AN INVESTIGATION INTO THE PERCEPTION OF GENDER IN FACES

Elizabeth Brown

A Thesis Submitted to the University of St Andrews for the Degree of Master of Science

October 1992

Department of Psychology, University of St Andrews, St Andrews, Fife,

KY16 9JU, Scotland

I, Elizabeth Brown, hereby certify that this thesis has been composed by myself, that it is a record of my own work, and that it has not been accepted in partial or complete fulfilment of any other degree or professional qualification.

Signed Date 3.2.93

In submitting this thesis to the University of St Andrews I understand that I am giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. I also understand that the title and abstract will be published, and that a copy of the work may be made and supplied to any bona fide library or research worker.

Acknowledgements.

I wish to thank Dr D. I. Perrett and Dr T. Jordan of the Psychology Department of St Andrews University for their advice and guidance during the course of this work. I also thank P. J. Benson for programming assistance and M. Oram for advice on statistics.

CONTENTS

Abstract

Experiments demonstrating categorical perception of phonemes have shown increased sensitivity to a change of stimulus across a category boundary (eg from /p/ to /b/) compared with a change within a category (eg two versions of /p/). Examination of the nature of the change in the acoustic properties of stimuli at the point of increased sensitivity has indicated how phonemic discriminations are made. In the present study an investigation was made to determine whether perception of gender in human faces is similarly categorical and whether the results of the investigation could be used to help determine how gender is perceived. A continuum of synthesised faces changing gradually from a prototype male to a prototype female in four equal steps was used in identification and discrimination tasks. A discrete categorical boundary was demonstrated in the identification task but discrimination of stimuli across that boundary was not more accurate than discrimination between two faces which had been identified as being of the same gender. A second experiment examined the attribution of gender to the isolated features of the two prototype faces and to composites of the faces when their features had been interchanged. The results showed that all the features except the nose carried some information about gender when they were viewed in isolation. They also showed that when brows, chin, brows & eyes together or the jaw were interchanged they effected significant change in the perceived gender of the recipient face. Another attempt was then made to demonstrate categorical perception of gender in faces using continua of faces in which all facial features remained neutral with regard to gender except the jaw and the brows. A more gradual change in perceived gender along the continuum instead of a discrete boundary between the two genders was now demonstrated in the identification task and discrimination between all pairs of faces became more difficult. It is concluded that perception of gender is not categorical but it is suggested that the same identification and discrimination paradigms might still be used to investigate further the perception of gender and identity in faces.

Chapter One.

The purpose of this study is to try to establish whether face perception might be investigated using similar methods to those already used in investigation of speech perception. Speech is not perceived by a process of simple transformation from fixed acoustic patterns which represent distinct units of sound making up the various words. The smallest unit of sound in any language which can distinguish one word from another is the phoneme and any one phoneme can be represented by very varied acoustic properties depending on its context within a word, the speaker's voice and dialect, and the speed at which s/he is talking. Lisker (1978) enumerated sixteen different cues which can be used to differentiate between /b/ and /p/ as they occur in the two words /rabid/ and /rapid/. In spite of this variation in the acoustic signal normal hearing listeners can easily understand rapid rates of speech even when the sound is distorted in some way; for example, in a noisy environment, over the telephone or if the speaker is whispering.

Categorical perception of phonemes.

It is well established that phonemes are perceived categorically; that is, discriminations are readily made between sounds when they represent two different phonemes such as /p/ and /b/ but are not made easily when they represent two different versions of /p/ from the same speaker. The listener seems to be noticing those differences which denote a change from one phoneme to another and ignoring those which do not, and the boundary between the two representations is quite sharp.

Categorical perception of phonemes was first demonstrated by Liberman, Harris, Hoffman and Griffith (1957). They found that it is the formant frequency transitions (the changes in frequency at the onset of the voiced part of the signal) which are the important auditory cues for the differentiation of the phonemes /b/, /d/ and /g/. They created synthetic analogues of natural speech syllables which varied in the formant frequency transitions creating a continuum of sounds ranging from /b/ to /g/. These sounds were then presented to listeners in two types of task, (i) an identification task in which they were required to say whether a syllable began with /b/, /d/ or /g/, and (ii) a discrimination task in which they had to say which two of three stimuli were the same. The results showed sharp distinctions between the categories /b/ and /d/, and /d/ and /g/ in the identification task. The discrimination task showed discrimination at the level of chance between syllables which fell into the same category, but much better discrimination between syllables which fell into different categories. This peak of discrimination corresponded to the point on the continuum where one phoneme had been identified 50% of the time; that is, the category boundary. There is a variety of opinion about what constitutes categorical perception (Harnad, 1987), but it is the foregoing definition of the phenomenon which is of interest in the present investigation.

Lisker & Abramson (1967) demonstrated that the category boundary between the phonemes /b/ and /p/ was determined by the voice-onset time (VOT), the time between the initial release of air and the beginning of the sound caused by vibration of the vocal cords. The boundary was found to be at about 30 ms. Eimas, Siqueland, Jusczyk & Vigorito (1971) found that infants aged 1-4 months noticed a difference between two syllables, one of which had a VOT under 30 ms and the other a VOT over 30 ms. They did not notice a difference between two syllables when both had VOTs over 30 ms even though the difference between the two VOTs in every case had been 20 ms. At that time no task had been devised for testing identification in infants