh (1967) which has already been described in section 2.1. An experiment by Stapley (1972) measuring the effects of noise on the recognition of faces will be described in chapter V. Julesz (1964) in his classic work on binocular perception made use of static, computer generated noise images. A research program has been started by Beevor (-) to investigate the findings of Julesz in the case of dynamic noise presented on television monitors.

Figure 4.3 An example of a face embedded in three different levels of noise.



4.3 The masking unit

The masking unit was built in order that any rectangular area of the monitor screen could be masked. Examples of masking are shown in figure 4.4. The unit consisted of four monostables which respectively determined the horizontal position, width, vertical position and height of the desired masking rectangle. The monostables used the horizontal and vertical drive pulses provided by the synchronizing generator. Their output consisted of adjustable blanking pulses used in defining the masked area. The blanking level of a video signal corresponds to the black level on a monitor screen. Fine and coarse adjustments were available in defining the desired rectangle.

The masking unit could be used at the same time as the noise generator and frame counter units.

4.4 The frame counter

The frame counter was designed to time the output from control unit A. Two decade counters could be set to provide an integral number of frames ranging from one to a hundred. Frame presentation was synchronized by the use of field pulses from the master unit. In this manner, a frame would only begin at the top of the monitor screen. A detailed circuit of the frame counter can be found in Appendix I.

In the first three experiments (chapter V) a frame counter was not used. Three and seven second presentations were required and this was achieved by means of simple RC circuits. The output from these circuits consisted of pulses lasting either three or seven seconds. These were fed to a gate which also received blanking pulses. The gate output then was able to time the output from control unit A. Figure 4.2b illustrates this operation. A detailed diagram can be found in Appendix I.

Figure 4.4 The top and bottom masking used in section 5.2



CHAPTER VNOISE AND MASKING IN RECOGNITION

In this chapter, three experiments are reported. Section one deals with the recognition of facial images embedded in noise. The effects of several variables are investigated. In the next two sections, two experiments examine the effects on recognition of masking certain portions of the facial image. The first experiment also compares the use of 'natural' and 'electronic' blanking. The chapter ends with a summary and general discussion of the results from the three experiments.

5.1 The recognition of facial images in noise

5.1.1 Introduction

The effects of visual noise on the recognition of facial images were originally investigated by Stapley (1972). He was interested in comparing the recognition of facial images embedded in noise with the recognition of faces presented in 'cartoon' form, i.e. with only the outlines visible. His subjects were familiar with the stimuli so that the task was really one of identification rather than recognition (Dember, 1963). His results indicated that the probability of correct identification increased with the signal-to-noise ratio in sigmoidal fashion. The signal-to-noise ratio (S/N) is the ratio of peak signal to noise power.

In the experiment which follows, the author was interested in measuring the effects of several variables. Eight faces were used as the stimuli (figure 5.1). These were photographed in five positions: full-frontal right and left profiles and right and left half-profiles. The eight faces were chosen from an original sample of about thirty according to a criterion of homogeneity. Subjects were provided with a reference board containing the eight faces in each of their five views. They were told that 24 slides would be presented to them in a random order. These would consist of the eight faces each shown in three positions: full-frontal, left half-profile and right profile.

A total of twenty subjects was used in this exploratory study. Each subject was tested with the same stimuli on four successive occasions. The order of presentation of the stimuli was randomized for each of the four sessions. Half the subjects were tested at one day intervals and the other half at one week intervals.

Stimuli were presented in the following manner; a face embedded in noise would appear for either three or seven seconds. Half the subjects received three second presentations and the other half received seven second presentations. If the subject did not respond to the stimulus by pressing an appropriate button corresponding to the face he recognised, then the same stimulus would reappear after a two second delay, but this time with less noise. This process would continue with less noise at each successive presentation until the subject pressed a button. At this point, the following slide would be presented without

Figure 5.1 The eight stimulus-faces of experiment 1. L.H. column - 1,3,5,7; R.H. column - 2,4,6,8.



any indication to the subject as to whether his previous choice had been correct.

The noise settings were controlled by the attenuators described in chapter four. The actual image signal was always kept constant at 1.2 volts peak. Only the noise level was varied. The setting at the noise generator was fixed at 0.2 volts R.M.S. and this level could be attenuated in steps of one decibel. When a stimulus was first presented to a subject, the noise present in the image was such that recognition would be unlikely. The level of this noise had been established beforehand in a pilot experiment. With successive presentations of the same image, the noise was progressively attenuated in one decibel steps until a button was pressed by the subject. The initial starting level of the noise had to be increased throughout each session and from session to session in order to accommodate for the improvements in subject performance.

The photographs of the subjects' reference board consisted of eight rows each containing the five views of one face. Each photograph measured about one and a half inches by one inch. The subject was free to place the reference board in any part of the cubicle. Ambient lighting was four and a half foot-candles. This value was kept constant throughout subsequent experiments. Screen brightness was measured by placing a photometer flush against the centre of the screen. A value of ten foot-candles was used in this and subsequent experiments. Camera and monitor settings were kept fixed throughout the entire test.

Each subject was given the following written instructions: "You will be given two minutes to study the photographs on the reference board placed in front of you. Testing will then begin. A face will be presented on the monitor for a period of three (seven) seconds. This face will be embedded in visual noise. If you recognise it, you are to press the appropriate button corresponding to the numbered photographs. After three (seven) seconds, the face and noise will disappear for about two seconds. If you have not pressed a button, the face will reappear, but this time with less noise. Once again the presentation will last for three (seven) seconds and then disappear for another two seconds. This process will continue, with less noise at each successive presentation, until you have pressed a button at which point the next face will be presented. You will not be told if you have

made any mistakes. A total of 24 slides will be presented and you will be shown a test slide at the beginning of the test. This slide will not count but is simply there to give you an idea of the task. The test slide face is not one of the eight faces on the reference board."

Six variables were investigated in this experiment and these appear in table 5.1.

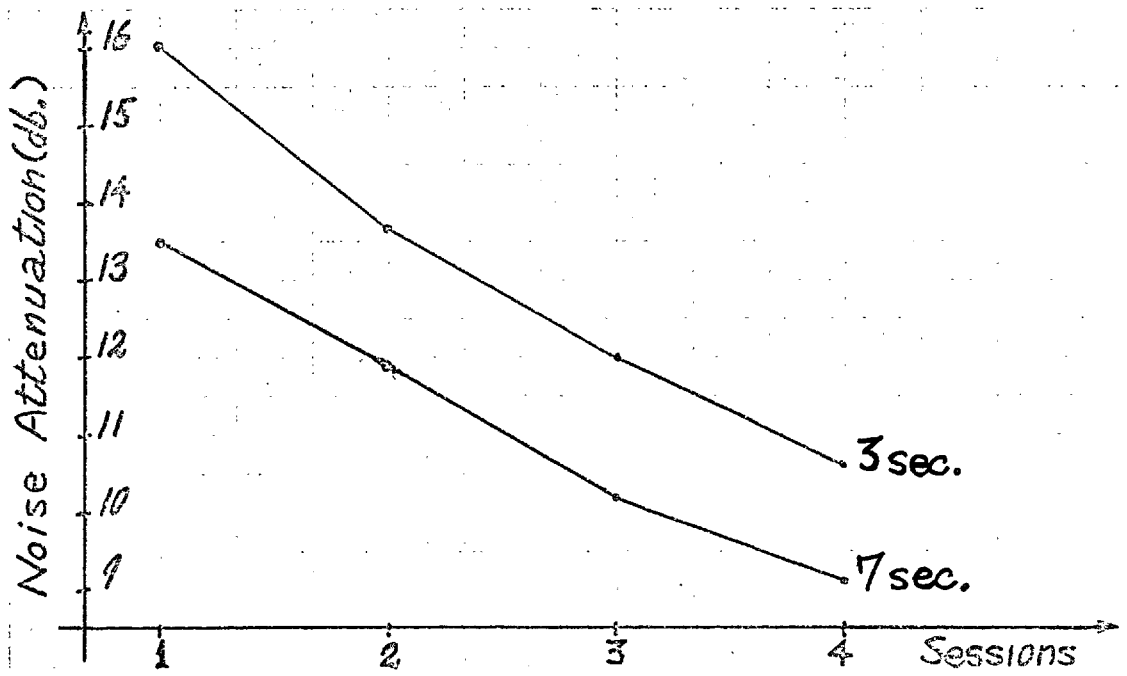
5.1.2 Results

A BMD02V programme was used to perform the analysis of variance on the results. The 'face' and 'subject' variables were considered random effects with the remaining variables considered fixed. The breakdown of the expected mean squares into components of variance was done according to Edwards (1968) and Schultz (1955). Values for the degrees of freedom and F ratios were measured according to Snedecor and Cochran (1967). A discussion on the method can also be found in Cochran (1951). Significant main effects and interactions are presented in table 5.2. The level of significance was set at 0.01 and the test was two-sided.

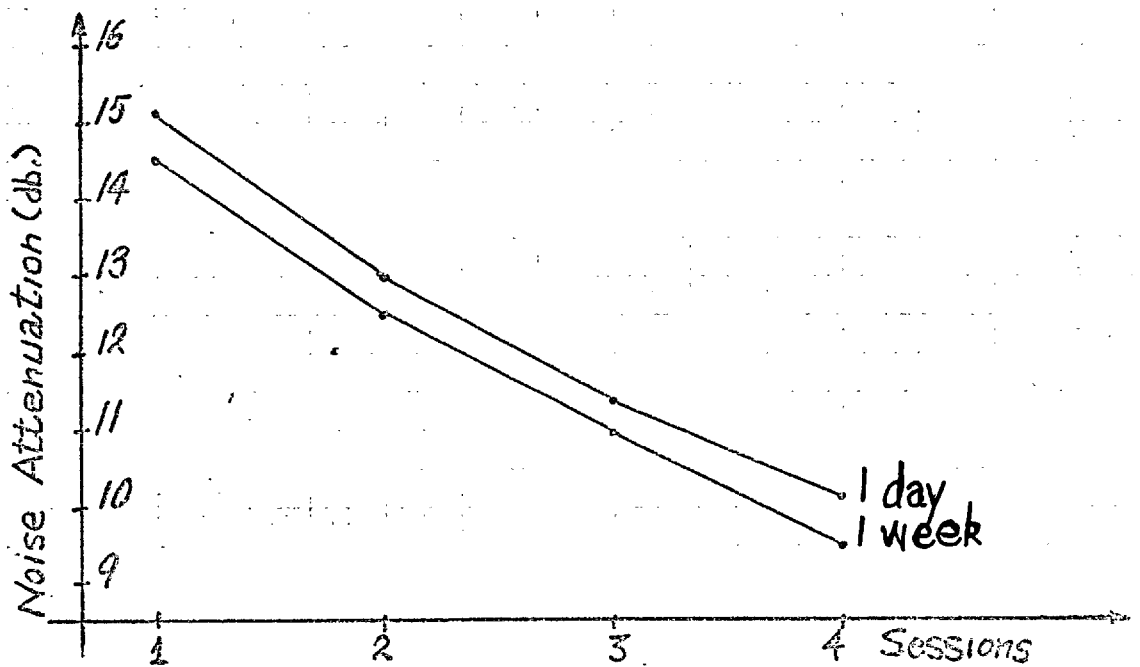
'Duration' and 'interval' were not significant in this experiment. Figure 5.2 illustrates the average performance in recognition over the four sessions as a function of these two variables. It should be noted that the ordinate axes for the two graphs are scaled in decibels of noise attenuation. The downward slope of the curves implies that recognition occurred at progressively higher noise levels, i.e. with less noise attenuation.

The 'sessions' variable was found to be highly significant and a Duncan multiple range test was applied at the 0.01 level in order to check whether the improvement in performance was in fact linear across the four sessions. This was found to be the case. Furthermore, analysis of regression indicated a strong linear component for the 'sessions' variable. It can be seen in figure 5.2 that the performance curves will eventually level off asymptotically. This is to be expected since a noise attenuation threshold will eventually be reached at which point the noise level will make recognition impossible except at the chance level. This threshold will correspond to complete noise saturation of the image.

Figure 5.2 Learning curves for both durations and intervals of experiment I.



Curves for the 3 and 7 second exposure durations.



Curves for the 1 day and 1 week intervals.

Table 5.1 The variables tested in Experiment 1

VARIABLE	ABBREVIATION	LEVELS OF EACH VARIABLE
INTERVAL	I	1 day, 1 week
DURATION	D	3 seconds, 7 seconds
SUBJECTS	S(DI)	5 subjects in each case
SESSIONS	C	Sessions 1, 2, 3 and 4
VIEWS	V	Full-frontal, half-profile, profile
FACES	F	8 different faces

Table 5.2 Analysis of variance for Experiment 1

Effect	Mean Square	F ratio	Degrees of Freedom		Level of Significance
			n ₁	n ₂	
D	1691.25	5.29	1	18.22	0.1
I	164.50	0.54	1.08	18.06	
DI	11.72	0.06	2.37	16.94	
S(DI)	298.74	47.49	16	112	0.01
C	2155.99	55.38	3	62.3	0.01
DC	24.53	1.13	3.76	60	
IC	2.91	0.25	12.8	57.3	
DIC	13.79	0.73	4.52	53.2	
CS(DI)	21.60	6.97	48	336	0.01
F	211.47	33.62	7	112	0.01
FD	22.07	3.51	7	112	0.01
FI	20.33	3.23	7	112	0.01
FDI	9.00	1.43	7	112	
FS(DI)	6.29	3.01	112	672	0.01
FC	17.39	5.61	21	336	0.01
FDC	2.91	0.94	21	336	
FIC	2.34	0.75	21	336	
FDIC	1.47	0.47	21	336	
FCS(DI)	3.10	1.48	336	672	0.01
V	43.37	1.75	2	20.3	
VD	1.33	0.65	20.75	46.4	
VI	2.49	0.79	8.92	46.4	
VDI	2.03	0.77	11.47	46.6	
VS(DI)	4.48	1.62	32	224	0.1
VC	2.53	0.46	20.25	66	
VDC	0.88	0.46	67.85	91.3	
VIC	1.80	0.89	28.02	148	
VDIC	1.20	0.79	45.08	164	
VCS(DI)	2.05	0.98	96	672	
VF	21.95	7.92	14	224	0.01
VFD	1.81	0.65	14	224	
VFI	2.16	0.78	14	224	
VFDI	1.75	0.63	14	224	
VFS(DI)	2.77	1.33	224	672	0.02
VFC	8.09	3.87	42	672	0.01
VFDC	4.42	2.11	42	672	0.01
VFIC	2.34	1.12	42	672	
VFDIC	2.10	1.00	42	672	
VFCS(DI)	2.09				

The 'views' variable was not found to be significant. This is attributed to the fact that all views were available to subjects on the reference board. A significant effect might have been obtained if only the full-frontal views had been available, but this possibility was not investigated (in subsequent tests, only the full-frontal case was studied). The 'face' variable was significant indicating that certain faces are easier to recognise than others.

Several first order interactions were found to be significant. The $F \times I$ interaction was examined and all 8 faces were found to be recognised at higher noise levels with the 1 week interval than with the 1 day interval. With the $F \times D$ interaction, all faces were recognised at higher noise levels with the 7 second exposure than with the 3 second exposure. The significance of the $C \times S(DI)$ interaction was examined and all subjects tested (within DI) were found to recognise the stimuli at progressively higher noise levels with each successive session. Similarly, the $C \times F$ interaction revealed that all 8 stimuli were recognised at higher noise levels with each successive session. However, subjects did recognise different faces at different noise levels and different subjects recognised the same face at different noise levels. This is evidenced by a study of the $S(DI) \times F$ interaction and was to be expected because of the significance of the two respective main effects. Finally, the $V \times F$ interaction revealed that certain stimuli were better recognised in certain views than in others. Four faces were best recognised in full-frontal view, 3 in profile and one in half-profile. However, the differences in each case were small.

The findings presented so far are based on the use of the entire set of data. Each data point consisted of the noise attenuation level at which recognition occurred for a particular subject examining a particular stimulus. This includes both 'correct' and 'incorrect' recognition responses. In this test, 9.7 per cent of the responses were incorrect, i.e. 186 confusions out of 1920. The number of confusions decreased with each session and was about the same for the two intervals tested. However, confusions for the seven second duration were two and a half times more frequent than for the three second duration (significant at the 0.001 level using the Binomial Test). Furthermore, confusions varied considerably amongst the different faces. For example, there is the paradoxical result that face 6 was recognised as another face on 38 occasions but that only

once was another face recognised as face 6. In other words, it would appear that face 6 was similar to certain other faces but that none of them were similar to face 6. Table 5.3 presents the confusion matrix.

Summary of results

- (a) There was no significant difference between scores for the one day and one week intervals between sessions.
- (b) Recognition improved linearly with progressively higher noise levels at each successive session.
- (c) There was no significant difference in recognition thresholds for the two exposure durations although more than twice as many confusions occurred in the case of seven second exposures (significant at 0.001 level).
- (d) Recognition scores for the three different stimulus views were not significantly different although some faces were more often recognised at the higher noise levels with a particular view.
- (e) Different faces were recognised at different noise levels by the same subject and different subjects recognised the same face at different noise levels.
- (f) Certain pairs of faces were often confused whilst others were not at all.

5.1.3 Discussion

Intuitively, most of the results of this experiment are not surprising. For instance, it has always been assumed that certain faces were easier to recognise than others presumably because they were less liable to be confused. One might expect such faces to be recognised at higher noise levels. In table 5.4, a comparison is made between the average noise attenuation level at which a face was recognised and the number of times it was not recognised correctly. A Spearman rank correlation test set at the 0.05 level indicated a significant correlation between both sets of results. Thus, the hypothesis that the less confused faces are 'easier' to recognise is satisfied to a certain extent, at least within the framework of the present experiment (correlation coefficient was 0.81).

An important result of this experiment concerns the relative confusability of different pairs of faces. While certain pairs were never

Table 5.3 The Confusion Matrix for the Eight Stimulus Faces

P: Stimulus Presented

R: Subject Responses

R \ P	1	2	3	4	5	6	7	8
1	215	0	0	0	1	10	0	5
2	1	218	4	4	0	5	14	0
3	0	4	228	2	1	2	4	0
4	3	2	0	208	4	2	8	6
5	7	3	5	8	228	9	1	0
6	1	0	0	0	0	202	0	0
7	5	13	3	3	2	6	211	5
8	8	0	0	15	4	4	2	224

Table 5.4 A Comparison Between the Average Noise Attenuation Levels at which Faces Were Recognised and the Number of Times They Were Recognised Incorrectly

FACE	1	2	3	4	5	6	7	8
AVERAGE RECOGNITION LEVEL (DB.)	13.85	11.83	11.79	12.00	10.94	12.81	12.60	11.20
NUMBER OF INCORRECT RECOGNITIONS	25	22	12	32	12	38	29	16

confused, others were on several occasions. It seems reasonable to assume that confusions occur on the basis of 'similar' features and that certain features may be more responsible than others for these confusions. For instance, a feature which is relatively similar for most faces is unlikely to be the principle source of a confusion. However, a feature which usually varies considerably from face to face might lead to confusion in a particular instance where it does happen to be similar for two faces. In searching for the latter type of feature, it is important to examine both confused and non-confused pairs of faces. One would expect such a feature to be relatively similar in the case of two confused faces and quite distinct in the case of non-confused faces. By definition, a feature which varies considerably from face to face is more important than one which is relatively similar for most faces. For that reason, it follows from the above argument that the features which are most responsible for confusions are also the most important ones.

All 8 faces used in the experiment were studied by the author with a view to establishing those features most responsible for the confusions which arose. The arguments discussed above were applied in scrutinizing both confused and non-confused pairs of faces and it was concluded that contour attributes were on the one hand the most similar features in confused faces and on the other hand the most distinct features in non-confused faces. Therefore, it is suggested that contour attributes are the most important features for recognition. This suggestion is obviously very tentative at the present stage but does provide guidelines for formal investigation in subsequent experiments. These will investigate the relative importance of several facial attributes.

An unexpected result in this experiment concerned the number of incorrect responses for the two exposure durations. Subjects receiving seven second exposures made more than twice as many incorrect responses as three second subjects. This finding is difficult to explain on the basis of stimulus or experimental variables. However, it may be related to the subjects' confidence in making decision responses. It is suggested that the seven second subjects had more confidence in their 'initial impressions' as a result of the longer exposure durations. This would imply the following decision process: when a stimulus is presented, the subject acquires a 'first impression'. Confidence in this 'first impression' increases with time. If the stimulus lasts for seven seconds, con-

confidence may be strong enough to entail a response. On the other hand, confidence may not be high enough after only three seconds. Consequently the subject waits for the next exposure, at which time he may form a new 'impression'. It will be seen from figure 5.1 that the three second group did not in fact perform as well as the seven second group in terms of noise attenuation thresholds even though the difference was only significant at the 0.1 level. This is consistent with the suggestion that they waited for more exposure presentations than the seven second group because of their lack of confidence. If the above explanation is to be accepted, then it follows that 'first impressions' may occasionally be unreliable.

The main findings of this experiment were summarised in the previous section. The discussion examined 3 more points, resulting in the following suggestions.

- (a) It is proposed that the ease of recognition of a face (as measured in terms of noise levels) is related to its distinctiveness, i.e. its 'lack of confusability' (significant at the 0.05 level).
- (b) Contour attributes are suggested as the most important attributes of a facial image in the recognition process. This suggestion will be tested formally in subsequent experiments.
- (c) It is suggested that subject confidence based on 'first impressions' may account for the greater number of incorrect responses which occurred in the seven second group.

The purpose of experiment 1 was to raise certain questions rather than to test specific predictions. Furthermore, because of the small number of subjects tested, results must be viewed cautiously. The question of attribute importance was raised and it was suggested that contour attributes were more important than other types. This suggestion will be explored more systematically in subsequent experiments. For instance, experiment 2 which follows examines the effects on recognition of partially masked facial images.

5.2 The first masking experiment

5.2.1 Introduction

In experiment 1, face recognition was tested in the presence of visual noise. As a result of this constraint, certain pairs of faces were

found to be confused more than others. From a study of the confused and non-confused pairs, it was suggested that contour attributes were the most important type of attribute for recognition.

In the present experiment, the effect of masking two different portions of the face will be examined (figure 4.4). In top masking, the lower portion of the facial outline contour is visible. In bottom masking, the upper portion of this contour is visible as well as the forehead-hairline contour. It is predicted that bottom masking will result in higher recognition scores than top masking. The prediction is based on the fact that more contour information is present in bottom masking than in top masking.

Bottom masking covers all information below the tip of the nose. Top masking covers all information above the eyebrow line. Obviously, more than just contour information will be available in both cases. In top masking, all discrete feature attributes and all feature relation attributes except forehead width are visible. In bottom masking, the nose and eyes are visible as well as the forehead width and the eyes to nose separation. Because of the presence of all types of attributes in both masking situations results will not exclusively reflect the importance of contour attributes. The role of other attributes will also have to be considered. It should be pointed out at this stage that four attributes will be visible in both masking situations: the eyes, the nose, the eye to eye separation and the eyes to nose separation.

Two types of masking will be used in the present experiment. The first, which has already been described in section 4.3, is 'electronic' masking and the second is 'natural' masking. Natural masking was obtained by asking those people who were photographed as the stimuli to hold a black piece of cardboard in the appropriate masking positions. In other words, the resulting slides already contained the desired masking. The two types of masking were used in order to compare their effects. It was suggested that subjects unfamiliar with electronic masking might be affected by it in some unpredictable manner.

Method. Sixty subjects were tested in this experiment: fifty male and ten female. Eight new stimulus faces were chosen from a large sample of students at the Royal Dental Hospital. These eight stimuli were presented in each of four conditions: top and bottom electronic masking and top and bottom natural masking. In the presentation of these 32 slides, natural and electronic masking alternated, all top masking slides preceded bottom

masking slides and the order of presentation of the eight faces was randomised for the first sixteen slides, the reverse order being used for the last sixteen. The presentation format described was used for half the subjects with the remaining half receiving the reverse format.

Slides were presented only once for a period of three seconds. A noise level corresponding to 35 decibels of noise attenuation was also present (same conditions as in experiment 1). This noise level was chosen in a pilot experiment in order to ensure an overall recognition score of about 70%. At this level, all facial features not covered by masking were clearly visible although not perfectly defined.

Subjects were given the following written instructions: "You will be given two minutes to study the photographs on the reference board placed beside you. You will then be shown a sequence of 32 slides. Each slide will be presented for three seconds only. After viewing the first slide, you must decide which of the eight stimuli has been presented. When you have reached your decision, you press the button corresponding to the appropriate stimulus. The same procedure is then followed for the remaining slides.

Each of the slide faces will be partially covered either at the top or bottom and the image will also be obscured by visual noise. A test slide will be presented at the beginning of the sequence to illustrate the effect of noise. This slide will not count in the test. The test face will not be one of the eight stimulus faces. The reference board will remain with you throughout the experiment."

The photographs on the reference board measured two inches by three inches.

5.2.2 Results

Five variables were examined in this experiment. These appear in table 5.5. The same analysis techniques were used as in the previous experiment. Data was in the form of binary scores: a 'one' was given for correct recognition and a 'zero' for incorrect recognition. Significant main effects and interactions are presented in table 5.6. The sex variable was analysed separately because of the uneven number of subjects for each sample.

The type of masking was not found to be significant in this experiment. Average percentages of recognition were 64% for natural

Table 5.5 The variables for Experiment 2

VARIABLE	ABBREVIATION	LEVELS OF EACH VARIABLE
ORDER	O	The two orders of presentation
MASKING	M	Top and bottom masking
TYPE OF MASKING	T	Natural and electronic
SUBJECTS	S(O)	30 subjects for each order
FACES	F	The eight stimulus faces

Table 5.6 The analysis of variance for Experiment 2

Effect	Mean Square	F ratio	Degrees of Freedom		Level of Significance
			n ₁	n ₂	
O	0.05208	0.391	22.69	55.32	
S(O)	0.492	2.46	58	406	0.01
M	41.41875	25.798	1	10.58	0.01
T	0.20833	0.804	3	13.47	
F	2.79137	13.957	7	406	0.01
OM	0.75208	1.446	1.62	21.96	
OT	0.13333	0.945	4.61	19.95	
OF	0.15327	0.766	7	406	
MS(O)	0.303	1.485	58	406	0.05
TS(O)	0.128	0.837	58	406	
FS(O)	0.200	1.212	406	406	
MT	0.03333	0.839	35.44	39.95	
MF	1.31042	6.424	7	406	0.01
TF	0.32143	2.101	7	406	
OMT	0.30000	1.196	2.4	18.01	
OMF	0.35804	2.340	7	406	0.05
OTF	0.17500	1.144	7	406	
MTS(O)	0.152	0.921	58	406	
MFS(O)	0.204	1.236	406	406	
TFS(O)	0.153	0.927	406	406	
TMF	0.08452	0.512	7	406	
OMTF	0.2369	1.436	7	406	
MTFS(O)	0.165				

masking and 66% for electronic masking. There was, however, a very significant difference between top and bottom masking. Average percentages of recognition for these two cases were 51% for top masking and 80% for bottom masking. Faces were also found to be significant with scores ranging from 50% to 76%.

The sex variable was examined by means of the Mann-Whitney U test (Siegel, 1956). No significant difference was found between the two sexes ($z = 0.26$).

Only two first order interactions were found to be significant at or below the 0.05 level. The significance of the $M \times F$ interaction indicates that the difference between scores for top and bottom masking varies according to which face is being presented. For example, the difference for face 4 was 57% whilst that for face 1 was 11% although in all cases bottom masking yielded the higher scores. The $M \times S(0)$ interaction indicated that the difference between scores for top and bottom masking varied, depending on the subject being tested.

5.2.3 Discussion

It was suggested in experiment 1 that contours were the most important facial attributes. On the basis of this suggestion, it was predicted that bottom masking would produce the higher recognition scores because it obscured less contour information. This prediction is analogous to an assumption made in section 3.2, namely that the masking of the more important attributes would yield the lower recognition scores.

The prediction was verified although the presence of other attributes in both masking situations must also be taken into account. Their relative contributions to the recognition process must be examined in both masking cases.

It was mentioned that four attributes were visible at all times: the eyes, the nose, the eyes to nose separation and the eye to eye separation. If any of these were to be considered more important than the contour attributes, it would be difficult to explain the 29% difference between the two masking cases. Attributes which were visible only in the bottom masking situation are the contours already mentioned and the forehead width. In the present experiment, it is impossible to evaluate the importance of this last attribute. Attributes which were visible only in the top masking situation are the mouth, the eyes to mouth separation, the

nose to mouth separation, the mouth to chin tip separation and the lower portion of the outer facial contour. On the basis of the results, it can be assumed that these attributes are less important than the contours visible in bottom masking. Therefore, it is concluded that the contours in the upper half of the face are the most important attributes for recognition, keeping in mind that the forehead width remains to be evaluated.

This conclusion receives strong support from the evidence presented in chapter II. It was stressed on many occasions that contours were high in information content and very easy to perceive. The boundary line between the hair and the forehead provides an excellent example of a highly defined contour line. The same applies to the outer contour of the face.

Summary. The analysis of variance indicated no significant difference between (a) the two orders of presentation, (b) the two sexes, and (c) the two types of masking. However, differences were found in (d) the particular face being presented, and (e) top and bottom masking. The $M \times F$ and $M \times S(O)$ interactions were also significant.

As a result of the discussion, one more finding can be added. (f) The contour attributes in the upper half of the face constitute the most important facial features for recognition. However, it must be added that the importance of the forehead width has not yet been evaluated.

The results of this experiment do not tell us which of the three discrete features is the most important. This question will be examined in the experiment which follows.

5.3 The second masking experiment

5.3.1 Introduction

In the previous experiment, two different portions of the face were masked. Results indicated that the most important facial attributes were the contours in the upper half of the face. No estimate of the relative importance of the other attributes could be given. In this section, the effects of masking the eyes, the nose and the mouth will be examined. It is, of course, difficult to mask any of these discrete features without destroying the information provided by other attributes. For instance, if the nose is masked, two other attributes (the eyes to

nose separation, the nose to mouth separation) can no longer be used in recognition. One way around the problem would be to use a mask of exactly the same shape as the nose. However, this would not eliminate one of the most important characteristics of the nose, namely its shape. Figure 5.3 illustrates the three masked areas tested in the present experiment. Electronic masking was used in all cases.

Two types of presentation were used for all subjects: wide masking and narrow masking. The latter is illustrated in figure 5.3 and the former is identical except that the masking strips extend over the boundaries of the face in all three cases. Wide masking was used because it divided the face into two separate portions unlike narrow masking which did not destroy the 'Gestalt' unity of the stimulus. In all presentations, only one of the three areas was masked and the 'no masking' (control) situation was also tested.

Method. The stimuli consisted of the eight faces used in the noise experiment. This time, however, only the full-frontal views were tested and the reference board presented only those views. Photographs measured three inches by four inches.

Thirty-six subjects participated in this experiment. Each subject was tested in three separate sessions, one for each masked area. 24 slides were presented at each session: eight with no masking, eight with wide masking and eight with narrow masking. A randomised block design was used to ensure a balanced effect of the variables. The four variables tested appear in table 5.7. Once again, visual noise was present in the stimuli. Different levels were used for each of the three sessions. These corresponded to the fifteen, thirteen and eleven decibel levels tested in the noise experiment. In that experiment, it was found that the average improvement with each successive session was approximately two decibels starting from fifteen at the first session.

Obviously, it is undesirable that visual noise should be used in this masking experiment. However, because of the enormous redundancy of information present in a facial image and because of the lesser importance of the discrete feature attributes with respect to the contour attributes, it becomes necessary to attenuate all facial information in order that a sensitive measure can be made of the relative importance of the discrete features. Otherwise, it is felt that subjects will ignore these attributes and base their recognition on the more important visible attributes.

Figure 5.3 Narrow masking of the eyes, nose and mouth strips



Table 5.7 The variables of Experiment 3

VARIABLES	ABBREVIATION	LEVELS OF THE VARIABLES
MASKED AREA	A	The eyes, nose and mouth areas
MASKING	M	Wide masking, narrow masking, no masking
FACES	F	The eight stimulus faces
SUBJECTS	S	The thirty-six subjects

Table 5.8 The analysis of variance for Experiment 3

Effect	Mean Squares	F ratio	Degrees of Freedom		Level of Significance
			n ₁	n ₂	
A	1.63	4.35	2	33	0.05
M	1.05	2.21	3	37	
F	3.52	12.15	7	245	0.01
S	1.16	4.00	35	245	0.01
AM	0.42	1.33	7	63	
AF	0.26	1.72	14	490	
AS	0.15	0.97	70	490	
MF	0.32	2.29	14	490	0.01
MS	0.21	1.51	70	490	0.02
FS	0.29	2.01	245	980	0.01
AMF	0.28	1.92	28	980	0.01
AMS	0.15	1.06	140	980	
MFS	0.14	0.99	490	980	
AFS	0.15	1.06	490	980	
AMFS	0.14				

If this were to be the case, no measure of the relative importance of the discrete features could be obtained.

In all three sessions, subjects were given the following written instructions: "You will be given two minutes to study the eight photographs placed on the reference board beside you. The test will then begin and you will be presented with a sequence of 24 slides. Each slide will last for only three seconds. After each presentation, you must decide which of the eight faces was presented. When you have made your decision, you press the appropriate button at which point the following slide will appear.

All of the slides will be covered in noise. Some will also have black areas covering certain portions of the face. A test slide will be presented before the start of the experiment to illustrate these effects. This slide will not count and the face will be different from the eight photograph faces. You will keep the reference board with you throughout the test."

5.3.2 Results

Once again, a BMD02V programme was used in the analysis of variance and data was in binary form. Table 5.8 presents significant main effects and interactions for a two-sided test at the 0.05 level.

No significant effect was found for the three masking situations which were 'no masking', wide masking and narrow masking. However, a Duncan multiple range test at the 0.1 level indicated a significant difference between narrow masking and 'no masking' and a near-significant difference between wide masking and 'no masking'. Recognition percentages for the three cases were 78.5% for 'no masking', 73.5% for wide masking and 71.8% for narrow masking. It will be seen that the effect of separating the facial image into two separate portions in wide masking was negligible with respect to the narrow masking case. In fact, performance was slightly better in the former case.

The 'area masked' variable was significant at the 0.05 level. Average percentages of recognition for the three areas, were as follows: 70.2% for the eyes area, 74.6% for the nose area and 78.9% for the mouth area. These values were averaged over the three types of masking which means that a common component from the 'no masking' situation was present in all three cases. If this component is removed, a more realistic esti-

mate of the average is obtained: 66% for the eyes area, 74% for the nose area and 78% for the mouth area. On the basis of these results, it can be concluded that the most important of the three areas for recognition is the eyes area. This is followed by the nose area and the mouth area. A Duncan multiple range test, set at the 0.05 level, indicated a significant difference between the eyes and nose areas and the eyes and mouth areas but not the nose and mouth areas.

The 'face' variable was again significant as expected. Furthermore, three first order interactions, $M \times S$, $F \times S$ and $M \times F$ were also significant. The latter interaction indicates that certain faces were better recognised in one masking state than in another. Four faces were best recognised with no masking, 3 with wide masking and one in both narrow and wide masking. However, this finding is influenced by the particular area being masked as evidenced by the significance of the second order interaction $A \times M \times F$.

5.3.3 Discussion

The fact that the eyes area when masked, resulted in lower scores than the nose or mouth areas indicates that visual noise provides a sensitive measure of attribute importance. Recognition for the three masked areas varied over a range of 12% which would not have been the case if subjects had ignored the discrete features in favour of the more visible and distinct attributes. At the same time, it becomes clear that the eyes area does contribute to recognition. On the other hand, masking of the nose and mouth areas yielded results very similar to those obtained in the 'no masking' situation. On that basis, it could be said that these areas have little effect on recognition when all other attributes are present. That is not to say that these areas are ignored but that their contribution is small compared to that of certain other attributes.

It was mentioned earlier that a drawback in masking facial areas is that other attribute information is destroyed unintentionally. For example, it may be that the eye to eye separation is actually more important than the eyes themselves. Unfortunately, this drawback cannot be rectified unless a more convenient set of attributes is chosen. This set would have to be composed of attributes defined exclusively in terms of facial areas. The author believes that this approach would not be in accord with present evidence in pattern recognition research (chapter II). For that reason,

the set of attributes defined in section 3.2 will be kept despite the drawback just mentioned.

In the experiment which has just been reported, discrete feature attributes were evaluated in terms of masked areas. Because it is not possible to compare discrete features and feature relation attributes within the same test, an experiment was designed to measure feature relations with all other attributes being kept identical. This experiment will be reported in chapter VI.

Because three noise levels were used for the three sessions of this experiment, it was decided to perform another analysis of variance on the data in order to test the effect of these levels. It was hoped that by lowering the noise attenuation level by two decibels at each session there would be no difference in the results for each session. This was found to be the case although an improvement was found within sessions. This finding lends indirect support to the results found in experiment 1 over the four sessions.

5.4 Summary of the first three experiments

The purpose of the first experiment was twofold: to investigate the use of visual noise as a tool in perception studies and to raise some questions about facial image recognition.

The experiment measured noise attenuation levels at which recognition occurred for facial images embedded in noise. Eight stimulus faces were tested over four sessions. Results and conclusions are summarised below:

- (a) There was no difference in performance between subjects tested at one day and one week intervals. The implications of this result were not discussed since not enough intervals were tested. However, it is possible that the constraints of the experiment might have ruled out any possible advantage in the case of one day intervals.
- (b) Subjects tested with three second exposures made more correct responses but at lower noise levels than subjects tested with seven second exposures. The difference in noise levels was not significant. It was suggested that the seven second group had more confidence in their 'initial impressions' but that these could often be incorrect.

(c) Performance improved steadily over the four sessions. It was predicted that performance would level off asymptotically at the noise saturation threshold.

(d) Some faces were easier to recognise than others. Partial support was given for correlation between face distinctiveness (lack of confusions) and ease of recognition (recognition at higher noise levels).

(e) Some pairs of faces were confused more than others. It was suggested from a study of the resulting confusion matrix that contour attributes were more important than discrete features or feature relation attributes.

(f) No significant difference was found between the three face views that were tested. This was explained by the fact that the reference board also contained these views.

Experiment 2 investigated the use of masking. Two types of masking were used on two different portions of the facial images. Subjects were given three second exposures for each stimulus and visual noise was present at a low level. The findings are now summarised:

(a) No difference was found between the results for male and female samples in this experiment.

(b) Electronic and natural masking did not produce different results.

(c) Some faces were easier to recognise than others.

(d) Bottom masking (everything below the tip of the nose) produced higher scores than top masking (everything above the eyebrow line).

(e) It was concluded that the contour attributes in the upper half of the face were the most important attributes for recognition.

Experiment 3 was designed to test the effect of masking three other areas of the facial images. These were designated the eyes, nose and mouth areas. Three masking situations were used: 'no masking' (control), wide masking and narrow masking. The following results were obtained:

(a) The eyes area when masked produced significantly lower scores than the nose and mouth areas.

(b) There was no significant difference in scores for the three masking situations although the 'no masking' case did produce higher scores than the wide and narrow masking cases.

(c) As expected, 'faces' were a significant variable.

The results of the three experiments reported in this chapter suggest that facial attributes vary in importance. In particular, contours are seen as the most important class of attributes (especially those in the upper half of the face) i.e. the forehead-hairline contour, and the head outline contour. Of the discrete feature attributes, it is suggested that the eyes are more important than the nose or mouth. However, this suggestion must await further evidence on the grounds that the effects of feature relation attributes have not been separated from those of the discrete features. The first experiment in the next chapter is designed to test the effects of changes in size of the feature relation attributes.

Overall, it is considered that the use of visual noise and masking techniques have aided in the design of the experiments. Some of the drawbacks have been discussed for both techniques.

CHAPTER VI

TWO EXPERIMENTS USING LINE DRAWINGS AND PHOTOFIT

Two experiments are reported in this chapter. Section 1 deals with an experiment designed to evaluate the effect of changes in the feature relation attributes. Subjects were given the task of rating 28 line drawings which differed only in the relative positions of the discrete features. In other words, only the feature relation attributes had been altered. The rating was one of similarity with respect to a reference drawing.

The second experiment involved the use of the Photofit kit. Subjects were tested in a recognition task under conditions of very short stimulus exposure. This test is described in section 2.

A summary of both experiments is given in section 3.

6.1 An experiment using line drawings

6.1.1 Introduction

So far, only two types of attributes have been examined - contour attributes and discrete feature attributes. In section 5.2, it was suggested that contours, especially those in the upper half of the face, were the most important group of attributes. In section 5.3, it was proposed that the eyes were the most important of the three discrete features. Both these conclusions were advanced on a tentative basis since the influence of feature relation attributes present in both cases had not been investigated. In this section, feature relations will be examined with all other attributes remaining constant. Line drawings will be used for this purpose.

The following procedure was used in the preparation of the line drawings: first, a negative transparency of a face was placed in an enlarger and its horizontal projection traced with a black pencil on a white sheet of paper. This drawing became the standard. Figure 6.1 shows the standard in its actual dimensions. From this standard, a set of 28 variations was prepared. In each case, only the positions of the eyes, nose and mouth were altered. In other words, only the feature relation attributes were altered with respect to their size in the standard. The entire set of variations appears in table 6.1. Discrete feature positions were altered by one of two values: $1/5$ th or $1/10$ th of an inch. Half the subjects were tested with $1/5$ th of an inch displacements and the remainder with $1/10$ th of an inch displacements. Furthermore, two standard faces were used in the experiment. Figure 6.2 illustrates the second standard.

Subjects were given the task of evaluating each of the 28 line drawings with respect to a standard on a similarity criterion. The set of drawings was split up into four smaller groups because it was felt that subjects could not compare and evaluate all 28 drawings simultaneously. The drawback to this method is that drawings in different groups cannot be compared unless all combinations of drawings within the four groups are tested. This would involve an unrealistic number of subjects. Arbitrary classification of the line drawings resulted in the following four groups:

Figure 6.I The first line drawing reference.

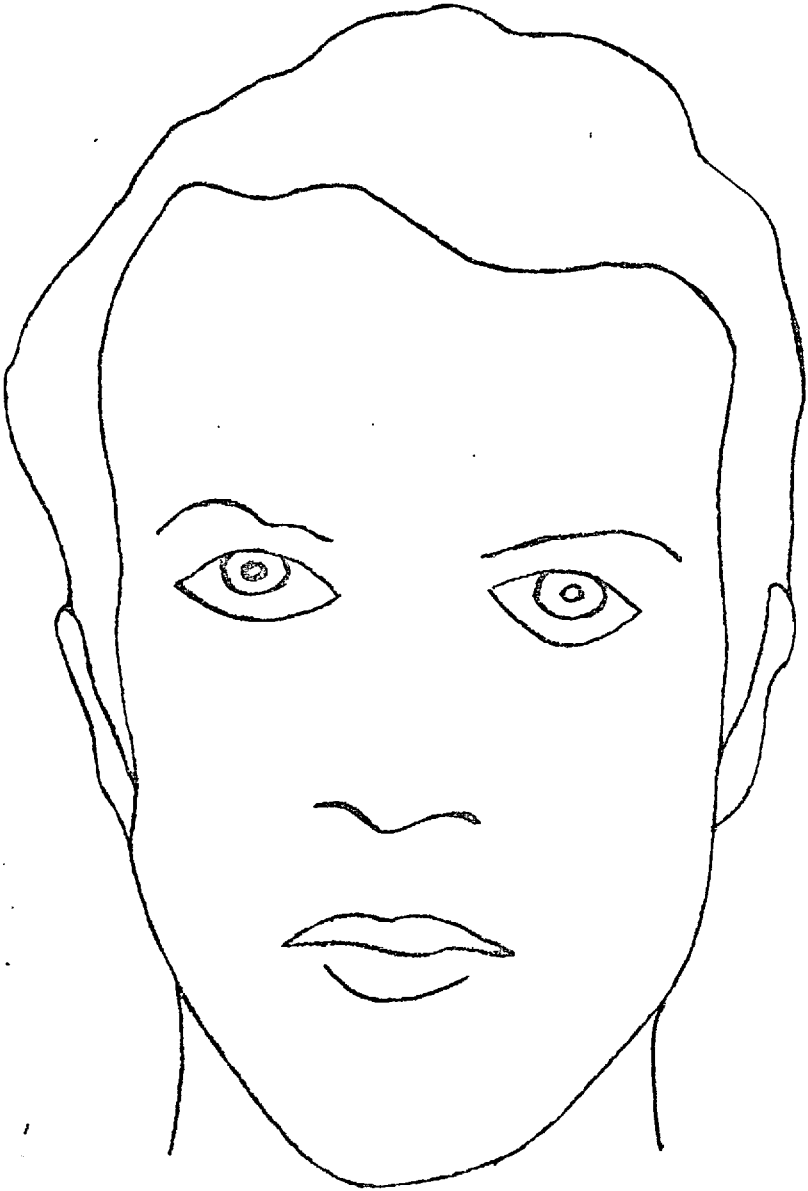


Figure 6.2 The second line drawing reference.

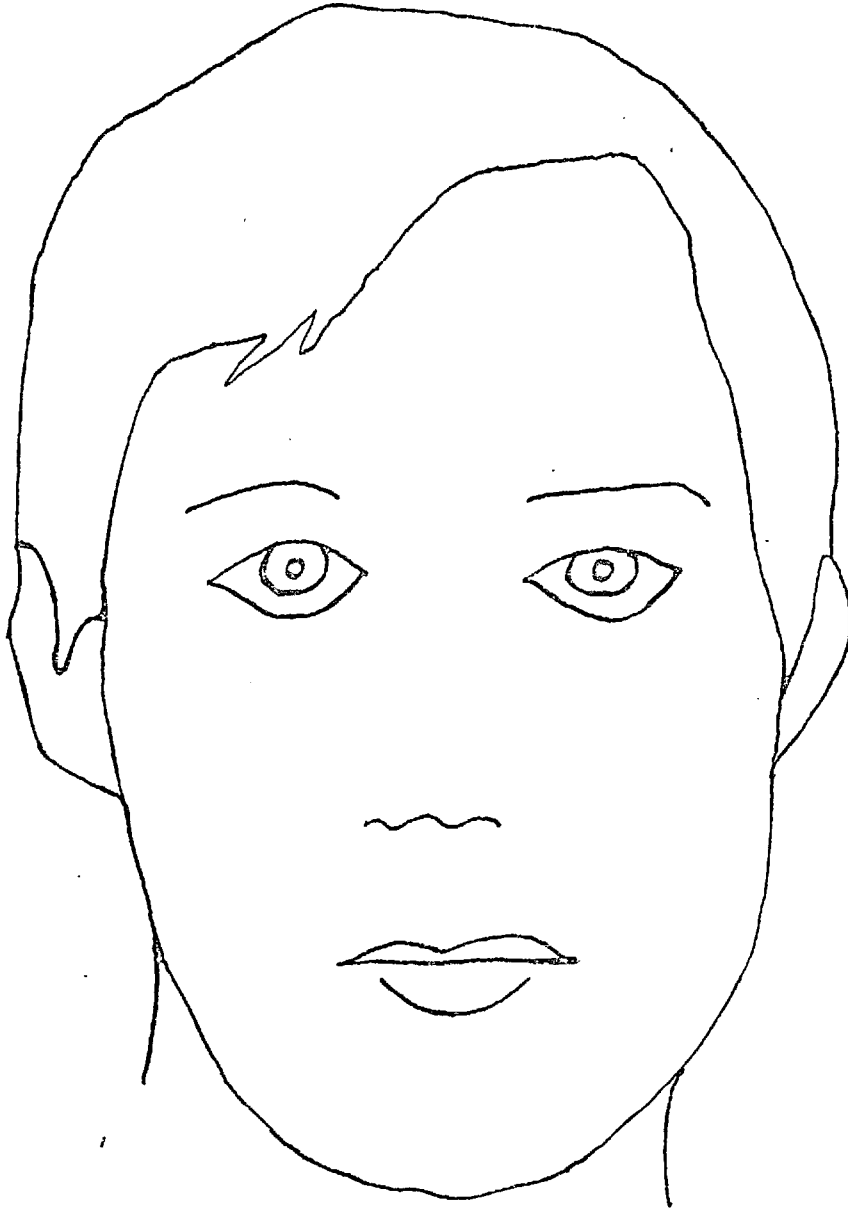


Table 6.1 The Ranking Orders

	No.	Rank Score	S.E. of the Mean	Variation
<u>Group 1</u>	1	2.99	0.226	Standard line drawing
	2	3.68	0.247	Eyes moved down
	3	4.49	0.250	Nose moved up
	4	4.95	0.256	Eyes moved up
	5	5.17	0.240	Mouth moved up
	6	5.25	0.212	Mouth moved down
	7	5.34	0.230	Nose moved down
	8	6.44	0.230	Eyes further apart
	9	6.67	0.257	Eyes closer together
<u>Group 2</u>	1	1.50	0.108	Standard line drawing
	2	3.50	0.175	Eyes moved down, Nose moved up
	3	4.01	0.181	Eyes moved up, Mouth moved down
	4	4.32	0.182	Mouth moved down, Nose moved up
	5	4.38	0.177	Eyes moved down, Mouth moved up
	6	4.84	0.157	Eyes moved up, Nose moved down
	7	5.54	0.179	Mouth moved up, Nose moved down
<u>Group 3</u>	1	1.78	0.112	Standard line drawing
	2	3.37	0.169	Nose and Mouth moved up
	3	3.59	0.186	Nose and Mouth moved down
	4	4.06	0.180	Nose and Eyes moved up
	5	4.69	0.178	Mouth and Eyes moved down
	6	4.81	0.175	Nose and Eyes moved down
	7	5.64	0.162	Mouth and Eyes moved up
<u>Group 4</u>	1	2.00	0.131	Standard line drawing
	2	2.67	0.187	Eyes moved down, Nose moved down, Mouth down
	3	3.45	0.175	Eyes moved up, Nose moved up, Mouth moved up
	4	5.38	0.220	Eyes moved down, Nose moved up, Mouth up
	5	5.50	0.215	Eyes moved down, Nose moved up, Mouth down
	6	5.65	0.208	Eyes moved up, Nose moved down, Mouth down
	7	5.88	0.207	Eyes moved up, Nose moved up, Mouth down
	8	7.18	0.196	Eyes moved up, Nose moved down, Mouth up
	9	7.29	0.176	Eyes moved down, Nose moved down, Mouth up

- (1) line drawings in which only one discrete feature position was altered with respect to the standard,
- (2) line drawings in which two discrete feature positions were changed in opposite directions, i.e. one upwards and one downwards,
- (3) line drawings in which two discrete feature positions were changed in the same direction, i.e. both upwards or both downwards,
- (4) line drawings in which all three discrete feature positions were altered.

Method Subjects proceeded in the following manner with each of the four groups of line drawings; a group was laid out in front of the subject and a standard face was placed on a board at eye level. Each drawing in the group was then rated from the most similar to the least similar with respect to the standard reference. A standard was included in each of the four groups. No time limit was imposed. All groups were rated by each subject and the total time used by each one varied between twenty minutes and one hour. Two types of reference standard were used in the testing. Half the subjects used a line drawing standard and the other half a photograph of the same dimensions prepared from the same original negative.

A total of 96 subjects was tested in this experiment. Each subject was given the following written instructions: "You have before you a group of line drawings which have been altered with respect to a reference also placed before you. The variations consist of displacements of the eyes, nose and mouth. Nothing else has been changed. I would like you to rank these line drawings, placing at one end of the table the face which you consider to be most similar to the reference, followed by the second most similar, the third and so on. When you are satisfied with your arrangement of this group, we will pass on to the second group. There are four groups in all and no time limit is imposed."

The four groups were always presented in the same order, i.e. group 1 followed by groups 2, 3 and 4. The line drawings within each group were randomized for each subject before presentation.

6.1.2 Results

Results were analysed separately for each of the four groups of drawings. Using group 1 as an example, the following procedure was used in scoring the ranked drawings: one point was given to the drawing ranked as most similar to the reference, two points for the second most similar, three points for the third and so on. Average scores were computed for each drawing by means of a BMD01D data description program and these appear in table 6.1 beside their respective drawings. Standard errors of the mean are also presented.

Average scores were used as the data for the analysis of variance performed on each group. The analysed variables are presented in table 6.2. Because of the balanced method in which the drawings were scored, only certain effects could be tested for significance. These consisted of the 'drawings' variable and all interaction effects. Results indicate that in fact only the 'drawings' variable turned out to be significant. These are presented for the four groups in table 6.3. The fact that none of the first order interactions were significant implies that subjects were consistent with their judgments in all experimental conditions. In other words, the use of two different references, two displacement values and two standard faces did not result in different similarity judgments (two-sided test at the 0.05 level).

The fact that both types of reference produced similar results indicates that the line drawing standard acted as an accurate representation of the photographed version. Furthermore, it is probable that line drawings of the type used in this experiment provide more realistic data than the more conventional schematic faces originally designed by Brunswik and Reiter (1938).

The non-significance of the displacement variable is interesting but also inconclusive since only two values were examined. Subjects reported that drawings with 1/5th of an inch displacements immediately appeared to be altered whilst subjects judging drawings with the smaller displacement reported that some changes were barely detectable. It should be noted, however, that subjects were not given the task of detecting displacements but rather that of ranking the drawings on a similarity criterion. Many subjects did in fact try to detect which positions had been altered before deciding on the ranking order, but there is no

Table 6.2 Variables for the line drawing experiment

Variable	Abbreviation	Levels of each variable
References	R	Line drawing and Photograph references
Faces	F	Standard Faces 1 and 2
Displacements	D	1/5 and 1/10 of an inch displacements
Drawings	V	9 drawings in groups 1 and 4, 7 drawings in groups 2 and 3

Table 6.3 Significant effects for the four groups of drawings (tested separately)¹

Group	Effect	Mean Squares	Degrees of freedom		F ratio	Level of Significance
1	V	11.4	8	14	3.04	0.07
2	V	12.8	6	12	8.67	0.01
3	V	12.9	6	12	25.64	0.01
4	V	28.7	8	14	15.73	0.01

¹The method of analysis used for the four groups was the same as that used in previous experiments. The test of significance was two-sided.

indication that this produced different responses from those of subjects who avoided this initial processing.

Since the main purpose of this experiment is to study the effects of feature relation attributes, it becomes necessary to define each line drawing in terms of the feature relation attributes which have been altered. This is illustrated in table 6.4. The following notation has been used: 'S' indicates that the attribute size has been made smaller as a result of changes in the position of discrete features, 'L' means that the size has been made larger and a dash indicates that the size has not been altered. All these changes are with respect to the attribute sizes established in the standard reference. The following abbreviations are used for each of the attributes:

- (1) mouth to chin tip separation (M-C),
- (2) mouth to nose separation (M-N),
- (3) mouth to eyes separation (M-E),
- (4) nose to eyes separation (N-E),
- (5) forehead width (FW),
- (6) eye to eye separation (E-E).

Because it is not possible to measure the effect of each feature relation attribute directly, a modified approach will be employed. This will consist in measuring the effect of 'changes' in each of the attributes. Subsequently, the effects of these changes will be related to the effects of the actual attributes. In all, there are twelve changes, two for each of the six attributes. The notation for each change appears in table 6.5. For example, M-N(s) is the abbreviation for the change which makes the mouth to nose separation smaller.

Table 6.4 indicates that each attribute change was present in nine drawings except for the case of E-E(s) and E-E(l) which appeared in one drawing only. These last two changes will not be considered for the moment in the computation of the effects of each change. An example will now be given to illustrate how the effects of the remaining ten changes were measured. Consider the nine drawings in which M-N(s) occurs. It can be seen that other attribute changes occur in each of these drawings. By summing all changes occurring in the nine drawings and equating the result to the sum of the scores assigned to each of the drawings, the following equation is obtained:

Table 6.4 The variations of the feature relation
attributes for each of the line drawings

(S): smaller separation (L): larger separation (-): no change

No.	Score	FW	M-C	M-N	M-E	N-E	E-E
1	2.99	-	-	-	-	-	-
2	3.68	L	-	-	S	S	-
3	4.49	-	-	L	-	S	-
4	4.95	S	-	-	L	L	-
5	5.17	-	L	S	S	-	-
6	5.25	-	S	L	L	-	-
7	5.34	-	-	S	-	L	-
8	6.44	-	-	-	-	-	L
9	6.67	-	-	-	-	-	S
1	1.50	-	-	-	-	-	-
2	3.50	L	-	L	S	S	-
3	4.01	S	S	L	L	L	-
4	4.32	-	S	L	L	S	-
5	4.38	L	L	S	S	S	-
6	4.84	S	-	S	L	L	-
7	5.54	-	L	S	S	L	-
1	1.78	-	-	-	-	-	-
2	3.37	-	L	-	S	S	-
3	3.59	-	S	-	L	L	-
4	4.06	S	-	L	L	-	-
5	4.69	L	S	L	-	S	-
6	4.81	L	-	S	S	-	-
7	5.64	S	L	S	-	L	-
1	2.00	-	-	-	-	-	-
2	2.67	L	S	-	-	-	-
3	3.45	S	L	-	-	-	-
4	5.38	L	L	-	S	S	-
5	5.50	L	S	L	-	S	-
6	5.65	S	S	-	L	L	-
7	5.88	S	S	L	L	-	-
8	7.18	S	L	S	-	L	-
9	7.29	L	L	S	S	-	-

$$\begin{aligned}
 &9M-N(s) + 4M-E(s) + M-E(l) + N-E(s) + 5N-E(l) + FW(s) + 3FW(l) \\
 &+ 6M-C(l) = 50.19
 \end{aligned}$$

This equation is based upon the nine drawings in which M-N(s) occurs. Equations can now be written down for the remaining nine changes which also appear in nine drawings. This will result in a system of ten equations with ten unknowns, each unknown being a change. Solution of these equations by matrix inversion yields values for each of the ten changes. These values represent the effects of each of the changes. The larger the value of a change, the more that change has contributed in altering a face with respect to the reference. In other words, larger effects make faces more dissimilar than smaller effects. Column 1 of table 6.5 represents the effects of the changes measured by the methods described above.

The method of computing the effects of attribute changes in this experiment is open to the following criticism: it is possible that certain line drawings with low scores in one group would have had higher scores if they had been placed in another group. This criticism follows from the drawback discussed in the introduction. For this reason, the results should be viewed cautiously. An alternate method of evaluating the same 28 variations will be reported in the next chapter. This will involve a sequential rather than simultaneous method of presenting the stimuli. Together, the results of both methods will produce a more reliable estimate of the effects of changes in feature relation attributes. For the moment, only the results of this experiment will be described.

In designing the present test, it was predicted that the effects of attribute changes would be correlated with their Weber fractions. On the assumption that all attribute sizes are changed by the same amount, and in the same direction, i.e. all sizes made smaller or bigger, then the value of a Weber fraction will be inversely related to the size of the attribute. In other words, larger attributes will have smaller Weber fractions than smaller attributes when both are changed by the same amount and in the same way. The prediction that the effects of changes in attribute size will be correlated with their Weber fractions therefore implies that smaller attributes, when changed, will have more effect than changes in larger attributes.

Table 6.5 A comparison between the measured effects of attribute changes and their Weber fractions.

Attribute Changes	Column 1 Effects	Column 2 Weber fractions
E-E(s)	-	0.375
E-E(1)	-	0.273
M-N(s)	2.21	0.250
M-N(1)	1.78	0.200
M-C(s)	1.04	0.167
M-C(1)	1.92	0.143
N-E(s)	0.67	0.125
N-E(1)	1.29	0.111
M-E(s)	0.52	0.083
M-E(1)	1.34	0.077
FW(s)	1.03	0.075
FW(1)	1.56	0.070

Table 6.6 The effects of the attributes

Attributes	Effects	Weber fractions
E-E	-	0.648
M-N	3.99	0.450
M-C	2.96	0.310
FW	2.59	0.145
NE	1.96	0.236
ME	1.86	0.160

Weber fractions were measured in the following manner: when an attribute was made smaller, its Weber fraction was measured as the ratio $\Delta S/S$ where ΔS is the incremental change and S is the original size of the attribute. When an attribute was made larger, its Weber fraction was measured as the ratio $\Delta S/(\Delta S + S)$. The choice of both Weber fractions was arbitrary and could equally well have been $\Delta S/(S - \Delta S)$ and $\Delta S/S$ as long as both values were different. This would ensure that different values occurred between changes to a smaller size and changes to a larger size. Both pairs of fractions also resulted in the same order or ranking of the attribute changes and therefore, the former pair was chosen as being equally acceptable as the latter. Weber fractions for each of the ten changes appear in column 2 of table 6.5. These values were computed on the basis of the separations present in standard face 2. These are 1.2 in. for N-E, 0.6 in. for M-N, 1.8 in. for M-E, 2.0 in. for FW and 0.9 in. for M-C. The incremental change ΔS was chosen as 0.15 in. which is the average between the two displacements used in this experiment.

A Kendall rank correlation test was used in checking the experimental prediction. The correlation coefficient was found to be significant at the 0.15 level. The same test was also applied a second time with the added changes E-E(s) and E-E(1). Although both these changes were only measured on the basis of one drawing each, it was felt that they had produced the largest effects when originally judged by subjects. For that reason, they were put at the top of the list in column 1 of table 6.5. Their Weber fractions, using $S = 0.8$ in. and $\Delta S = 0.3$ in., were larger than those of the other ten changes, i.e. $E-E(s) = 0.375$ and $E-E(1) = 0.273$. The resulting correlation coefficient for the second test was found to be significant at the 0.01 level.

6.1.3 Discussion

The prediction that changes in smaller attributes had a greater effect than changes in larger attributes was partially verified by the results of the Kendall rank correlation tests. However, because of the limitations discussed in the previous sub-section, no definite conclusions can be drawn until further evidence is provided. This will be reported in chapter VII.

It is possible, nevertheless to make certain general statements about the results. For instance, it appears that certain changes in attribute sizes have more effect than others. In particular, changes in small attributes seem to alter faces more than changes in larger attributes. Whether these changes have more effect than changes in other types of attributes remains to be examined. This will be the purpose of the second experiment in chapter VII. Another interesting finding concerns the consistency with which subjects judge faces on a similarity criterion. None of the experimental variables used in the experiment seemed to alter this consistency. Subjects were also accurate in identifying the standard face present in each of the four groups of drawings, even when their reference was in photographic form. This suggests a considerable acuity in judging the relationship between the discrete features.

So far, only the effects of changes in attribute size have been measured. It is not possible to infer from these results which attributes are the most important, in terms of the definition of importance given in section 3.2. Attribute importance can only be measured in a recognition task and not in a test involving subjective evaluations of similarity. Nevertheless, it is possible to classify the feature relation attributes in terms of the effects which they produce when altered. It is suggested that the effect of an attribute be taken as the sum of the two effects produced by that attribute, in one case when it is made smaller and in the other case when it is made larger. For example, the effect of the attribute M-N can be taken as the sum of the effects M-N(s) and M-N(l). On that basis, the following order emerges, beginning with the attribute with the greatest effect: the eye to eye separation, the mouth to nose separation, the mouth to chin tip separation, the forehead width and finally the remaining two attributes which have about the same effect, i.e., the nose to eyes separation and the mouth to eyes separation. Values for these effects are presented in table 6.6. Roughly speaking, they agree with those obtained by summing Weber fractions except in the case of the forehead width.

6.1.4 Summary

The following results were obtained from this experiment:

(a) Subjects were found to be consistent in the way in which they judged line drawings of faces on a similarity criterion.

(b) Certain attributes, when altered in size, were found to have greater effects than others. The magnitude of the effect of an attribute reflected the degree to which a face had been altered by the presence of that attribute when changed. It was suggested that changes in smaller attributes had more effect than changes in larger attributes.

(c) The following order of attribute effect was found:

- (1) The greatest effect -- the eye to eye separation.
- (2) The mouth to nose separation.
- (3) The mouth to chin tip separation.
- (4) The forehead width.
- (5) The mouth to eyes and nose to eyes separations.

It was stressed that these results should be taken on a tentative basis because of the limitations discussed. It was mentioned that further evidence would be presented in experiment 2 of chapter VII.

6.2 The first experiment using Photofit

6.2.1 Introduction

The experiments described in chapter V made use of two techniques designed to limit the amount of information available for recognition, i.e. visual noise and masking. In the present experiment, a third technique will be used, namely tachistoscopic presentation. This will be done by means of the frame counter unit described in section 4.4. Furthermore, Photofit faces will be used as the stimuli.

In chapter III, a brief introduction was given to the Photofit kit. It was mentioned that the kit could be very useful in face recognition experiments because of the ease with which faces could be assembled and individual components changed. These advantages were put to use in the present experiment. It will be recalled that a Photofit face consisted of five components: the forehead-hairline section, the eyes, nose and mouth strips and the chin section. Because of the large quantity of variations for each component, it is possible to assemble an extremely large number of different faces.

The purpose in using Photofit faces will now be explained by means of an example. Suppose that a subject is provided with the reference board containing three different Photofit faces. By different is meant that none of the faces have any components in common. The subject is now told that one of these three faces will be presented on the monitor screen for a very short duration, i.e. less than 100 milliseconds. His task is to recognise which of the three faces has been presented. Now suppose that the stimulus face is actually made up of components from the three reference faces. Then, provided that the subject is not aware of this, it may be assumed that his response will reflect the importance of the different components. For example, if the stimulus is composed of the forehead-hairline section of face 1, the eyes, nose and mouth strips of face 2 and the chin section of face 3, then if the subject responds by pressing button 3, it can be inferred that the chin section is the most important component for recognition, at least for that particular stimulus.

It has already been stressed that pattern recognition does not consist in coding specific facial areas but rather a set of attributes,

some of which can not be neatly enclosed within any facial area. For that reason, the five components which are used to make up a Photofit face cannot be looked upon as five independent perceptual units. When components from different reference faces are used in assembling a new face, some of the attributes in this new face may not bear any relation whatsoever to the attributes in the reference faces. This is particularly evident in the case of the feature relation attributes. For this reason, the results of the present test must be viewed carefully.

Eight stimulus faces were prepared from the Photofit kit for a recognition task. All components were different and their selection was random with the restriction that the resulting sample should be fairly homogeneous. This last criterion was developed at length in section 3.3. A set of 80 slides was prepared from the eight stimulus faces. The slides were distributed in the following manner:

- (1) Faces 1, 6, 7 and 8 were each presented four times, i.e. a total of 16 slides.
- (2) Faces 2, 3, 4 and 5 were used in preparing the remaining 64 slides. It was decided that a Photofit face should be considered as an assembly of three parts: the forehead-hairline section, the eyes-nose-mouth section considered as one unit and the chin section. These are respectively abbreviated as F, M and C. On the basis of this classification, it is possible to construct 64 faces which make use of the three parts of each of the four reference faces, i.e. $4^3 = 64$. Within this set, some faces will be made up of parts from three different reference faces (24), some faces will be made up of parts from two different reference faces (36) and finally, four of the faces will be the actual reference faces.

Method These 80 slides were presented to subjects in the same manner as in previous tests, except that the duration of presentation for each slide was 40 milliseconds for half the subjects and 120 milliseconds for the other half. Furthermore, an eight second interval was used between each presentation. A warning buzz preceded every presentation by about 500 milliseconds. The entire set of slides was pre-recorded using an Ampex 7003 video taperecorder. Half the subjects received one order of presentation and the other half received the reverse order.

A total of 60 subjects was tested. Each was given a reference board containing photographs of faces 1 to 5. Response buttons were provided corresponding to the five reference faces. A sixth button used for indicating the NONE response was also available. This button was added for the case when faces 6, 7 and 8 would be presented. Presumably, it could also be pressed when a 'combination' face was presented, i.e. a face made up of different parts from amongst the components of reference faces 2, 3, 4 and 5. Subjects were not told that any 'combination' faces would be presented. It was established in a pilot test, that the likelihood of subjects noticing the presence of 'combination' faces was extremely small for presentations lasting less than four frames, i.e. 160 milliseconds.

Because of the very short exposure durations used in this experiment, it was felt that an objection might be raised against the method of presenting the slides. This is done by means of conventional picture scanning. At the monitor end, the image is presented by means of a flying spot which scans from the top to the bottom of the screen. This constitutes one television field lasting 20 milliseconds. A second field is interlaced between the lines of the first scanned field and together, the two fields constitute one frame lasting for 40 milliseconds. It was mentioned that half the subjects received 40 millisecond presentations (1 frame) and the other half, 120 millisecond presentations (3 frames). The objection to the scanning comes from the fact that the information in the upper half of the facial image is presented before the information in the lower half of the image although all information is presented for the same duration. This might possibly lead to a bias in favour of information in the upper half of the monitor screen. To avoid this objection, the following changes were made in the presentation. First, slides were mounted sideways in the projector and second, the monitor was rotated 90 degrees in a special cradle. The resulting image on the monitor screen was an upright facial image but with scanning from left to right instead of from top to bottom.

Subjects were given the following written instructions: "You will be presented with a sequence of 80 slides, each at eight second intervals. A warning buzz will precede each presentation. Beside you is a reference board containing five numbered photographs. When you

think that one of these has been presented on the monitor screen, you press the appropriate button. If you consider that none of these faces has been presented, you press the button marked NONE. You will be given two minutes to study the reference board and it will remain with you throughout the session. Before the actual test begins, you will be presented with five demonstration slides. These do not count in the test but are simply there to illustrate the presentation format. Slides will be presented for only 1/25th (3/25th) of a second and you will not be told whether you have made any mistakes."

The five demonstration slides consisted of the five reference faces presented in their numbered sequence. The eight second interval between presentations was chosen in a pilot study. With this interval, subjects were not too rushed on the one hand and on the other hand, they did not have too much time in which to become bored.

6.2.2 Results

It is convenient to separate the set of 80 stimuli into the following three categories: slides which were 'entities', slides which used components from two reference faces and slides which used components from three reference faces. The last two categories contain the 'combination' faces and it is these which offer the most important data. Each category will now be examined in greater detail.

1. 20 slides which were 'entities', i.e. reference faces 1 to 5 and faces 6, 7 and 8.

The twenty slides consist of faces 1, 6, 7 and 8 each presented four times and faces 2, 3, 4 and 5 each presented once. Table 6.7 presents a summary of the results for the two exposure durations. Two general observations can be made from the results; first, recognition was above the chance level and second, recognition improved with exposure duration. On the whole, it can be seen that recognition of faces is quite accurate even when the exposure duration is very short.

2. 24 slides which were made up of components from three different reference faces.

Table 6.8 summarises the results for this group of stimuli. For both exposure durations, recognition appeared to be based more often on

Table 6.7 Results for the 20 slides which were 'entities'.

SLIDES	Number of Presentations	Recognised	Exposure Duration	
			1/25 sec.	3/25 sec.
Faces 6,7 and 8	4 times each	Correctly as NONE As 1,2,3,4 or 5	74.2% 25.8%	81.4% 18.6%
Face 1	4 times	Correctly As NONE As 2,3,4 or 5	63.3% 25.0% 11.7%	75.8% 14.2% 10.0%
Faces 2,3,4 and 5	Once each	Correctly As NONE Wrongly as 2,3,4 or 5	73.3% 7.5% 19.2%	76.7% 13.3% 10.0%

Table 6.8 Results for the 24 slides made up of components from three reference faces

Exposure Duration	% of times the slide was recognised on the basis of			% of NONE responses
	Forehead-hairline	Eyes-nose-mouth	Chin	
1/25 sec.	29.7%	18.7%	10.9%	27.2%
3/25 sec.	21.1%	15.0%	12.8%	40.0%

Table 6.9 Results for the 36 slides made up of components from two reference faces

SLIDES	RECOGNITION SCORES	1/25 sec.	3/25 sec.
12 slides with F and M from the same reference face	Recognised on the basis of F-M	49.2%	48.1%
	Recognised on the basis of C	13.3%	11.1%
	Recognised as NONE	16.1%	25.5%
12 slides with F and C from the same reference face	Recognised on the basis of F-C	42.5%	40.6%
	Recognised on the basis of M	21.0%	13.5%
	Recognised as NONE	23.0%	32.5%
12 slides with M and C from the same reference face	Recognised on the basis of M-C	36.4%	38.1%
	Recognised on the basis of F	22.8%	22.5%
	Recognised as NONE	23.6%	28.1%

the forehead-hairline section (F) than on the other two parts (M and C). Of the latter, the eyes-nose-mouth section (M) was more often used than the chin section (C). The percentage of NONE responses increased with exposure duration. Although, only one subject actually noticed the use of 'combination' faces (personal comment to the author), it seems probable that subjects were beginning to distinguish more clearly between the presented stimuli and the reference faces as exposure time increased.

3. 36 slides which were made up of components from two different reference faces.

Results for this group appear in table 6.9. For both exposure durations, recognition appeared to be based more often on the joint F-M unit than on either of the other two joint units. Of the latter, the joint F-C unit was more often used than the M-C unit. When the remaining unit is considered for all three cases, it appears that the forehead-hairline section had more effect than the eyes-nose-mouth section which in turn had more effect than the chin section. The number of NONE responses increased in all cases with increased exposure duration.

It is interesting to note the correspondence between results for this category and the previous one. In category 2, the sequence of facial parts in terms of their use for recognition was F, M and C. In category three, the sequence of pairs of facial parts in terms of their use for recognition was F-M, F-C and M-C. Furthermore, the remaining part in this category produced the following sequence in terms of its use for recognition: F, M and C. On the basis of this finding, it is concluded that the most important part for recognition is the forehead-hairline section, followed by the eyes-nose-mouth section and finally the chin section.

6.2.3 Discussion

Before discussing the results of this experiment, it is useful to review the findings of previous experiments. In experiment 1, it was suggested that contour attributes were more important than discrete features and feature relation attributes. In experiment 2, it was concluded that contours in the upper half of the face were more important than contours in the lower half of the face. Experiment 3 examined the importance of discrete features and suggested that the eyes were more

important than the nose or mouth. Since the effect of feature relation attributes had not been measured in any of these experiments, it was decided to examine the question in a separate test. Indirect measurement of the effect of changes in size of the feature relation attributes revealed that changes in smaller attributes altered faces more than changes in larger attributes.

In the present study, three parts of the facial image were examined: the forehead-hairline section, the eyes-nose-mouth section and the chin section. Each of these parts contained a certain number of attributes. The forehead-hairline section contained the upper portion of the head-outline contour and the forehead-hairline contour. The eyes-nose-mouth section was composed of the three discrete features as well as three feature relation attributes: the eyes to mouth separation, the eyes to nose separation and the nose to mouth separation. The chin section contained the lower portion of the head-outline contour. In preparing the 'combination' faces, some attributes could not be transferred intact from the reference faces to the new 'combination' faces. For example, the forehead width in a new face bore no relation to any of the forehead widths in the reference faces unless the forehead-hairline and the eyes-nose-mouth sections in the new face came from the same original reference. The same argument applies in the case of the mouth to chin tip separation. Finally, the head-outline contour was sometimes partially altered in 'combination' faces using different references for the forehead-hairline and chin sections.

Having described the attributes present in each of the three parts under study, it appears that the following order of attribute importance emerges: first, the contours in the upper half of the face are once again found to be most important. Second, the eyes-nose-mouth section as one unit appears to be more important than the lower portion of the head-outline contour. In terms of single attributes, none of the features present in the eyes-nose-mouth unit may be more important than the lower portion of the head-outline contour, but the unit itself, by virtue of the number of attributes involved, automatically takes on greater importance. This follows from the fact that the unit must necessarily belong to a less homogeneous class than any of its respective components.

In the next chapter, the individual components of the eyes-nose-mouth unit will be evaluated separately.

6.2.4 Summary

The results of this experiment confirmed prior evidence suggesting that the contours in the upper half of the face were the most important attributes for recognition. The eyes-nose-mouth section was found to be more important than the chin section although it was suggested that this might not be the case for any of its component attributes.

NONE responses were found to increase with exposure duration indicating that the recognition process became more efficient with increased presentation time. Finally, it was mentioned that conclusions regarding the importance of attributes should be interpreted with reserve because of the possible effects of the feature relation attributes which were not measured.

6.3 Summary of the chapter

Two experiments were reported in this chapter. The first examined the effects of changes in feature relation attributes and the second investigated recognition of facial images in conditions of short exposure duration.

In the first experiment, it was suggested that changes in small feature relation attributes would have a greater effect in altering a face than equivalent changes in larger attributes. This prediction was partially confirmed by the results although it was stressed that more evidence was required. Nevertheless, there was clear evidence to indicate that feature relation attributes did in fact play a part in recognition judging from the effects produced by small alterations in their size.

An important result of the experiment was the consistency with which subjects made their similarity judgments under different conditions. The use of two references, two standard faces and two displacement values did not seem to affect the outcome of the results.

The second experiment was concerned with the importance of attributes in terms of Photofit components. Stimuli were exposed for very short durations and recognition was found to depend on certain components more than on others. Contours in the upper half of the face were once again found to be the most important attributes.

Both experiments carry certain implications for users of identification kits. For example, it becomes clear that the relationships between discrete features must be considered carefully in assembling a face. These can have a considerable influence on the appearance of a face and it may even be that changes in certain discrete features are less important than certain changes in the feature relation attributes. This possibility will be examined in chapter VII. Another implication concerns the importance of the different Photofit components. Results indicate that recognition is based more often on the forehead-hairline component than on any other. Whether the eyes, nose and mouth strips are individually more important than the chin section remains to be proved but their effect as a unit does seem to be.

CHAPTER VII

TWO MORE EXPERIMENTS USING PHOTOFIT

Two experiments using Photofit are reported in this chapter. Section 1 describes a test designed to measure how Photofit components are evaluated with respect to references on a similarity criterion. The five components used in preparing facial images are individually tested. In other words, several variations of a component are judged with respect to a standard reference of the same component and this is repeated for all five components. The standard components used as references are then used to prepare a standard face for the experiment described in section 2. This test compares similarity ratings for individual components evaluated in the first experiment with similarity ratings for the same components when they are used to alter the standard face. The test also compares the effect of changes in feature relation attributes with the effect of changes in discrete feature attributes. The possible additivity of the effects of changes in Photofit components is also investigated. Finally, the effects obtained in changing feature relation attributes in a Photofit face are compared with those obtained with a line drawing (section 6.1). Section 3 summarises the results of the two experiments.

7.1 A test of similarity for Photofit components

The two experiments to be reported in this chapter are related. For this reason, a brief explanation will be given to describe this relation before proceeding with the report of the first experiment.

As mentioned, Tversky and Krantz (1969) suggested that the effects of changes in certain schematic facial features were additive (in the ordinal sense). They used drawings of faces which were composed of three features: an elliptic head-outline contour, circular eyes and a slit mouth. Each feature could be in one of two forms. This resulted in eight stimuli which were rated by direct and comparative judgments on a similarity criterion.

In the two experiments described in this chapter, the additivity of feature effects will be examined in the case of facial images in photographic form. The first experiment will be concerned with the rating of Photofit components. This will involve similarity judgments on fourteen variations of a component with respect to a reference of the same component. Variations will be evaluated in the case of all five components used in preparing a Photofit face. The five components used as references in this test will then be assembled into a reference face in the second experiment. The components found to be the most similar to the five references in experiment 1 will be used to alter the reference face of the second experiment. This face will be altered by one to five components and subjects will make similarity judgments on the altered versions with respect to the reference face. The results will be examined to see whether in fact feature effects are additive. Other aspects will also be examined in both of the experiments, but these will be described in their respective sections.

7.1.1 Introduction

In this test, subjects were given the task of evaluating variations of Photofit components with respect to chosen references. Fourteen variations of each of the five Photofit components were rated in this manner. The chosen references for the five components were F51, E10, N21, M8 and C86. F, E, N, M, and C are the abbreviations for the forehead-hairline section, the eyes, nose and mouth strips and the chin section. The initial choice of the 14 variations of each component was made by the

author according to a rough similarity criterion. It should be noted however, that similarity between variations and the references was not a prerequisite for the purposes of this test.

Subjects were presented sequentially with the five samples of fourteen versions, beginning with the forehead-hairline group and followed by the eyes, nose, mouth and chin section groups. In each case, they were given the following verbal instructions: "You have before you a reference feature and fourteen other versions of the same feature. I would like you to extract from the group of 14 the five versions which you consider to be most similar to the reference. When you have done this, you are to rank the chosen five from the most similar to the least similar. There is no time limit."

When subjects had completed the task for the five groups, they were given the job of assigning numerical values to each of the five ranked versions of each group. This was done in the following manner; the five ranked versions were presented to the subject in the order in which he had ranked them initially. He was told that the least similar of the five would be given a fixed value of one point and that he must assign values to the remaining four versions with a possible maximum of five points. Subjects were restricted to integers in their ranking. They were also told that a score of five points given to a version implied that that version was identical to the reference. A total of 42 subjects was tested.

7.1.2 Results

The results of the test appear in table 7.1. Scores were obtained by summing all values assigned by subjects to each version. With 42 subjects, the possible maximum score for any version was 210, i.e. 5×42 .

The main purpose of this test was to extract from each group of 14 versions the two most similar versions. These appear in Table 7.2. Scores for each of the ten cases are presented in column 1. Column 2 indicates the number of times each version was present in the final selection of five. The third column indicates the average score given to each version out of a possible maximum of 5. Finally, in column 4 are the differences between the scores of column 3 and the maximum possible score

TABLE 7.1 Scores for the 14 versions of each component

No.	Forehead-Hairline Scores	No.	Eyes Strip Scores	No.	Nose Strip Scores	No.	Mouth Strip Scores	No.	Chin Scores
F48	109	E4	96	N23	118	M6	107	C88	85
F58	56	E13	86	N20	106	M10	93	C50	78
F79	44	E8	83	N8	80	M22	89	C84	76
F116	39	E1	57	N38	42	M2	80	C87	56
F41	38	E49	46	N28	41	M5	67	C65	48
F39	28	E31	28	N41	31	M29	27	C36	46
F76	25	E16	25	N9	31	M30	15	C18	17
F37	23	E24	23	N25	19	M32	13	C34	16
F112	19	E48	18	N7	17	M58	11	C45	16
F156	18	E43	16	N63	12	M33	10	C6	6
F38	15	E42	12	N60	9	M70	3	C72	6
F19	14	E66	9	N24	2	M36	1	C48	3
F11	6	E38	2	N5	1	M79	1	C19	0
F15	2	E62	0	N1	0	M34	0	C25	0

TABLE 7.2 Scores for the two most similar versions of each of the five components

Version	Column 1 Scores (Max: 210)	Column 2 No. of times in group of 5	Column 3 Mean Scores $\bar{x}_{\max.} = 5$	Column 4 Deviation $5 - \bar{x}$
N23	118	31	2.81	2.19
F48	109	32	2.60	2.40
M6	107	34	2.55	2.45
N20	106	32	2.52	2.48
E4	96	32	2.29	2.71
M10	93	33	2.21	2.79
E13	86	30	2.05	2.95
C88	85	30	2.02	2.98
C50	78	33	1.86	3.14
F58	56	25	1.33	3.67

of 5. This last column will be used in a comparison with scores obtained in the next experiment. In the latter case, the ten versions will be used individually to alter a reference face composed of the five reference components used in the first experiment. The effect of each of the ten versions in altering the reference face will be compared with the results of column 4. The absolute values in this column are not important in themselves since the comparison will be based on the ranking order in each case.

7.1.3 Discussion

The results of this experiment consist of 'similarity scores' for the two most similar versions of each of the five Photofit components. The values in column 4 of table 7.2 represent 'deviations' from the reference in the sense that their magnitude increases with dissimilarity from the reference. As an illustration of the comparisons which will be made, in the next experiment, consider the following example: F48 deviates from the reference F51 by 2.40 (column 4). Now suppose that a face is made up of all the references used in the first test. Furthermore, a second face is made up of the same references except that F51 is replaced by F48. If we consider the first face to be the reference, then the similarity score obtained from rating the second face with respect to that reference will provide us with a deviation score which will be a function of the effect of F48. It is this score which will be compared with the score of 2.40 in experiment 1. Scores and ranks for all ten versions will be involved in the comparison.

7.2 A test on the additivity of component effects

The main purpose of this experiment was to investigate the possible additivity of Photofit component effects measured from similarity judgments. The design of the test also allowed results to be used for the following purposes:

- (1) to compare the effects of changes in feature relation attributes with those of changes in discrete feature attributes,
- (2) to compare the effects obtained from changes in feature relation attributes using Photofit with those obtained previously with line

drawings (section 6.1),

(3) to make the comparison described in the previous section.

7.2.1 Introduction

The present test is in many ways similar to that of section 6.1 except for the following changes; faces were varied not only by means of changes in feature relation attributes but also by changes in other attributes (or more specifically, by changes in Photofit components). Judgments of similarity were made from sequential presentation rather than from simultaneous presentation of groups of faces. Finally, facial images were presented in Photofit form rather than in the form of line drawings. The reference face used throughout the test was made up of the component references of the previous test. This face is illustrated in figure 7.1.

64 altered faces were prepared for the experiment. Table 7.3 presents a list of these faces. Seven groups are outlined, each with the following characteristics:

- Group 1: This group consists of the reference face, repeated twice.
- Group 2: This group is composed of ten faces which deviate from the reference by only one component. The ten components used in the alterations correspond to the ten most similar components evaluated in the previous test (2 versions of each of the 5 Photofit components).
- Group 3: This group consists of ten faces which deviate from the reference by two components. The components used in the alterations correspond to the most similar versions in the previous test, i.e. the 5 versions most similar to the 5 references (one version per reference).
- Group 4: This group is composed of ten faces which deviate from the reference by three components. The components used in the alterations are the same as those used in group 3.
- Group 5: This group is made up of 5 faces which deviate from the reference by 4 components. The components used in the alterations are the same as those used in groups 3 and 4.
- Group 6: This group consists of two faces. One face is composed of the 5 components used in altering the reference in the three previous

Figure 7.1 The Photofit reference for section 7.2

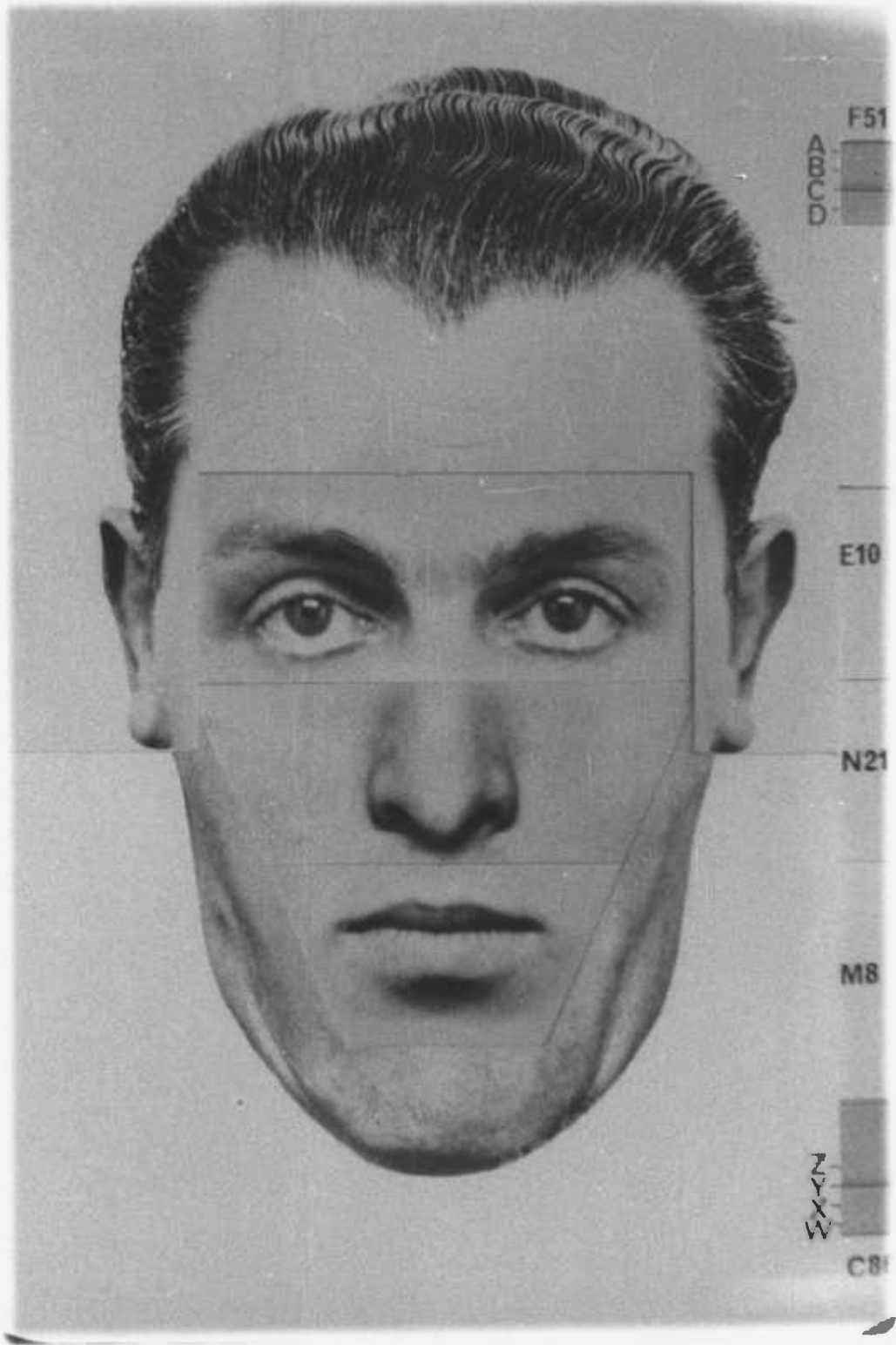


TABLE 7.3 A list of the altered faces

No.	Group	F	E	N	M	C
1	1	51	10	21	8	86
2		51	10	21	8	86
3	2	51	10	21	8	88
4		51	10	21	8	50
5		51	10	21	6	86
6		51	10	21	10	86
7		51	10	23	8	86
8		51	10	20	8	86
9		51	4	21	8	86
10		51	13	21	8	86
11		48	10	21	8	86
12		58	10	21	8	86
13	3	51	10	21	6	88
14		51	10	23	8	88
15		51	4	21	8	88
16		48	10	21	8	88
17		51	10	23	6	86
18		51	4	21	6	86
19		48	10	21	6	86
20		51	4	23	8	86
21		48	10	23	8	86
22		48	4	21	8	86
23	4	48	4	23	8	86
24		48	4	21	6	86
25		48	4	21	8	88
26		48	10	23	6	86
27		48	10	23	8	88
28		48	10	21	6	88
29		51	4	23	6	86
30		51	4	23	8	88
31		51	4	21	6	88
32		51	10	23	6	88
33	5	48	4	23	6	86
34		48	4	23	8	88
35		48	4	21	6	88
36		48	10	23	6	88
37	6	51	4	23	6	88
38		48	4	23	6	88
39		58	13	20	10	50
No.	Group	E		N		M
40	7	-	-	-	-	Up
41		-	-	-	-	Down
42		-	-	Up	-	-
43		-	-	Down	-	-
44		Up	-	-	-	-
45		Down	-	-	-	-
46		-	-	Down	-	Up
47		-	-	Up	-	Down
48		Up	-	Down	-	-
49		Down	-	Up	-	-
50		Down	-	-	-	Up
51		Up	-	-	-	Down
52		-	-	Down	-	Down
53		Up	-	Up	-	-
54		Down	-	Down	-	-
55		Down	-	-	-	Down
56		Up	-	-	-	Up
57		Up	-	Up	-	Up
58		Down	-	Down	-	Down
59		Up	-	Down	-	Up
60		Up	-	Down	-	Down
61		Down	-	Up	-	Up
62		Down	-	Up	-	Down
63		Down	-	Down	-	Up
64		Up	-	Up	-	Down

groups. The other is composed of the remaining 5 components used in altering the faces in group 2.

Group 7: This group consists of 25 faces which correspond to the feature relation variations tested in section 6.1 with the exception of three missing variations. These are the two cases in which the eye to eye separation was altered and the case in which the nose and mouth were moved upwards. This last variation was unintentionally missed out. It was found after the test, that another variation had in fact been used in its place.

Presentation of the 64 stimuli was randomised for each subject. Stimuli were presented sequentially for evaluation and no time limit was imposed. Subjects were provided with the reference face in its actual dimensions. They were free to place it anywhere in the testing cubicle. Each face was rated on a 1 to 5 integer scale. Values corresponded to the following subjective evaluations:

- 1 point - completely different from the reference
- 2 points - weak similarity to the reference
- 3 points - medium similarity to the reference
- 4 points - strong similarity to the reference
- 5 points - identical to the reference.

A total of 56 subjects participated in the experiment. Each subject was given the following written instructions: "You will be given two minutes to study the reference placed before you. This will be followed by the presentation of 65 slides. The first slide will correspond to the reference and will not count in the experiment. Your job is to evaluate the remaining 64 slides with respect to their similarity to the reference. The following scale should be used in the rating: (the above scale was reproduced at this stage). Each slide will be presented sequentially and no time limit is imposed. You have five buttons corresponding to the scores you assign to each slide. The reference may be placed anywhere in the cubicle and will remain with you throughout the experiment."

7.2.2 Results

This section will be divided into four parts corresponding to specific aspects of the results which will be examined individually.

A. The effects of single components

In section 7.1, a measure of similarity was assigned to the two most similar versions of each of the five component references. The effect of each version was described as the deviation which that version entailed from a possible maximum score of 5. These effects appeared in column 4 of table 7.2 and are reproduced in the first column of table 7.4. These same versions were used in the present experiment (group 2) in order to produce 10 faces, each differing from the reference face by only one component. The effect obtained for each of the components used in the alterations was measured in the following manner: first, the average score obtained for the two references (group 1) which were evaluated, was computed. The first reference had an average score of 4.64 and the second of 4.43, making an overall average of 4.53. Average scores for each of the faces in group 2 were then found and the differences between each of these scores and 4.53 were taken as measures of the effects of the ten components used in the alterations. These appear in column 2 of table 7.4. The values in columns 1 and 2 can now be compared. The object of the comparison can be illustrated by an example. Consider the two forehead-hairline versions F48 and F58. The latter was found to have the greater effect in the previous test, i.e. it was found to be less similar to the reference than F48. Examining the effect of these two versions in column 2, F58 is once again found to have had the greater effect. In other words, when two references were altered, in one case by F58 and in the other by F48, the former face was found to be less similar to the reference than the latter. Therefore, in the case of the forehead-hairline component, the order of effects between columns 1 and 2 correspond. The same applies to all other components except the nose strip. The different result occurring in this last case may be indicative of the effects of changes in feature relation attributes. It will be noted that the ten faces of group 2 differ from the reference not only by one component change but also by changes in feature relation attributes. It is essential that this should be taken into consideration when assembling a facial image made up of several components.

TABLE 7.4 The effects of single components

Version	Column 1 Effects (previous test)	Column 2 Effects (present test)
N23	2.19	0.76
N20	2.48	0.42
F48	2.40	1.74
F58	3.67	2.67
M6	2.45	0.69
M10	2.79	0.99
E4	2.71	1.01
E13	2.95	1.82
C88	2.98	1.35
C50	3.14	1.60

TABLE 7.5 The effects of changes in feature relation attributes: a comparison between the results of two experiments (section 6.1 and section 7.2)

Attribute Changes	Results from Section 6.1 using line drawings		Results from Section 7.2 using Photofit	
	Effects	Weber fractions	Effects	Weber fractions
FW(s)	1.03	0.075	0.86	0.079
FW(1)	1.56	0.070	0.22	0.073
MN(s)	2.21	0.250	0.23	0.375
MN(1)	1.78	0.200	0.16	0.273
ME(s)	0.52	0.083	1.04	0.083
ME(1)	1.34	0.077	0.35	0.077
NE(s)	0.67	0.125	0.48	0.107
NE(1)	1.29	0.113	0.00	0.097
MC(s)	1.04	0.167	0.87	0.115
MC(1)	1.92	0.143	0.29	0.103

B. The effects of changes in feature relation attributes: further evidence

In section 6.1, an experiment using line drawings was designed in order to measure the effects of changes in feature relation attributes. 28 drawings were prepared and split into four groups. It was felt necessary to use four separate groups because of the problems involved in ranking a large number of drawings simultaneously. This entailed the drawback that faces in different groups could not be compared. Results indicated that the effects of attribute changes were correlated to their Weber fractions and that consequently smaller attributes had greater effects than larger attributes when both were changed by the same amount.

Another measure of the effect of changes in feature relation attributes was obtained from the present experiment. However, three changes were made; to begin, faces were in Photofit form as opposed to line drawing form. Secondly, the variations were mixed in with 39 other faces which had been altered in terms of Photofit components. Finally, presentation was sequential rather than simultaneous within four groups. It was decided not to test the effects of changes in the eye to eye separation because of limitations in the Photofit kit.

The effects of attribute changes were measured in the same manner as described in section 6.1. Table 7.5 presents the results from both the line drawing and Photofit versions of the test. Weber fractions measured in the latter case, were based on the following values: 0.4 in. for M-N, 1.3 in. for M-C, 1.4 in. for N-E, 1.8 in. for M-E and 1.9 in. for FW. The incremental change for all variations was 0.15 in. A Kendall rank correlation test was applied to the data of the Photofit test and it was found that no correlation existed between the effects of the attribute changes and their respective Weber fractions. This result was in contradiction to that obtained in the case of the line drawing test. In evaluating the effects of each attribute (same method as in section 6.1) it was discovered that the effects of the M-N and M-E attributes for the present test were reversed. This is indicated in table 7.6 (columns 1 and 2).

Because of this discrepancy, a third test was conducted. This involved the presentation of 26 line drawings in a sequential manner. The variations corresponded to those used in section 6.1 with the exception of the two drawings containing different eye to eye separations. The test was

based on the following assumption: if results agreed with those of the first line drawing test, then the discrepancy mentioned above was due to the different nature of the stimuli, i.e. one type being in line drawing form, the other being in Photofit form. On the other hand, if results agreed with those used in the Photofit test, then the discrepancy was due to the manner of presentation, i.e. sequential as opposed to simultaneous within four groups.

Twenty subjects were used in the test and scores for similarity were assigned on a zero to five scale. The reference was that used as standard face two in section 6.1 and incremental changes were of 0.15 in. The presentation was randomised for each subject and no time limit was imposed. The viewing angle corresponded to that used in the Photofit test. A standard line drawing was included among the stimuli. Results were computed in the same manner as in both previous tests and the effects of attribute changes were again defined as the deviations with respect to the score obtained in the case of the standard line drawing. Before describing the results, a brief outline will be given of the parameters involved in all three tests.

Section 6.1. Line drawings were used and presentation was simultaneous within four groups. Results indicated that the effects of attribute changes correlated with their Weber fractions.

Section 7.2 using Photofit. Variations were presented sequentially and were mixed in with 39 other Photofit faces which had been varied in terms of component changes.

Section 7.2 using line drawings. Line drawings were presented sequentially.

Results from the last test indicated a strong correlation with Weber fractions on the Kendall rank correlation test (significance level was .023 on a one-sided test). The effects of the attributes appear in the last column of table 7.6. It will be seen that the results for both line drawing tests place the M-N and M-C attribute effects at the top of the list and that only the positions of the M-E and FW attribute effects are reversed. Furthermore, both tests indicate that the effects of changes in attribute size are correlated with their Weber fractions.

Figure 7.2 The Photofit reference with altered feature relation attributes.

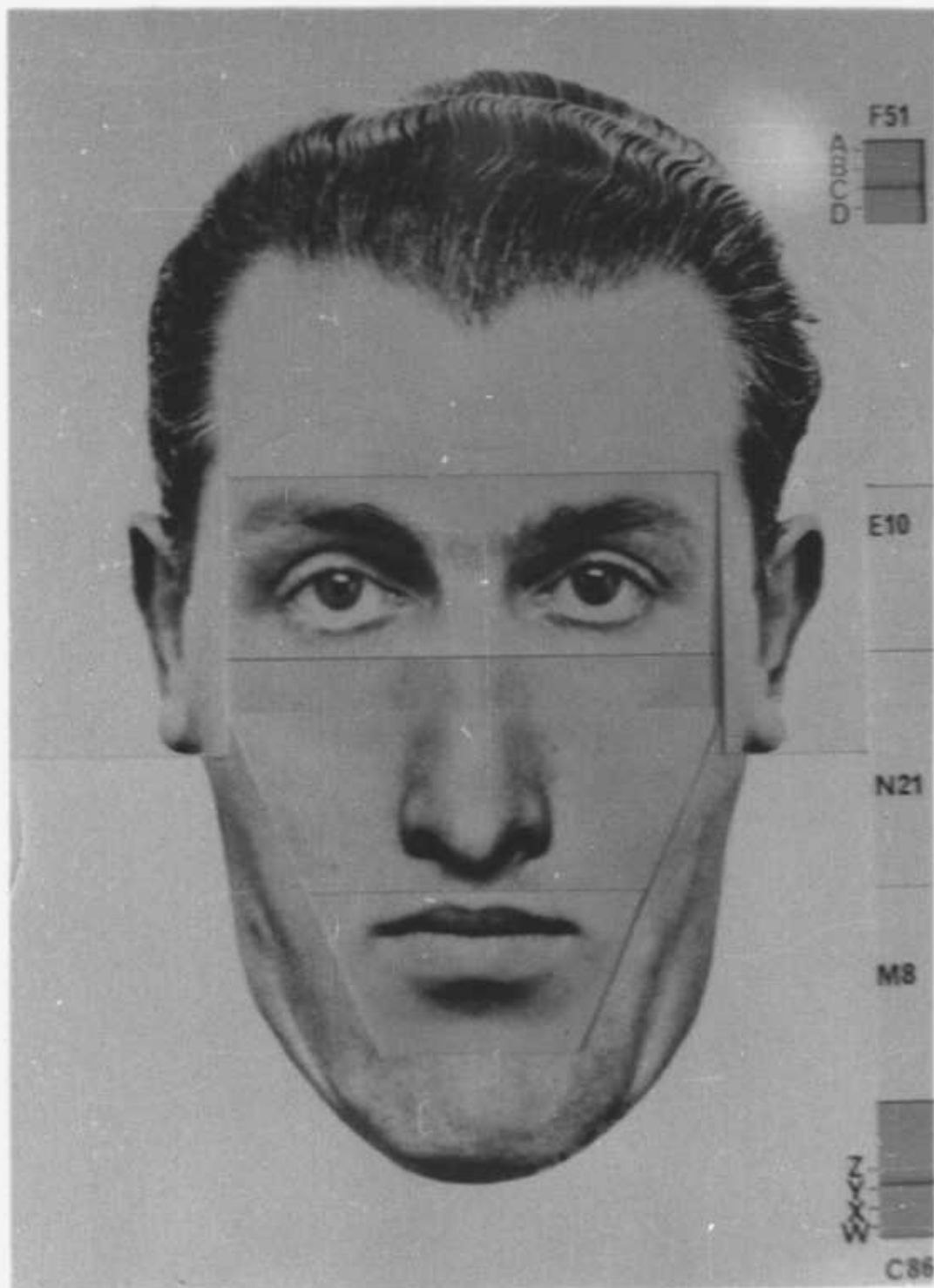


Table 7.6 The effects of the attributes

Attributes	Results using line drawings in section 6.1		Results using Photofit in section 7.2		Results using line drawings in section 7.2	
	Effect	Rank	Effect	Rank	Effect	Rank
M-N	3.99	First	0.39	Fifth	1.19	First
M-C	2.96	Second	1.16	Second	0.78	Second
FW	2.59	Third	1.08	Third	0.42	Fifth
N-E	1.96	Fourth	0.48	Fourth	0.61	Fourth
M-E	1.86	Fifth	1.39	First	0.69	Third

Table 7.7 The effect of component changes

Group	Component(s) in change	Effect		E4	N23		1.16		
1	None	Ref.	3	F48	N23		1.69		
	None	Ref.		F48	E4		2.28		
	C88	1.35		F48	E4	N23	2.30		
2	C50	1.60	4	F48	E4	M6	2.46		
	M6	0.69		F48	E4	C88	2.78		
	M10	0.99		F48	N23	M6	2.37		
	N23	0.76		F48	N23	C88	2.51		
	N20	0.42		F48	M6	C88	2.76		
	E4	1.01		E4	N23	M6	1.67		
	E13	1.82		E4	N23	C88	2.10		
	F48	1.74		E4	M6	C88	2.42		
	F58	2.67		N23	M6	C88	2.21		
	3	M6 C88		2.03	5	F48	E4	N23	M6
N23 C88		1.57	F48	E4		N23	C88	3.05	
E4 C88		2.21	F48	E4		M6	C88	2.74	
F48 C88		2.10	F48	N23		M6	C88	2.78	
N23 M6		1.44	E4	N23		M6	C88	2.62	
E4 M6		1.44	6	F48	E4	N23	M6	C88	3.17
F48 M6		1.80		F58	E13	N20	M10	C50	3.42

This is not the case for the Photofit test. On that basis, it is concluded that the main difference between the test of section 6.1 and the Photofit test of section 7.2 is due to the nature of the stimuli rather than to the method of presentation. In other words, the effects of attributes will be different according to whether the stimuli are line drawings or Photofit faces whilst the method of presentation is not critical. At the same time, it should be pointed out that there is some degree of correlation between the results for line drawings and those for Photofit even though the latter do not correlate with their Weber fractions. In particular, the ranking for the effects of the M-C, N-E and FW attributes seem to be roughly in agreement. The most apparent discrepancy between results concerns the reversed ranking for the M-N and M-E effects. Because of this, it can not be assumed that similarity effects in the case of Photofit faces are related to their Weber fractions. Other more complex factors must come in to play.

C. A check on the possible additivity of component effects

It was mentioned that the reference face used in the present experiment was made up of the reference components used in section 7.1, i.e. F51, E10, N21, M8 and C86. Furthermore, a set of 39 faces was prepared (groups 1 to 6) in which one to five components differed from the reference. The components used in the alterations were established in section 7.1. The 39 variations of the reference face are presented in table 7.7 which includes the components used in each alteration and the measured effects. The latter were measured in the same way as before, i.e. an effect represented the difference between the reference face score and the score obtained for the altered face whose effect was being measured.

The question of concern in this subsection is whether the effects of component changes are additive. For example, if a change in feature A produced an effect 'x' and a change in feature B produced an effect 'y', then how would the effect 'z' produced by the simultaneous change of features A and B be related to 'x' and 'y'? Would it be equal to the sum 'x + y' or would some interaction between features A and B create a third term in the sum? Obviously, if feature changes were found to be additive, it would be possible to establish beforehand the effect produced by a particular feature change. In that sense, similarity judgments could be

quantified and it would be possible to compare faces on the basis of numerical ratings of similarity.

The possibility of feature effects being additive will first be examined at the ranking (or ordinal) level. Additivity at this level must only satisfy the following conditions:

- (a) Consider the effects obtained individually by features A, B and C. Let these effects be x , y and w . Suppose that $x > y > w$.
- (b) Consider the effects obtained by simultaneously combining two of the three features A, B and C. Let these be 'g' for AB, 'r' for AC and 's' for BC. Furthermore, let 't' be the effect obtained by simultaneously combining all three features A, B, C.
- (c) For additivity at the ranking level, the following inequality must be satisfied, i.e. $t > g > r > s$.

To verify this, the following procedure is used: consider the effects of the component changes in group 2. Faces in this group consisted of references which had been altered by only one component. For the moment, only five of the components in group 2 are relative to the discussion, i.e. F48, E4, N23, M6 and C88. The remaining 5 form a special case which was studied in part A of this section. Now, consider the effects of the five relevant components:

F48	=	1.74
E4	=	1.01
N23	=	0.76
M6	=	0.69
C88	=	1.35

If the effects of components are additive in the ordinal sense, then it should be possible to predict the ranking of the effects in the case where the reference face is altered by 2, 3 or 4 components. Consider for example, group 3 where the reference was altered by 2 components. The effects obtained for these faces are presented in column 1 of table 7.8. In column 2 are the effects obtained by summing the individual effects of the two components involved in each change. For example, the effect obtained by altering the reference by F48 and E4 is 2.28. The sum of the individual effects of F48 and E4 is 2.75. In the former case, the effect

Table 7.8 The correlation between the effects of 2, 3 or 4 changes and the sum of the individual effects of the component changes involved

2 Changes			3 Changes			4 Changes		
Components	Effects Column 1	Sum ¹ Column 2	Components	Effects Column 3	Sum Column 4	Components	Effects Column 5	Sum Column 6
F48-E4	2.28	2.75	F48 E4 C88	2.78	4.10	F48 E4 N23 C88	3.05	4.86
E4-C88	2.21	2.36	F48 M6 C88	2.76	3.78	F48 N23 M6 C88	2.78	4.54
F48-C88	2.10	3.09	F48 N23 C88	2.51	3.85	F48 E4 M6 C88	2.74	4.79
M6-C88	2.03	2.04	F48 E4 M6	2.46	3.44	E4 N23 M6 C88	2.62	3.81
F48-M6	1.80	2.43	E4 M6 C88	2.42	3.05	F48 E4 N23 M6	2.60	4.20
F48-N23	1.69	2.50	F48 N23 M6	2.37	3.19			
N23-C88	1.57	2.11	F48 E4 N23	2.30	3.51			
N23-M6	1.44	1.45	N23 M6 C88	2.21	2.80			
E4-M6	1.44	1.70	E4 N23 C88	2.10	3.12			
E4-N23	1.6	1.77	E4 N23 M6	1.67	2.46			
Significance level of correlation	0.019		Significance level of correlation	0.0023		Significance level of correlation	0.117	

¹The term 'sum' means the sum of the individual effects involved in a particular change.

of the changes was the largest amongst all changes in group 3, i.e. all changes involving 2 components. In the latter case, the sum of the two individual changes produced the second largest summed effect. Using a Kendall rank correlation test to compare the ranking orders in columns 1 and 2, a significant correlation is found at the 0.02 level (one-sided test). Using the same procedure for columns 3 and 4 where three changes are involved, correlation is found to be significant at the 0.003 level. Finally, correlation between ranking orders for columns 5 and 6 which involve four component changes is significant at the 0.12 level. Thus it would appear that the effects of component changes are additive at the ranking level in the case of 2 and 3 changes but less so in the case of four changes.

The Kendall rank correlation test provides a measure of agreement between two sets of ranked values. It does not take into account the difference between absolute values in each set. It can be seen from table 7.8 that the values in columns 2, 4 and 6 are always larger than the values in columns 1, 3 and 5. Two possible explanations are given to account for this: first, it may be that the range over which faces were ranked (1 to 5) was not large enough to allow for a proper distribution in the values assigned to all faces. This might give rise to clusters of results which might normally be more spread out. The second explanation is that the effects of components are not independent. In other words, the effect produced by two simultaneous component changes is not the same as the sum of the individual effects of each of these changes. In the present test, the effect of 2 or more component changes was always smaller than the sum of the individual effects. This can be expressed mathematically by the use of weightings. Consider, for example, the case in which three changes are made; suppose these consist of changes in the mouth, nose and eyes components. Then we can use the following notation:

$$xdM + vdN + ydE = d(MNE)$$

$$x(0.69) + v(0.76) + y(1.01) = 1.67$$

where: x , v , y are weightings assigned to changes in the mouth, nose and eyes components.

dM , dN , dE are the individual effects produced by changes in the mouth, nose and eyes components, one at a time.

d(MNE) is the effect produced by simultaneous changes in the eyes, nose and mouth components.

The weightings for all five components can be computed in the case of 2, 3 and 4 changes. Average values are presented in table 7.9.

Table 7.9 The weightings for each component for the case of 2, 3 and 4 changes

Number of changes	Forehead-hairline weighting	Eyes weighting	Nose weighting	Mouth weighting	Chin weighting
2 changes	0.66	0.88	0.63	1.10	0.86
3 changes	0.65	0.64	0.60	1.01	0.74
4 changes	0.48	0.66	0.93	0.58	0.63

It can be seen from table 7.9 that weighting values for each component generally decrease with the increase in the number of total changes involved. However, the manner in which they decrease does not appear to conform to any particular mathematical relation. In fact, in the case of the eyes and nose components, the weighting values for four changes are larger than those for three changes. This would seem to imply that only a very complicated mathematical relation would be capable of satisfying a criterion of additivity above the ranking level, i.e. a relation which took into account the absolute values of the effects instead of simply their ranking order.

In an effort to find such a relation, the author made use of a formula proposed by Prosser et al. (1964) and developed by Lewis and Allnatt (1965). This formula related the effect produced by several simultaneous visual noise impairments judged by subjects on a television (using a 5 point rating scale) to the individual effects of the impairments when judged by themselves. The relation is described below in terms of its use for the author's results:

$$\left(\frac{1}{P_T} - 1\right) = \sum_{r=0}^n \left(\frac{1}{P_r} - 1\right) \dots\dots\dots (1)$$

where $P_1, P_2 \dots P_n$ are the mean scores obtained (each individually) when components 1, 2...n were used to alter the reference face.

Mean scores were computed in the following manner: the 1 to 5 rating scale used by subjects was replaced by a scale with the following values: 0, 0.25, 0.50, 0.75, 1. Furthermore, mean scores were the actual average scores based on a maximum of 1 and not the deviation scores. P_0 is the score obtained by the reference face and P_T is the score obtained for a face in which n components have been simultaneously changed. The value of n varies from 2 to 5 depending on the number of components changes in a particular face.

Relation (1) proved to be extremely accurate in the particular case for which it was designed. However, its use rested on certain assumptions which cannot be said to be satisfied in the author's case. Chief amongst these was the assumption of independence between the different types of impairment. In the present case, this would amount to assuming that the effects of changes in different components were independent, i.e. that feature relation attributes were not important. This assumption could definitely not be expected to hold in the present case.

Results using relation (1) appear in table 7.10. It will be seen that the difference between scores in columns 3 and 4 (see column 4) tends to increase with the number of changes involved. Furthermore, these differences are sometimes positive, sometimes negative and widely varying for a particular number of changes. Generally, results tend to reflect the non-uniformity obtained for the weightings presented in table 7.9. Thus it appears that a relation suitable for demonstrating the additivity of feature effects would have to be of a fairly complicated nature. It would have to take into account the interactions which occur between changed features, i.e. it would have to provide for the effects caused by changes in feature relation attributes. Furthermore, it would have to consider the fact that changes in different variations of the same component are likely to result in different effects.

The present subsection has examined the question of additivity of feature effects. Although this has been shown to exist at the ranking level, it has not been possible to present a formula capable of demonstrating additivity above the ranking level, i.e. the interval or ratio level. Reasons for this have been given. Principal among these is the fact that

complicated interactions occur when several features (components) are altered in a face. Not only are the features altered, but also feature relation attributes. Nevertheless, it is in some ways remarkable that additivity of effects should exist even if only at the ranking level.

Table 7.10 Scores using the relation $\left(\frac{1}{P_T} - 1\right) = \sum_{r=0}^n \left(\frac{1}{P_r} - 1\right)$

Components involved in altering the reference	Mean scores P_T for each altered face (maximum=1)	COLUMN 3 $\left(\frac{1}{P_T} - 1\right)$	COLUMN 4 $\sum_{r=0}^n \left(\frac{1}{P_r} - 1\right)$	Difference ¹ between values for columns 3 and 4
F48	0.446	1.242		
E4	0.629	0.590		
N23	0.692	0.445		
M6	0.710	0.408		
C88	0.545	0.835		
M6 C88	0.375	1.667	1.374	+0.293
N23 C88	0.491	1.037	1.411	-0.374
E4 C88	0.330	2.030	1.556	+0.474
F48 C88	0.362	1.762	2.208	-0.446
N23 M6	0.522	0.916	0.984	-0.068
E4 M6	0.522	0.916	1.129	+0.213
F48 M6	0.433	1.309	1.781	-0.472
E4 N23	0.594	0.683	1.166	-0.483
F48 N23	0.460	1.174	1.818	-0.644
F48 E4	0.312	2.205	1.963	+0.242
F48 E4 N23	0.308	2.247	2.408	-0.161
F48 E4 M6	0.268	2.731	2.371	+0.360
F48 E4 C88	0.192	4.208	2.798	+1.410
F48 N23 M6	0.290	2.448	2.226	+0.222
F48 N23 C88	0.254	2.937	2.653	+0.284
F48 M6 C88	0.214	3.673	2.616	+1.057
E4 N23 M6	0.464	1.155	1.574	-0.419
E4 N23 C88	0.357	1.801	2.001	-0.200
E4 M6 C88	0.277	2.610	1.964	+0.646
N23 M6 C88	0.330	2.030	1.819	+0.211
F48 E4 N23 M6	0.232	3.310	2.816	+0.494
F48 E4 N23 C88	0.121	7.264	3.243	+4.021
F48 E4 M6 C88	0.196	4.102	3.206	+0.896
F48 N23 M6 C88	0.187	4.348	3.061	+1.287
E4 N23 M6 C88	0.228	3.386	2.409	+0.977
F48 E4 N23 M6 C88	0.089	10.236	3.651	+6.585
Reference face (average of the 2 faces used)	.884	0.131		

¹The positive signs mean that the values of column 3 were larger than the corresponding values of column 4.

D. A comparison between the effects of changes in components and changes in feature relation attributes

The effects of feature relation attributes were discussed in part B and those of component changes in part C. This subsection compares the effects of both types of changes. These are presented below, beginning with the largest effect:

(1)	F48	=	1.74
(2)	C88	=	1.35
(3)	ME(s)	=	1.04
(4)	E4	=	1.01
(5)	MC(s)	=	0.87
(6)	FW(s)	=	0.86
(7)	N23	=	0.76
(8)	M6	=	0.69
(9)	NE(s)	=	0.48
(10)	ME(1)	=	0.34
(11)	MC(1)	=	0.29
(12)	MN(s)	=	0.23
(13)	FW(1)	=	0.22
(14)	MN(1)	=	0.16
(15)	NE(1)	=	0.00

Any comparison between the effects of both types of changes must be made cautiously. First of all, the five components used in altering the reference can only be said to have satisfied one criterion, i.e. they were chosen as the most similar versions to the individual reference components when compared with a larger group of 14 versions. This does not in any way imply that they are representative of their component class in the sense of being 'average' representations, assuming that such things exist. It can only be suggested that they might result in lower effects than the other 13 versions judged in section 7.1. Even then, it was seen in part A, that this was not the case for the nose component. It must also be kept in mind that component changes automatically entail feature relation changes and that these will contribute to the size of the effects.

Having stated these reservations, it is possible to consider certain general points about the order of effects obtained above. For

instance, it can be seen that certain changes in feature relation attributes had greater effects than certain component changes. This applies in the case of the eyes, nose and mouth components. In a limited sense, this supports previous conclusions which suggested that contours were the most important set of attributes but that some discrete features might be less important than certain feature relation attributes. Of course, the importance of an attribute is not the same as the effect it produces when changed and the eyes, nose and mouth components are not the same as the eyes, nose and mouth attributes. Nevertheless, it is interesting that certain changes in feature relation attributes should be found to have a greater effect than certain changes in components with respect to a similarity criterion.

7.3 Summary

Two experiments were reported in this chapter. The first (section 7.1) involved similarity ratings for different variations of Photofit components with respect to chosen reference components. The main purpose of the experiment was to determine the most similar variation of each component with respect to the reference components. These were then used for the second experiment which examined several aspects:

- (a) The effects of single components used to alter a reference face,
- (b) The effects of changes in feature relation attributes using Photofit,
- (c) The possible additivity of feature effects,
- (d) The comparative effects of changes in Photofit components and changes in feature relation attributes.

All effects were measured according to a criterion of similarity and all stimuli were in Photofit form. A Photofit face was made up of the reference components used in the first experiment and this served as the reference face for the second experiment. 64 variations of this face were prepared by either altering individual components or by altering the feature relation attributes. The components used to alter the reference were established in the first experiment.

The following general conclusions were reached from the results of the second experiment:

(a) The rankings for the effects of single Photofit components judged with respect to single reference components were the same as those for the case when these components were used to alter a reference face made up of the reference components. Only one exception occurred in the case of the nose component.

(b) Feature relation attributes which were altered in Photofit faces did not produce the same effects as in line drawings. It was established that the main reason for this was the different nature of the stimuli. The effects of altered feature relation attributes in Photofit faces did not correlate with their Weber fractions.

(c) The effects of Photofit components used to alter the reference were found to be additive at the ranking level. However, because of the effects of altered feature relation attributes involved in component changes, it was not possible to establish a relation which would satisfy a criterion of additivity at the interval or ratio level.

(d) Some changes in feature relation attributes were found to have a greater effect on similarity ratings than certain changes in Photofit components.

CHAPTER VIIISUMMARY OF THE RESULTS

This chapter reviews the evidence presented in the author's experiments and discusses possible areas of application.

In chapter II, some of the research dealing with visual pattern recognition was described. Several suggestions and models were proposed in order to explain how pattern recognition occurred. However, the great majority of the stimuli used in both the psychological and neurophysiological research were of the 'simple' type, i.e. geometrically simple patterns. In most cases, a stimulus could be defined in terms of a few straight lines or curves, a set of angles and perhaps a few surface areas. In the author's research, the problem of defining a set of features for a stimulus such as the human face is far greater. Keeping in mind the general approach used in defining a pattern, the author proposed in section 3.2 that a human face be described in terms of three types of facial features: contour attributes, discrete features and feature relation attributes.

The main purpose of the research reported in this thesis was to evaluate the relative importance of different facial features. In studying facial features of a particular type, the main problem to confront was one of isolating or controlling the effect of features not under study in a particular experiment. It will be remembered (section 2.1) that a considerable difference of opinion existed in the case of research using stabilised images as regards the relative importance of lines and angles as perceptual elements. Obviously, the principal obstacle in solving that particular dichotomy was the fact that lines could not be studied independently of the presence of angles or vice-versa. In the case of human faces, the same problem exists. For example, feature relation attributes cannot be studied in the absence of discrete features or contours. Because of this problem, certain indirect approaches were used in each of the experiments and consequently, any statement about the relative importance of a particular attribute or set of attributes was made with the proviso that the influence of other attributes had been present in the case of each evaluation. Nevertheless, in examining the results as a whole, a general organisation does seem to emerge in the ranking of features with respect to their importance for recognition.

The first experiment (section 5.1) was intended to be general and exploratory. In it, recognition in the presence of visual noise was investigated. This technique was seen as an alternative to tachistoscopic presentation methods in that it allowed stimuli to be presented for longer durations but still with a finite probability of incorrect recognition.

Several variables were investigated in the experiment. Each subject was given the same stimuli to recognise on four successive sessions.

Improvement was found to be linear over the four sessions, i.e. subjects recognised the stimulus faces in progressively noisier conditions in a linear^{*} fashion (figure 5.2). Subjects tested at one day intervals did not perform differently from subjects tested at one week intervals. Subjects given 7 second stimulus exposures did perform better

than the 3 second group but also made significantly more incorrect recognitions. This was explained by suggesting that the 7 second group felt more confidence in their initial impressions but that these could be misleading. This hypothesis was not however tested in later experiments.

Three different stimulus views were tested in the experiment: full-frontal, full-profile and half-profile. No significant difference was found in the recognition scores for each view although some faces were better recognised with one view than with another. The overall non-significance of the 'views' factor was attributed to the presence of all views on the subjects' reference board.

In studying both the confused and non-confused pairs of stimulus faces in the experiment, it was proposed that contours were the most important set of attributes. This suggestion, although extremely subjective was seen as a tentative starting point for subsequent experiments.

The second experiment (section 5.2) investigated recognition in the presence of masking. Two areas were obscured by electronic methods: with top masking, all facial detail above the eyebrow line was covered and with bottom masking, all information below the tip of the nose. Bottom masking resulted in significantly higher recognition scores (80% vs. 51%) than top masking. This was attributed to the fact that the upper half of the face contained more contour information than the lower half. However, the actual contribution of the forehead width attribute present in the bottom masking case, was left undetermined. It was also proposed that certain facial features present in both masking cases could not be responsible for the large discrepancy which occurred between both sets of results.

The results obtained with electronic masking were compared with those obtained by the use of 'natural' masking, i.e. the stimulus faces were masked by means of black cardboard sheets. It was felt that the use of electronic masking might have some unpredictable 'distracting' effect

* i.e. when noise is measured in decibels.

on subjects by virtue of the novel effect which it produced. However, no significant difference was found.

In the second experiment, the relative importance of the discrete features, i.e. the eyes, the nose and the mouth could not be directly compared. This was the purpose of experiment 3 (section 5.3). In it, masking was used to cover either the eyes, the nose or the mouth by means of masking strips which terminated within the face or extended over its boundaries. Recognition was found to be significantly lower in the case of the eyes strip being masked. Results for nose and mouth masking were very similar to those obtained in the no masking control case. No significant difference was found between the two lengths of masking strip used (wide and narrow masking).

As a general summary of the two masking experiments, it can be said that the portion of the face which includes the eyes and all detail above them is definitely more important than the remaining portion in terms of both contour and discrete feature attributes. However, the presence of feature relation attributes in both portions has not been considered and it is this fact which led to the design of the fourth experiment (section 6.1).

Experiment 4 was designed to measure the subjective effects produced by altering the sizes of different feature relation attributes. Subjects were provided with a reference face and asked to judge a group of altered faces with respect to it on a similarity basis. The following average effects were found for the attributes tested:

Attribute	Description	Effect
M-N	Mouth to nose separation	3.99
M-C	Mouth to chin tip separation	2.96
FW	Forehead width	2.59
N-E	Nose to eyes separation	1.96
M-E	Mouth to eyes separation	1.86

The larger the effect, the greater was the degree of change caused by a particular attribute. The effects presented above are not important in terms of absolute values but rather in terms of their relative values since the choice of rating scale will influence the size of the absolute values. All faces tested in the experiment were in line drawing form and alterations in feature relation attributes were of two sizes. Two reference faces were tested and these were in either line drawing or photographic form. However, the use of two alteration sizes, two reference faces and two types of reference did not result in significantly different effects. In the author's view, this provided strong evidence for the ordering of effects presented above. Effects were also found to correlate with their Weber fractions implying that changes in smaller attributes produced greater effects than equivalent changes in larger attributes.

A possible criticism concerning experiment 4 was the fact that the altered faces (the stimuli) were presented as four arbitrarily chosen groups, each judged separately instead of all being presented simultaneously. The reason underlying this procedure was the fact that the entire set of stimuli (28 faces) could not possibly have been judged all in one presentation. Because of this criticism, it was decided to repeat the experiment and results will be described in the summary of experiment 7 (section 7.2). For the moment though, experiment 5 will be discussed in order to preserve the chronological order of the author's work.

Experiment 5 (section 6.2) made use of the Photofit kit and examined the question of face recognition in conditions of very short exposure duration, i.e. 40 and 120 milliseconds. Because of the versatility of the Photofit kit, it was possible for the author to interchange different facial components between faces. In fact, 4 reference faces, each composed of three chosen facial units (1: the forehead-hairline section 2: the eyes, nose and mouth sections as a unit, and 3: the chin section) were used to prepare 64 different faces, i.e. $4^3 = 64$. An additional set of 4 faces, each shown 4 times was included in the test. One of these was the fifth reference face and the other three were irrelevant faces. Subjects were given the task of recognising which reference face was presented. However, they were not told that the majority of the stimuli (60) were actually composed of features from different references. In order to allow for the fact that some irrelevant

faces would be presented, subjects were given the option of pressing a NONE button to indicate that their choice was not one of the references. This NONE button could obviously be used when subjects were presented with combination faces.

Results indicated that subjects based their responses more often on the forehead-hairline section than on either the chin section or the eyes-nose-mouth unit. Furthermore, recognition of the actual references (not combinations) was above the chance level, indicating that recognition was still quite accurate even with very short exposures. The percentage of NONE responses did increase as expected when exposure duration was lengthened, implying that recognition or rather discrimination capacity was improved with longer exposure. These results provided clear evidence in the author's opinion, that the contours present in the forehead-hairline section were the most important attributes for recognition. Furthermore the eyes-nose-mouth unit was found to be more important than the chin section although it was not established whether any of the attributes in this unit were more important by themselves, than the chin section. It should be remembered that the eyes-nose-mouth unit contains not only all three discrete features but also several feature relation attributes.

Experiment 5 was followed by two more tests using Photofit (chapter VII). The first of these was concerned with similarity ratings made on different variations of each of the five Photofit components with respect to chosen component references. The main purpose of the experiment (section 7.1) was to determine the most similar variation (amongst 14) of each component with respect to its reference. These were then used for the second experiment which examined several points: the effects of components used to alter a reference face, the effects of changes in feature relation attributes using Photofit, the possible additivity of feature effects and the comparative effects of changes in components and changes in feature relation attributes.

All effects were measured according to a similarity criterion and all stimuli were in Photofit form. A reference face was made up of the reference components used in the previous experiment and 64 variations of this face were prepared by altering either individual components (39) or feature relation attributes (25). The components used in the alterations were chosen from the previous experiment. The following general

results were obtained from the experiment. First, it was found that the rankings for the effects of components judged with respect to single reference components were the same as those for the case when these components were used to alter the reference face. Only one exception to this finding occurred in the case of the nose component. Feature relation attributes which were altered in Photofit faces did not produce the same effects as in line drawings (section 6.1). In particular, the effects of the M-N and M-E effects were reversed. Consequently the effects for the case of Photofit faces did not correlate with their Weber fractions. It will be remembered that in the case of line drawings, simultaneous presentation techniques involving four separate groups were used whilst in the case of Photofit faces, sequential presentation was used. It was suggested that either the different method of presentation or the different nature of the stimuli would account for the different results just mentioned. A third test using sequential presentation of line drawings was performed and it was found that effects correlated with their Weber fractions. Consequently, it was concluded that the different results obtained in sections 6.1 and 7.2 were due to the different nature of the stimuli and not the method of presentation.

Another aspect of the results to be examined was the possible additivity of feature effects. These were found to be additive at the ranking level but no relation was found capable of demonstrating additivity above this level. This was attributed to the fact that interaction effects also came into play when component effects were added. These were due to the alterations which occurred in feature relation attributes when several components were used to alter the reference face. It was concluded that the effects of feature relation changes precluded the possible use of a mathematical relation capable of satisfying a criterion of additivity above the ranking level.

A final aspect of the results to be examined concerned the relative size of the effects produced by both component changes and feature relation changes. Although the effects which were examined were only applicable to one particular face, i.e. one set of Photofit components, the ordering of the effects confirmed the general trend of results presented in previous experiments. For example, the two components containing contour attributes were found to produce the greatest effects. In particular the forehead-hairline section produced a greater effect

than the chin section. Amongst the components containing discrete features, the eyes strip was found to be more important than either the nose or mouth strips. Finally, certain feature relation changes resulted in greater effects than certain discrete feature changes.

As a summary of the research reported in this thesis, the following general conclusions can be drawn from the results:

- (a) Contour attributes, especially those in the upper half of the face are the most important facial attributes for recognition;
- (b) Amongst the discrete features, the eyes are found to be more important than either the nose or mouth;
- (c) Certain changes in feature relation attributes produce greater effects on similarity than others. Furthermore, certain of these changes appear to produce greater effects than certain changes in discrete features.

Probably the most direct application of the results reported in this thesis lies in the field of criminal identification. Here the problem is not one of recognition per se, but one of matching by recall. Limitations in both the identification kit and the user make this task extremely difficult. For that reason, it becomes crucial to have some idea of the relative importance and effect of different facial features. Thus, although recognition is probably largely an unconscious and automatic process, there is probably scope for manipulating the recall process in reconstructing a suspect's face. This manipulation is, in the author's opinion, most likely to be successful if it is based on some ordered system of priorities. The same applies in the case of a possible face 'pattern recognition system'. The results of this thesis may be of some help in both cases.

APPENDIX A

Appendix A contains circuits for some of the apparatus used in the author's experiments. Figure A.1 represents the circuit of the 3 or 7 second timing gate. The output of this gate was used to time the output from control unit A. Figure A.2 illustrates the circuit used in adding the visual noise to the video signal. The noise frequency distribution from the General Radio white noise generator is shown in figure A.3. Finally, figure A.4 represents the circuit for the frame counter. This apparatus was used to time the output from control unit A in the case of the Photofit experiment reported in section 6.2.

Figure A.1 The 3-7 second timing gate

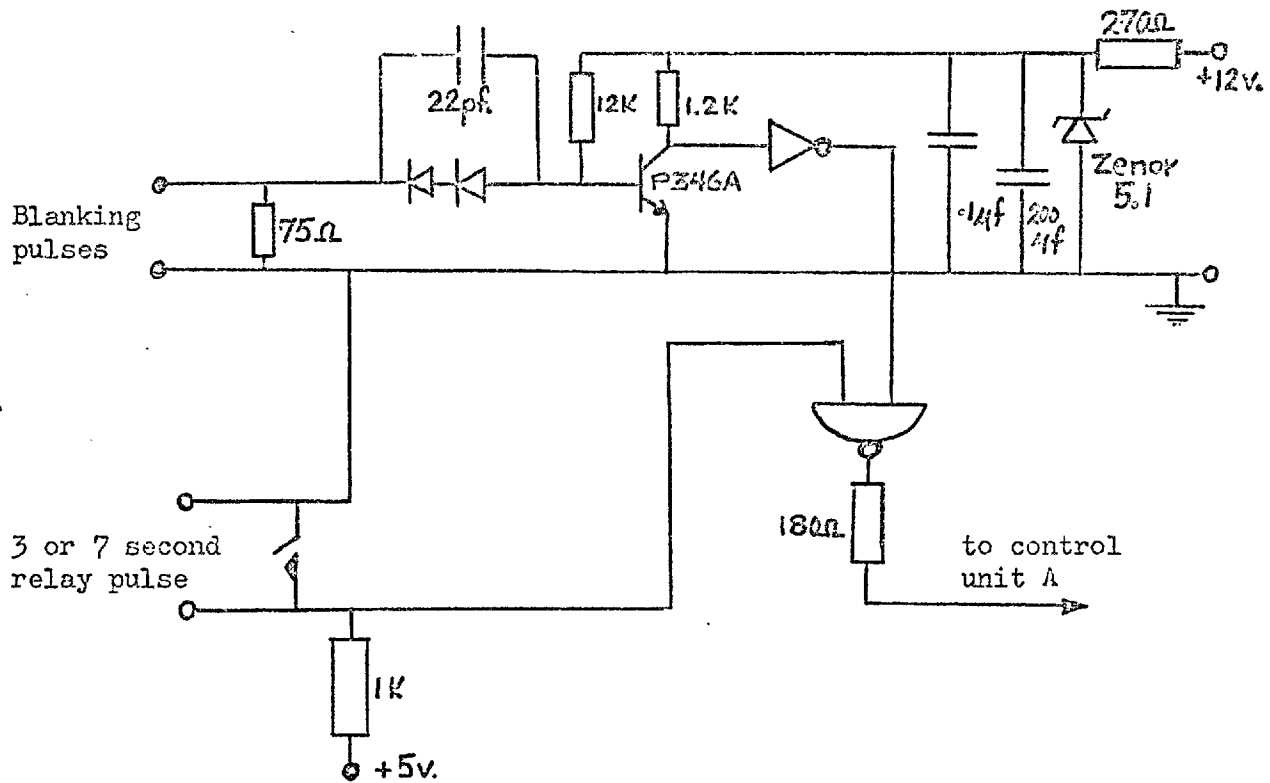


Figure A.2 The noise-video gate

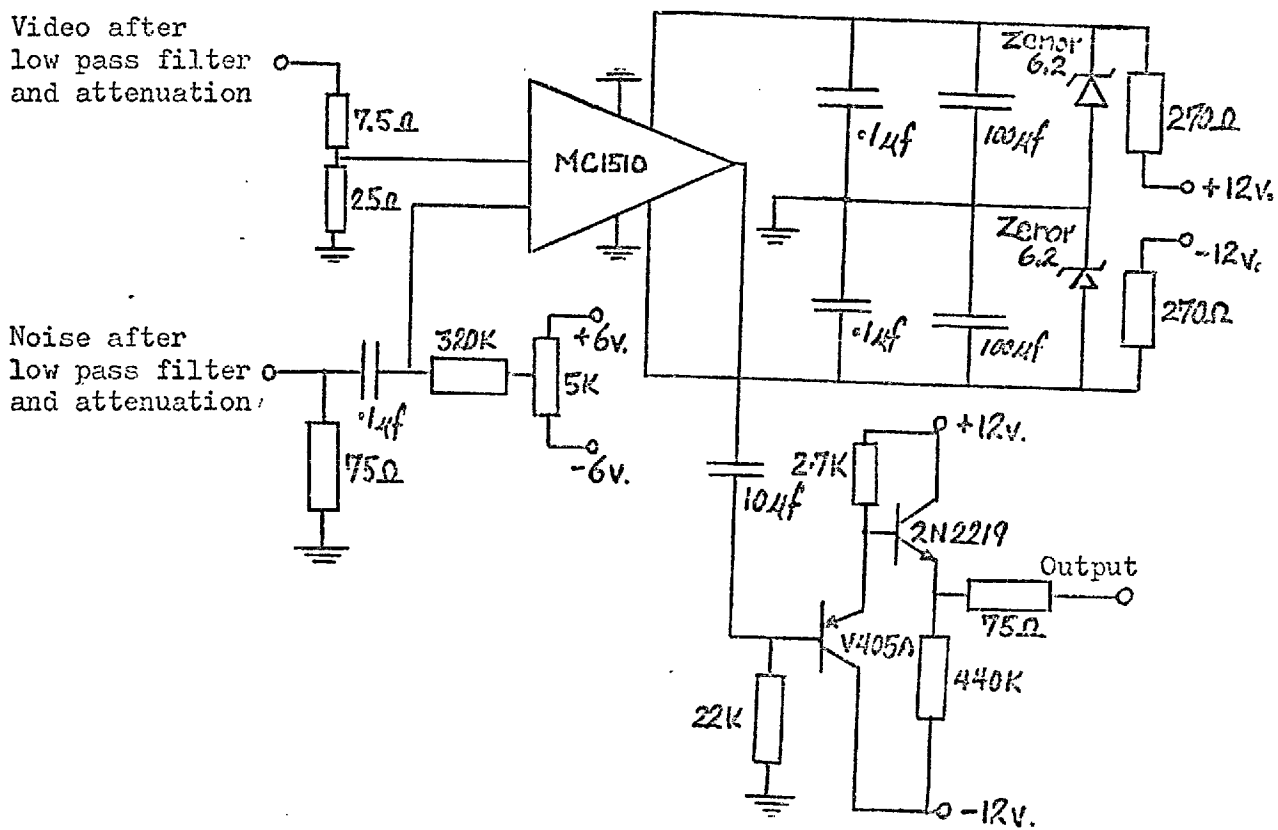
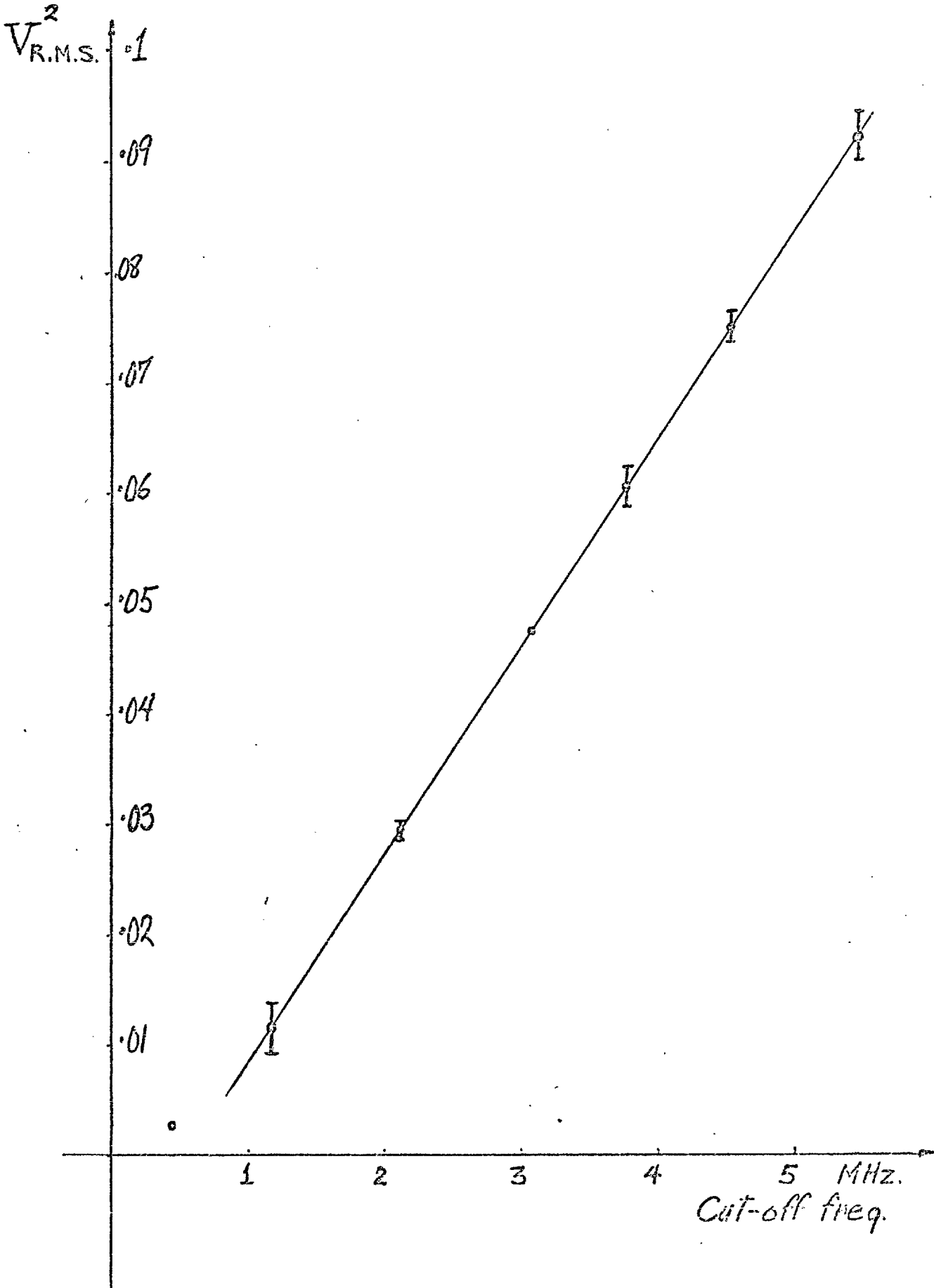


Figure A.3 Frequency distribution of the visual noise from the General Radio noise generator.



APPENDIX B

THE RECOGNITION OF FACIAL EXPRESSIONS

This section reports two experiments performed by the author in connection with the recognition of facial expressions. Both experiments are essentially replications of previous work.

B.1 Introduction

The introduction presents a general review of some of the research which has dealt with facial expressions. The areas investigated and the methodological problems encountered are discussed.

Probably the best known early work dealing with the study of expressions was that of Darwin (1873). In it two problems were discussed, namely whether facial expressions were innate or acquired and secondly, how recognition of the underlying emotion occurred. Darwin concentrated on the first of these questions and concluded with an argument in favour of innateness. Basically, he considered that facial expressions were originally crude physiological actions consciously performed and that these slowly evolved into innate reflexes associated with particular emotional states. As for the actual recognition process, Darwin acknowledged in a general manner that different facial configurations were understood and interpreted as implying different emotions and that in some cases, a particular feature might act as a strong clue by itself in aiding recognition.

Darwin's stand in favour of innateness was strongly attacked at the beginning of this century by the empiricists who believed that facial expressions developed maturationally and socially rather than being innate. For an interesting exposé of the two schools of thought, see Honkavaara (1961). Probably the most acceptable solution to the dichotomy can be found in terms of a compromise such as that proposed by Thomson (1941). In a revealing study of expression development in blind and seeing children, she concluded that certain actions such as laughing and crying were innate but that all expressions were subject to biological and social development. Her evidence came from the fact that blind children lacked certain expressions and that their range of 'expressiveness' for known expressions tended to be more uniform than that of seeing children.

Because of the intractable nature of the problem of expression development, research into facial expressions has tended to drift to the more tangible question of recognition. Here however, a host of methodological problems have resulted in a great deal of ambiguous and confusing data. See Vine (1969) for a review of recent work in this area. Broadly speaking, recognition experiments can be criticized for one of 3 reasons: the nature of the stimuli, the method of indicating recognition and the interpretation of the results.

Stimuli often consisted of drawings or photographs of faces portraying particular emotions. Some were presented in full-frontal view, others in profile. Generally, it was not realised that such stimuli were completely lacking in contextual information implicit in the display of an emotional state. This information is dynamic in the temporal sense. Emotions are expressed over finite periods of time and cannot be truly represented in a static photograph. This drawback was partially remedied in some research by making use of film, e.g. Thomson (1941), Frijda (1953).

An important aspect in the recognition of emotional states concerns the relative contribution of the available types of contextual information: verbal content, body movement, facial expression and other less specific cues. These are all considered as components which may be relevant to communication and the actual contribution of the facial expression component has yet to be assessed, except at a qualitative level, e.g. Ekman (1965). It may indeed be that facial expressions are useful but not necessary in most dyadic interactions. For example, telephone conversations which make use of verbal and tonal content may be sufficient in most cases for proper interpretation of most emotional states. These questions remain to be answered. In the meantime though, studies in expression recognition can be useful as part of the larger area of research concerned with non-verbal communication.

Probably the most important criticism to be levelled against most recognition experiments has been the method by which subjects have been told to signal recognition. In most cases, word-labels have been used. For instance, the experimenter has given an actor a set of emotions to portray. The words used to indicate the desired expressions to be portrayed by the actor have subsequently been accepted as the proper labels for the recognition task. In other words, subjects only scored successfully if they used the same words in their recognition. This method of

testing tacitly assumed that the actor was, in some way, capable of portraying 'exactly' some emotion specified by a particular word. Obviously this led to low recognition scores. An improvement was obtained when the word-labels were actually available as references to the subjects or if synonyms were allowed. The acceptability of the latter was usually decided upon by a group of judges. However, in the long run, the semantic difficulties inherent in specifying some facial expression by a specific word could not be avoided.

A solution to the problem was suggested by Schlosberg (1952, 1954) and developed at greater length by Osgood (1966). Instead of using word-labels as their responses, subjects made use of a 'semantic space' into which stimuli could be classified. The semantic space can be described as a bounded volume, with each point within it specified by three orthogonal co-ordinates. Each orthogonal axis is bipolar and can be scaled to the point of being continuous rather than discrete. The axes chosen by Schlosberg were called the sleep-tension, the pleasant-unpleasant and the attention-rejection axes. A particular stimulus was assessed in the following manner: a subject assigned a value of x to the sleep-tension component (or dimension) y to the pleasant-unpleasant component and z to the attention rejection component. This resulted in a point (x,y,z) in the semantic space corresponding to a particular stimulus. Other subjects might assign very similar values to the 3 components of the stimulus in which case, a cluster of points would result around (x,y,z) . The tighter the cluster, the greater was the inter-subject agreement. A very loose cluster indicated considerable variation of agreement amongst subjects. Different stimuli could be compared within the semantic space by measuring the geometrical distances separating clusters.

It will be seen that the purpose of the semantic space is to get rid of the labelling problem and to provide everyone with the same references. Furthermore, the idea of emotional intensity or magnitude is conveyed by means of scaled axes. In Osgood's case, a set of 40 stimuli were tested. This resulted in points which could almost all be bounded within a volume similar to a truncated pyramid. The choice of the axis labels is of course important and must be done with the idea of describing completely any emotional state by only 3 components. Ekman (1965) criticised the labels used by Schlosberg on the grounds that they were not independent. He decided to reject the attention-rejection axis. Further-

more, he hypothesised the following: "head cues carry information primarily about what particular affect is being experienced (emotion conveyed through expression) but provide relatively little information about the intensity of the affect or the level of arousal; body cues reverse this pattern, communicating information primarily about the level of arousal or the degree of intensity of the affective experience, but providing relatively few cues about what particular affect is being experienced". In other words, Ekman classified emotions according to their nature and intensity, each to be evaluated along one axis, i.e. the sleep-tension axis for the intensity and the pleasant-unpleasant axis for the nature of the emotion.

That the nature of an emotion should be so specific as to be judged along one axis is highly doubtful and Ekman did in fact propose 3 new axes for his semantic space: aroused-unaroused, active-passive and physically mobile-physically immobile. Later, Ekman and Friesen (1967) found more evidence for their previous hypothesis using the Woodworth system of six emotion categories. They also separated the perception of emotion into two states, the general and the specific. The general state of perception was the recognition of a gross or general state of emotion and the specific state involved the recognition of the particular emotion. To them, the face was essentially a display system whilst the body indicated adaptive or dynamic efforts regarding the affect.

Ekman and Friesen's conclusions, which are perhaps too general, can be contrasted to the highly specific proposals made by Tomkins and McCarter (1964). In their research, they assert that recognition studies prove the existence of eight 'primary' affective states, each with its own 'subcortical innate program' and with characteristic and recognisable facial expressions. Although Tomkins and McCarter may have obtained certain clear results, it is doubtful whether they have sufficient grounds to make unequivocal statements regarding the existence of 'primary affects' and 'subcortical innate programs'. Such assertions are not acceptable at this stage. Nor are the results obtained by Cuceloglu (1970) concerning the perception of facial expressions in different cultures. The latter proposed a cross-cultural semantic space with the following bipolar dimensions: pleasantness, irritation and non-receptivity. These accounted for more than 70% of inter-subject variance. However, in the author's opinion, the use of highly schematized line drawings invalidates the

results obtained unless conclusions are completely reformulated by introducing the concept of stereotypes. It is after all unrealistic to propose that line drawings, especially those used by Cuceloglu, can be acceptable as representations of emotional states. Line drawings are probably recognised on the basis of their stereotyped nature and in that sense, any consistency between subjects is useful only in so much as it establishes that stereotypes do exist. As a general summary to this introduction, it can be said that three areas of research dealing with facial expressions have been investigated:

(a) The origin of facial expressions: two views have been presented to account for the origin of facial expressions. One view suggests that facial expressions are innate, the other that they are learned through social mimicry with possible biological development. Research into this area was prominent in the first half of this century but has declined in recent years. Neither view seems to be wholly acceptable.

(b) The recognition of facial expressions: a great deal of work has gone into the evaluation of our recognition capability concerning facial expressions. However, results have often differed or been contradictory for methodological reasons which have been described and which will be summarised below.

(c) Recognition in different cultures: Darwin based his belief in the innateness of facial expressions partly on the study of expressions in different cultures. Unfortunately, his data was rather anecdotal. A modern example of a study of expression recognition for 3 cultures was that of Cuceloglu. Here however, the use of line drawings of a highly stylised nature made the results unacceptable in the author's view.

It was seen in the introduction that different methodological approaches to the problem of recognition created a wide variation in the results. Three main criticisms were levelled against most of the work. These are summarised below:

(a) The type of stimulus: line drawings and photographs were described as lacking in contextual information usually present in the display of an emotion. This information is both temporal and spatial and appears in different perceptual modes: visual, auditory and perhaps tactile. The use of film was recognised as an improvement in the choice of stimuli.

(b) The method of indicating a recognition response: the use of word-labels was criticised especially when references were not provided and synonyms were not accepted. The use of a semantic space was seen as an improvement in evaluating expressions since it did not require the use of vocabularies which differed for every individual. Furthermore, the dimensions of a semantic space being scaled, allowed for judgments of magnitude or intensity. However, it was seen that criticisms were levelled against the choice of axes in establishing the space dimensions (this is usually done on the basis of factor analysis).

(c) The interpretation of results: experiments using the same or different methods of testing have come up with different interpretations. Some have been very general, others highly specific. Very few have been in complete agreement.

B.2 The first experiment

B.2.1 Aim

The aim of this experiment is simply to point out by means of an example, the problems inherent in testing for the recognition of facial expressions. Most of these problems were discussed in the introduction. The experiment is in the traditional format, i.e. it makes use of photographs and word-labels. For that reason, many of the drawbacks discussed will become evident.

B.2.2 Method

In preparing his stimuli, Osgood (1966) made use of a set of 40 words which are reproduced in table B.1 (column 1). These same words were used by the author and given to a professional actor (Alfie Bass) in order to make up a set of stimuli. 5 'neutral' face stimuli were prepared instead of only one and this resulted in a set of 44 stimuli, each prepared as slides in the same manner as in previous tests. These slides were presented randomly to a group of 21 subjects and recognition was indicated by means of a word describing the emotion portrayed by each slide. No time limit was imposed. Subjects were not given any references and were told that they could use the same word to describe different stimuli.

Table B.1 The words most used in recognising facial expressions

Number	COLUMN 1 Actor's word list	COLUMN 2 Most used words	COLUMN 3 No. of most used words	COLUMN 4 No. of synonyms
1	Silent laughter	Happy	7	7
2	Puzzlement	Puzzlement	3	5
3	Dreamy sadness	Thoughtfulness	2	4
4	Surprise	Surprise	11	5
5	Annoyance	Anger	4	6
6	Neutral	Neutral	15	1
7	Boredom	Tiredness	2	4
8	Glee	Happy	9	8
9	Rage	Anger	2	13
10	Quiet pleasure	Happy	8	7
11	Fear	Worry	4	8
12	Sulkiness	Disapproving	3	2
13	Expectancy	Surprise	7	1
14	Physical pain	Worry	3	4
15	Bewilderment	Worry	7	7
16	Neutral	Neutral	13	0
17	Scorn	Disbelief	2	3
18	Excitement	Happy	7	10
19	Adoration	Neutral	7	1
20	Dread	Worry	2	4
21	Distrust	Annoyance	2	6
22	Complacency	Neutral	6	2
23	Worried laughter	Amused	6	11
24	Acute sorrow	Neutral	6	0
25	Horror	Fear	4	4
26	Neutral	Neutral	12	1
27	Neutral	Neutral	8	1
28	Pity	Depressed	2	1
29	Intrigued interest	Doubt	2	4
30	Dismay	Surprise	4	2
31	Suspicion	Disbelief	3	1
32	Awe	Surprise	9	6
33	Sullen anger	Thoughtful	2	1
34	Despair	Sadness	5	1
35	Disgust	Disgust	7	7
36	Stubbornness	Anger	3	4
37	Joy	Happy	11	10
38	Loathing	Disgust	6	6
39	Neutral	Neutral	15	0
40	Amazement	Surprise	15	6
41	Cynical bitterness	Dislike	2	7
42	Determination	Serious	2	2
43	Anxiety	Worry	3	3
44	Incredulous doubt	Worry	3	5

B.2.3 Results

Results appear in table B.1. Column 2 contains the words most often used to describe the stimuli of column 1. It should be stressed that the words in column 1 were simply given as cues to the actor who portrayed the implied emotions and no comparison between columns 1 and 2 is intended. The results of interest are those of column 2 and those of column 1 were only used in order that a set of stimuli could be prepared. Column 3 indicates the frequency of occurrence of the words presented in column 2. The maximum possible value is 21. Column 4 shows the number of times, words considered to be synonyms of those in column 2 were used. The choice concerning synonyms was made by the author and is of course completely arbitrary. Column 4 is presented merely as an illustration of the problems involved in interpreting subjects' responses.

B.2.4 Discussion

It will be seen from column 2 of table B.1 that the vocabulary of words used by subjects in response to 44 stimuli is quite restricted. 'Happy' is present 5 times, 'surprise' 5 times, 'worry' six times and 'neutral' 8 times. Osgood (1966) maintained that 'the language of emotions appeared to be very much richer than the territory it was designed to cover, and it was therefore full of quasi-synonyms'. This statement is true in the sense that many different words are used for the same stimulus, some being synonyms, others being unrelated. In the latter case, one can only assume that the facial expression has been interpreted differently. Returning to Osgood's statement, it should be pointed out that a rich vocabulary for the emotions does not imply an equally rich set of interpretations or recognition responses. Indeed, the fact that 24 of the 44 stimuli were most often described by one of 4 words seems to imply either a limitation in our recognition capabilities or a possible limitation in our vocabularies or even possibly a bad actor.

Examining the results of column 2, we can single out the word 'happy' for study. 'Happy' was the most often used word to describe 5 of the stimuli. To explain this, several interpretations are possible:

- (a) The actor was incapable of producing the subtle nuances which exist between 'silent laughter', 'glee', 'quiet pleasure', 'excitement' and 'joy' assuming that such nuances exist in the display of facial expressions.

- (b) Subjects were incapable of detecting the subtle nuances which did exist between the five stimuli.
- (c) Subjects were not equipped with an adequate vocabulary to describe the five different stimuli.
- (d) It may be that 'silent laughter', 'glee', 'quiet pleasure', 'excitement' and 'joy' naturally result in the same facial expression.

In the author's opinion, the results obtained in table B.1 can be explained as a combination of the four interpretations just described. Beginning with the actor, it is extremely difficult for him to convey related emotions simply through static facial expressions. Conveying an emotion is rarely done exclusively on the basis of an expression. The actor may try to momentarily create his own context with internal verbalisation and body positioning, but these will not be picked up in a stimulus consisting of only a facial expression.

Over and above this limitation, it may be that there are simply not as many facial expressions as there are emotions. There may be a finite number of distinguishable expressions which could be described as 'primary expressions'. Notice that this is quite different from the 'primary affects' of Tomkins and McCarter (1964). In their case, they have suggested the existence of eight primary emotional states whilst 'primary expressions' are facial configurations rather than emotional states. If one accepts the existence of primary expressions then it is possible to understand why the emotion language exceeds the range of possible expressions. Consider once again the 5 stimuli labelled as 'happy'. In each case a different although related emotional state was meant to be conveyed. However, it appears that the 5 resulting facial expressions were labelled by the same word. One might conclude from this that all 5 emotional states do in fact give rise to the same facial expression and that any difference between the states will only be detectable from contextual cues not present in the photographic stimuli, i.e. verbal content, body posture and movement and other cues. In other words, the five emotional states labelled as 'happy' give rise to the same primary expression and consequently, photographic presentation of emotional states will necessarily be insufficient for recognition of the differences which actually do exist between the five states.

In the author's opinion, the concept of 'primary expressions' offers the most important explanation of the results obtained in table B.1.

Although the actor and the subjects are limited in their own respective ways, it is suggested that the stimulus presentation imposes the greatest limitation on the task of recognising emotional states. If only primary expressions are available for recognition, subjects cannot be expected to differentiate between closely related emotional states. To recapitulate, it is suggested that only a limited number of facial expressions are capable of being portrayed. These are termed 'primary expressions' and each one is representative of several emotions. To detect the actual emotional differences, one must have more than static facial expression information. It becomes obvious that the use of movies represents a distinctive improvement in the method of presenting the stimuli.

Having considered the actual display of emotions, one must now examine the problem of recognition. If we accept that there are only a limited number of primary expressions, then we cannot expect subjects to recognise subtle variations in emotions if these are represented by the same primary expression unless we provide them with a complete context of the situation. This is certainly not the case with photographs or slides. Table B.1 indicates that certain stimuli resulted in high scores in column 3 and others in low scores. In the author's opinion, this is taken to be an indication of the degree to which a primary expression has become stereotyped. In other words, all expressions are primary but some are more stereotyped than others, i.e. those which resulted in the greatest response uniformity. 'Happy', 'worried' and 'surprise' are examples of strong stereotypes.

Stereotypes are artifacts of the graphic arts. In Ancient Greece, masks were used to indicate the happy and unhappy emotional states. Paintings, advertisements and caricatures make use of stereotyped expressions to convey mood when no other information is present. Sometimes a written caption or body cues will supplement the information present in the stereotyped facial expression. We recognise strong stereotypes immediately because of the learnt associations implicit in them. Often we need no other clues for recognition. The emotion of 'awe' may be represented by the primary expression of 'surprise' which is a strong stereotype. The subtlety between awe and surprise is lost but could be recognised in a fuller context.

B.2.5 Summary

A short experiment using photographs and word-labels was conducted in order to investigate some of the methodological problems which occur in tests dealing with the recognition of facial expressions. The basic problems which cropped up were similar to those described in the introduction and an attempt was made to interpret the results as they stood. It was suggested that only a limited number of 'primary expressions' could be portrayed and that each of these was representative of more than one emotional state. To recognise the differences between emotions resulting in the same primary expression, it was necessary to have available information other than the particular facial expression. Examples of primary expressions were 'happy', 'surprise' and 'worry'.

The concept of 'stereotype' was proposed to account for the different scores in column 3. It was suggested that a primary expression resulting in a high score was strongly stereotyped. Actually scores could be based on column 3 or a combination of columns 3 and 4 if synonyms were considered. Facial expressions presented in static form, i.e. photographs, were considered to be stereotyped in the sense that they were often sufficient to convey general emotional state without the presence of other information. However, it was recognised that expressions varied in the degree to which they were stereotyped, weaker stereotypes giving rise to a greater variety of unrelated responses.

B.3 The second experiment

B.3.1 Aim

The purpose of this second test is to investigate the relative usefulness of the upper and lower halves of the face in conveying emotion through expression. Masking techniques will be used for this purpose.

B.3.2 Introduction

Considerable controversy existed in the first half of this century as to whether the upper or lower half of the face was more capable of portraying emotions. Generally, the mouth and the eyes were seen as the main components involved in both halves and results either concluded in favour of the upper, lower or both halves (depending on the expression): Buzby (1924), Dunlap (1927), Coleman (1949), Frois-Wittmann (1930) and Hanawalt (1942, 1944). Because of the variety of results obtained, it was decided to carry out a separate experiment to check previous interpretations.

B.3.3 Method

The test was composed of three parts: one control test in which the entire face was presented, one test where the upper half was presented and finally one test where the lower half was presented. Masking was used and the dividing line between the upper and lower halves of the face passed through the nose tip. Faces were presented full-frontally in the usual slide format. Ten stimuli were chosen from experiment 1 and these appear in table B.2 (column 1). Strong and weak stereotypes are present.

Subjects were presented with a reference list containing the 10 word-labels of column 1. Sixty-six subjects were tested, 22 for each part of the experiment. No time limit was imposed and all subjects received the same order of presentation.

B.3.4 Results

Table B.2 lists the percentages of correct recognition for the ten stimuli in all three tests. A Kendall rank correlation test performed on the results obtained for the whole and upper-half cases was found to be significant at the 0.036 level. Correlation between the whole and lower-half cases was significant at the 0.019 level. No correlation was found between the upper and lower-half cases. Analysis of variance based on 2 variables, i.e. the 10 stimuli composing the 'stimulus' variable and the 3 methods of presentation composing the 'part' variable, indicated that the 'stimulus' variable was significant at the 0.01 level and the 'part' variable was significant at the 0.02 level (2 sided-test). However, a Duncan multiple range test set at the 0.01 level indicated that no significant difference existed between the upper and lower-half parts of the 'part' variable. As for the 'stimulus' variable, a Duncan multiple range test set at the 0.05 level indicated a significant difference between the following stimuli: between 'happiness' and all other stimuli, between 'surprise' and the first five stimuli and finally between 'neutral' and the first 3 stimuli.

Stimuli for which the upper half of the face resulted in higher scores than for the lower half were: 'sadness', 'fear', 'worry', 'surprise' and 'happiness'. Stimuli for which the reverse situation occurred, were 'determination', 'puzzlement', 'disgust' and 'neutral'. 'Anger' resulted in the same score for both the upper and lower half cases. For all stimuli,

Table B.2 Scores for the stimuli of experiment 2
presented in three viewing conditions

STIMULI	WHOLE PRESENTATION	UPPER-HALF PRESENTATION	LOWER-HALF PRESENTATION
Sadness	32%	23%	5%
Determination	45%	9%	41%
Fear	32%	27%	0%
Puzzlement	59%	27%	32%
Worry	36%	27%	18%
Disgust	77%	9%	45%
Anger	27%	18%	18%
Neutral	77%	50%	68%
Surprise	82%	64%	14%
Happiness	96%	100%	96%
AVERAGE	56.3%	35.3%	33.7%

except 'happiness', the whole face proved more useful than either half by itself.

It should be noted that the use of a reference list and a relatively small number of stimuli resulted in scores which were all above the chance level in the control test. In fact, certain stimuli scored very highly and once again these can be taken to be the stronger stereotypes.

The results of the analysis of variance and the correlation tests will now be discussed briefly.

B.3.5 Discussion

Analysis of variance indicates that certain stimuli were better recognised than others to a significant degree and that the whole face resulted in significantly higher scores than either half by itself. However, it will be noted that considerable variation occurred between each part of the test depending on the stimulus. It appears certain that some stimuli are better recognised with one half than with the other but in no apparently systematic fashion. As mentioned above, no correlation was found between the ranks for the upper and lower-half cases and it does not appear that those stimuli which had higher scores in the lower-half case have any common properties which make them stand out as a set which could be readily differentiated from the set arising from the upper-half case. Therefore, it must be concluded that certain emotions are better recognised with the upper-half of the face, others with the lower but in no systematic fashion, i.e. with no common properties identifying the stimuli in each case.

Before concluding, it should be noted that certain stimuli were recognised almost as well with one half of the face as with the whole. Usually, in these cases, the remaining half proved to be of very little use. Therefore, there is very little doubt that features in either half carry most of the information for certain expressions, i.e. 'sadness', 'fear' and 'surprise' were well recognised (in comparison with the control) with the upper half and 'determination' and 'disgust' were well recognised with the lower half. The point of the test was not to investigate this fact but rather to see whether either half was most useful for a particular 'category' of stimulus. The latter point was not verified.

B.4 General summary

The two experiments reported in this appendix were concerned with the recognition of facial expressions. A brief review of the research done in this field was presented in the introduction and this was followed by the first experiment. The purpose of this first test was to provide an example of the problems which occur in work of this type and for that reason, the test was designed in the traditional fashion which makes use of photographs and word-labels. Results led to the following general conclusions:

(a) It was suggested that only a limited number of expressions (not specified) existed to cover a rather larger number of emotions. These were labelled 'primary expressions'. It was felt that each primary expression served as the expressive component (from the face point of view) for several emotions and that consequently, related emotions resulting in the same primary expression required additional contextual information such as verbal content or body cues, in order that the particular emotion be recognised. In other words, primary expressions may be sufficient to identify a 'gross' emotional state but the 'actual' emotion to be recognised requires more information. That emotion, and possibly a few others will all be classified as variations of the gross emotional state which has resulted in a particular primary expression.

(b) It was suggested that primary expressions were stereotyped to differing degrees. The stronger stereotypes were recognised more often, in the sense that there was less variation in the different subject interpretations for a strong stereotype. Weaker stereotypes were subject to wide variations in interpretation. Consequently, a primary expression which was weakly stereotyped could be said to be representative of a gross emotional state which depended mostly on non-facial expression information for its recognition.

A second experiment was carried out in order to investigate the relative superiority of the upper and lower half of the face in expression recognition. 10 stimuli from the first test were used and three viewing conditions tested: the whole face, the upper-half and the lower-half. The following conclusions were reached:

- (a) Some expressions were better recognised with one half of the face than with the other. However, no apparent common properties identified the groups in either case. In all cases except one, a whole face resulted in higher scores than either half by itself.
- (b) Correlation was found between the score ranks for the whole face case and each of the half-face cases. However, none was found between the two half-face cases.
- (c) On the whole, certain stimuli were significantly better recognised than others. This was taken to be another measure of the degree to which certain expressions were stereotyped.

APPENDIX CA TEST CASE ON BIASING REFERENCE PHOTOGRAPHS

Shortly after the completion of his experiments, the author was contacted by the firm of Sampson & Co., a firm of solicitors engaged in the defence of a Mr. Ince, standing trial for the 'Braintree Barn restaurant' murder. Mr. Ince was identified by one of two witnesses in the following manner: his photograph was presented along with 11 others on a large sheet and witness 1 was asked if she recognised the murderer. She was present for a considerable period of time at the scene of the crime (about 30 minutes) and picked out the photograph of Mr. Ince when given the sheet with the 12 photographs.

The Defence contacted the author and asked him to examine whether a bias could be shown to exist in the case of Mr. Ince's photograph. The sheet used by the witness was given to a group of 54 subjects who were asked the following question: "If you had to choose one of these 12 photographs as being particularly biased from the remainder because of some characteristic(s) which you are free to choose, what would be your first choice? The criteria which you use in making your choice are entirely up to you and cannot be specified by the experimenter."

28 of the 54 subjects chose Mr. Ince's photograph as being the most biased. None of them knew what he looked like beforehand. Neither did the author until after the experiment. The remaining 11 photographs were chosen around or below chance level. The bias of Mr. Ince's photograph was significant at the 0.001 level (Duncan Multiple Range Test) with no other photographs being significant.

Subjects volunteered reasons for their choice of photograph no. 3 (Mr. Ince). The following were given:

- (a) He looked more educated than the rest,
- (b) He was better groomed, and
- (c) There was a "department of employment" stamp at the base of the photograph (no stamps appeared on any of the remaining photographs).

The use of a sheet with 12 photographs is apparently standard procedure in preliminary suspect identification situations. For that

reason, it is important to ascertain beforehand whether any strong biases exist for certain members of the sample. It is probably impossible to come up with a completely unbiased set of photographs in any situation, but certainly there should be an effort to reduce bias where possible before a witness produces evidence from a chosen sample. In the particular case just described, the bias seems to be very strong.

The witness who identified Mr. Ince was very much involved in the case. In fact she and her father were shot whilst her mother actually died as a result of injuries. The father was unable to pick out Mr. Ince but the daughter was and the case rests largely on her identification evidence.

REFERENCES

- Ambrose, J.A. The development of the smiling response in early infancy.
1961 In B.M. Foss (Ed.), *Determinants of Infant Behaviour*.
New York : Wiley, pp. 179-196.
- Ambrose, J.A. The concept of a critical period for the development of
1963 social responsiveness in early human infancy. In B.M.
Foss (Ed.), *Determinants of Infant Behaviour*. New York :
Wiley, pp. 201-225.
- Attneave, F. Some informational aspects of visual perception.
1954 Psychol. Rev., 61, pp. 183-193.
- Baker, E. Perceiver-Variables involved in the recognition of faces.
1967 Ph.D. Thesis, University College, University of London.
- Bartlett, F.C. Remembering: a study in experimental and social
1967 psychology. London : Cambridge University Press.
- Bradshaw, J.L. & Wallace, G. Models for the processing and identificat-
1971 ion of faces. Percept. Psychophys., 9, pp. 443-448.
- Brooks, R.M. & Goldstein, A.G. Recognition by children of inverted
1963 photographs of faces. Child Dev., 34, pp. 1033-1040.
- Brunswik, E. & Reiter, L. Perception and the representative design of
1938 psychological experiments. Z. Psychol., 142, pp. 67-134.
- Buzby, D.E. The interpretation of facial expression. Am. J. Psychol.,
1924 35, pp. 602-604.
- Campbell, F.W. & Robson, J.G. A fresh approach to stabilized retinal
1961 images. J. Physiol., Lond., 158, pp. 1-11.
- Chance, J., Goldstein, G. & Schicht, W. Effects of acquaintance and
1967 friendship on children's recognition of classmates'
faces. Psychon. Sci., 7, pp. 223-224.
- Cochran, W.G. Testing a linear relation among variances. Biometrics,
1951 7, pp. 17-32.

- Coleman, J.C. Facial expressions of emotion. Psychol. Monogr., 63,
1949 no. 296, pp. 1-36.
- Connolly, K. & McKellar, P. Forensic Psychology. Bull. Br. psychol.
1963 Soc., 16, pp. 16-24.
- Cuceloglu, D.M. Perception of facial expressions in three different
1970 cultures. In A.T. Welford & L. Houssiadas (Eds.),
Contemporary problems in Perception. London : Taylor
& Francis, pp. 93-100.
- Darwin, C. Expression of the Emotions in Man and Animals. London,
1873 Murray.
- Dember, W.N. The Psychology of Perception. New York : Holt, Rinehart
1963 and Winston.
- Deutsch, J.A. A theory of Shape Recognition. Br. J. Psychol., 46,
1955 pp. 30-37.
- Deutsch, J.A. The plexiform zone and shape recognition in the octopus,
1960 Nature, 185, pp. 443-446.
- Dodwell, P.C. A coupled System for coding and learning in shape dis-
1964 crimination. Psychol. Rev., 71, pp. 148-159.
- Dodwell, P.C. Perceptual processing: stimulus equivalence and pattern
1971 recognition. New York : Appleton, Century & Crofts.
- Dunlap, K. The role of eye-muscles and mouth-muscles in the express-
1927 ion of the emotions. Gen. Psychol. Monogr., 2, no. 3,
pp. 195-235.
- Eagle, M.N. & Klein, G.S. Fragmentation phenomena with the use of
1962 stabilised retinal images. Percept. mot. Skills, 15,
pp. 579-582.
- Edwards, A.L. Experimental Design in Psychological Research; (third
1968 edition), New York; Holt Rinehart & Winston.
- Ekman, P. Differential Communication of Affect by Head and Body
1965 cues. J. Person. Soc. Psychol., 2, pp. 726-735.

- Ekman, P. & Friesen, W.V. Head and Body cues in the judgment of emotion:
1967 a reformulation. Percept. mot. Skills, 24, pp. 711-724.
- Evans, C.R. Some studies of pattern recognition using a stabilised
1965 retinal image. Br. J. Psychol., 56, pp. 121-133.
- Evans, C.R. Further studies of pattern perception and a stabilised
1967 retinal image; the use of prolonged after-images to achieve perfect stabilisation. Br. J. Psychol., 58, pp. 315-327.
- Fantz, R.L. Pattern vision in newborn infants. Science, 140,
1963 pp. 296-297.
- Fantz, R.L. Pattern discrimination and selective attention as determinants of perceptual development from birth. In A.H. Kidd & J.L. Rivoire (Eds.), Perceptual Development in Children. New York : Int. University Press.
- Fantz, R.L. Visual perception and experience in early infancy: A
1967 look at the hidden side of behaviour development. In H.W. Stevenson, E.H. Hess & H.L. Rheingold (Eds.), Early Behaviour: Comparative and developmental approaches. New York : Wiley, pp. 181-224.
- Fantz, R.L., Ord, J.M. & Udelf, M.S. Maturation of pattern vision in
1962 infants during the first six months. J. comp. physiol. Psychol., 55, pp. 907-917.
- Fisher, G.H. & Cox, L. Development and Initial application of the facial
1971 attribute differential. Personal Communication.
- Fraisse, P. & Blancheteau, M. The influence of the number of alternatives
1962 on the perceptual recognition threshold. Q. Jl.exp. Psychol., 14, pp. 52-56.
- Friedman, M.P., Reed, S.K. & Carterette, E.C. Feature saliency and
1971 recognition memory for schematic faces. Percept. Psychophys., 10(1), pp. 47-50.

- Frijda, N.H. The understanding of facial expression of emotion.
1953 Acta. Psychol., 9, pp. 294-362.
- Frois-Wittmann, J. The judgment of facial expressions. J. exp.
1930 Psychol., 13, pp. 113-151.
- Galper, R.E. Recognition of faces: the role of expression. University
1970 Microfilms abstract no. 70-16, 061.
- Galper, R.E. & Hochberg, J. Recognition memory for photographs of faces.
1971 Am. J. Psychol., 84, pp. 351-354.
- Garneau, M. Facial Description Ability. Police Res. Bull., 19,
1972 pp. 44-45.
- Gibson, E.J. Principles of Perceptual Learning and Development, New
1966 York : Appleton, Century & Crofts.
- Goldstein, A.G. & Chance, J.E. Visual recognition memory for complex
1970 configurations. Percept. Psychophys., 9(2B), pp. 237-241.
- Goldstein, A.G. & Mackenberg, E.J. Recognition of human faces from
1966 isolated facial features: a developmental study.
Psychon. Sci., 6(4), pp. 149-150.
- Goldstein, A.J., Harmon, L.D. & Lesk, A.B. Identification of human faces.
1971 Proc. IEEE., 59(5), pp. 748-750.
- Goldstein, A.J., Harmon, L.D. & Lesk, A.B. Man-machine interaction in
1972 human-face identification. Bell. System Tech. J., 51(2),
pp. 399-427.;
- Grimsdale, R.L., Sumner, F.H., Tunis, C.J. & Kilburn, T. A system for
1959 the automatic recognition of patterns. Proc. IEEE., 106,
pp. 210-221.
- Hake, H.W. Contributions of Psychology to the study of pattern
1957 vision. WADC. Tech. Rep. no. 57-621, pp. 79-92.
- Hanawalt, N.G. The role of the upper and lower parts of the face as a
1942 basis for judging facial expressions (1), J. gen.
Psychol., 27, pp. 331-346.

- Hanawalt, N.G. The role of the upper and lower parts of the face as a
1944 basis for judging facial expressions (2), J. gen. Psychol.,
 31, pp. 23-36.
- Harmon, L.D. Some aspects of recognition of human faces. In O.J.
1971 Grusser & R. Klinke (Eds.), Pattern recognition in
 Biological and Technical systems, Berlin : Springer
 Verlag. pp. 196-217.
- Hecren, H. & Angelerques, R. Agnosia for faces. Arch. Neurol., 7,
1962 pp. 92-100.
- Hochberg, J. & Galper, R.E. Recognition of faces: 1. An exploratory
1967 study. Psychon. Sci., 9, pp. 619-620.
- Home Office Report The use of Photofit: a progress report. Personal
1971 Communication..
- Howells, T.H. A study of ability to recognise faces. J. abnorm. soc.
1938 Psychol., 33, pp. 124-127.
- Hubel, D.H. Integrative processes in central visual pathways of the
1963 cat. J. opt. Soc. Am., 53(1), pp. 58-65.
- Hubel, D.H. & Wiesel, T.N. Receptive fields, binocular interaction and
1962 functional architecture in the cat's visual cortex.
 J. Physiol., 160, pp. 106-154.
- Hubel, D.H. & Wiesel, T.N. Receptive fields and functional architecture
1965 in two nonstriate visual areas (18 and 19) of the cat.
 J. Neurophysiol., 28, pp. 229-289.
- Hubel, D.H. & Wiesel, T.N. Receptive fields and functional architecture
1968 of monkey striate cortex. J. Physiol., 195, pp. 215-245.
- Kabrisky, M., Tallman, O., Day, C.M. & Radoy, C.M. A theory of pattern
1970 perception based on human physiology. In A.T. Welford &
 L. Houssiadass (Eds.), Contemporary problems in Perception,
 London : Taylor and Francis, pp. 129-148.

- Kagan, J., Henker, B.A., Hen-Tov, A., Levine, J. & Lewis, M. Infant's
1966 differential reactions to familiar and distorted faces,
Child Dev., 37, pp. 519-532.
- Kaufman, L. & Richards, W. Spontaneous fixation tendencies for visual
1969 forms. Percept. Psychophys., 5(2), pp. 85-88.
- Kuffler, S.W. Discharge patterns and functional organisation of
1953 mammalian retina. J. Neurophysiol., 16, pp. 37-68.
- Lashley, K.S. The Mechanism of Vision: XV. Preliminary studies of
1938 rat's capacity for detail vision. J. gen. Psychol., 18,
pp. 123-193.
- Lashley, K.S. The problem of cerebral organisation in vision. Visual
1942 Mechanisms: Biological Symposia, 7, Cattell Press,
pp. 301-321.
- Laughery, K.R., Alexander, J.F. & Lane, A.B. Recognition of human faces:
1971 effects of target exposure time, target position and type
of photograph. J. appl. Psychol., 55, pp. 477-483.
- Lettvin, J.Y., Matwrana, H.R., McCulloch, W.S. & Pitts, W.H. What the
1959 frog's eye tells the frog's brain. Proc. IRE., 47,
pp. 1940-1951.
- Lindzey, G., Prince, B. & Wright, H.K. A study of facial asymmetry.
1952 J. Personality, 21, pp. 68-84.
- Mackay, D.M. Visual noise as a tool for research. J. gen. Psychol.,
1965 72, pp. 181-197.
- Malpass, R.S. & Kravitz, J. Recognition of faces of own and other race.
1969 J. pers. soc. Psychol., 13, pp. 330-334.
- Marsh, A.H. Perception of visual forms in intensely noisy television
1967 images. Ph.D. thesis, Imperial College, University of
London.
- McCurdy, H.G. Experimental notes on the asymmetry of the human face.
1949 J. abnorm. soc. Psychol., 44, pp. 553-555.

- McFarland, J.H. Sequential part presentation: a method of studying
1965 visual form perception. Br. J. Psychol., 56, pp. 439-446.
- McFarland, J.H. Parts of perceived visual forms: new evidence. Percept.
1968 Psychophys., 3, pp. 118-120.
- Messick, S. & Damarin, F. Cognitive styles and memory for faces.
1964 J. abnorm. soc. Psychol., 69, pp. 313-318.
- Morgan, G.A. & Ricciuti, H.W. Infant's responses to strangers during
1969 the first year. In B.M. Foss (Ed.), Determinants of
infant behaviour IV, London; Methuen.
- Noton, D. & Stark, L. Eye movements and visual perception. Scient. Am.,
1971 224, pp. 34-43.
- Osgood, C.E. Dimensionality of the semantic space for communication
1966 via facial expressions. Scand. J. Psychol., 7, pp. 1-30.
- Pasnak, R. Pattern complexity and response to distinctive features.
1971 Am. J. Psychol., 84, pp. 235-245.
- Pearson, D. Fidelity criteria for visual image transmission over
1965 noisy communication channels. Ph.D. Thesis, Imperial
College, University of London.
- Penry, J. The face of Man. London : Hutchinson.
1952
- Penry, J. Looking at faces and remembering them: a guide to facial
1971 identification. London : Elek.
- Phillips, R.J. Why are faces hard to recognise in photonegative?
1972 Personal Communication. Percept. Psychophys., 12(5),
p. 425.
- Pritchard, R.M., Heron, W. & Hebb, D.O. Visual perception approached by
1960 the method of stabilised images. Can. J. Psychol., 14,
pp. 67-77.

- Ratliff, F. Inhibitory Interaction and the detection and enhancement
1961 of contours. In W. Rosenblith (Ed.), *Sensory Communication*, Cambridge : M.I.T. Press.
- Richards, W. & Kaufman, L. Center of gravity tendencies for fixations
1969 and flow patterns. Percept. Psychophys., 5(2), pp. 81-84.
- Rosenthal, R. On the social psychology of the psychological experiment.
1963 Am. Scientist, 51(2), pp. 268-280.
- Scapinello, K.F. & Yarmey, A.D. The role of familiarity and orientation
1970 in immediate and delayed recognition of pictorial stimuli.
 Psychon. Sci., 21, pp. 329-331.
- Schlosberg, H. The description of facial expressions in terms of two
1952 dimensions. J. exp. Psychol., 44(4), pp. 229-237.
- Schlosberg, H. Three dimensions of emotions, Psychol. Rev., 61(2), p. 81.
1954
- Schultz, E. F. Rules of thumb for determining expectations of mean
1955 squares in analysis of variance. Biometrics, 2,
 pp. 123-135.
- Siegel, S. Nonparametric statistics. London : McGraw Hill.
1956
- Snedecor, G.W. & Cochran, W.G. Statistical Methods. (Sixth edition)
1967 Iowa State Press, pp. 364-369.
- Stapley, B. Visual enhancement of telephone conversations. Ph.D.
1972 Thesis, Imperial College, University of London.
- Stechler, G., Bradford, S. & Levy, L. Attention in the newborn, Science,
1966 151, pp. 1246-1248.
- Stent, G.S. Cellular Communication. Scient. Am., 227(3), pp. 42-52.
1972
- Sutherland, N.S. Visual discrimination of orientation and shape by the
1957 octopus. Nature, 179, pp. 11-13.

- Thomson, J. Development of facial expressions of emotions in blind
1941 and seeing children. Archs Psychol., 264.
- Tomkins, S.S. & McCarter, R. What and where are the primary affects?
1964 Some evidence for a theory. Percept. mot. Skills, 18,
pp. 119-158.
- Tversky, A. & Krantz, D.H. Similarity of schematic faces: a test of
1969 interdimensional additivity. Percept. Psychophys., 5,
pp. 124-127.
- Uttley, A.M. The classification of signals in the nervous system.
1954 EEG and Clinical Neurophysiology, 6, pp. 479-494.
- Van Doorn, A.C. The Plumbicon compared with other television camera
1966 tubes. Philips Tech. Rev., 27, pp. 1-14.
- Vine, I. Communication by facial-visual signals. In J.H. Crook
1970 (Ed.), Social behaviour of Birds and Mammals, Academic
Press, pp. 296-336.
- Vitz, P.C. & Todd, T. C. A model of the perception of simple geometric
1971 figures. Psychol. Rev., 78(3), pp. 207-228.
- Wallace, G., Coltheart, M. & Forster, K.I. Reminiscence in recognition
1970 memory for faces. Psychon. Sci., 18, pp. 335-336.
- Weisstein, N. What the frog's eye tells the human brain: Single cell
1969 analysers in the human visual system. Psychol. Bull.,
72, pp. 157-176.
- Werblin, F.S. The control of sensitivity in the retina. Scient. Am.,
1973 228(1), pp. 70-70.
- Witryol, S.L. & Kaess, W.A. Sex differences in social memory tasks.
1957 J. abnorm. soc. Psychol., 54, pp. 343-346.
- Yarmey, A.D. Recognition memory for familiar "public" faces: effects
1971 of orientation and delay. Psychon. Sci., 24, pp. 286-287.

- Yin, R.K. Looking at upside down faces. J. exp. Psychol., 81,
1969 pp. 141-145.
- Yin, R.K. Face recognition by brain-injured patients: a dissociable
1970 ability? Neuropsychologia, 8, pp. 395-402.

ADDITIONAL REFERENCES

- Lewis, N.W. & Allnatt, J.W. Subjective quality of television pictures
1965 with multiple impairments, Electronics Letters, 1,
 pp. 187-188.
- Prosser, R.D., Allnatt, J.W. & Lewis, N.W. Quality grading in impaired
1964 television pictures, Proc. IEE, 3, pp. 491-502.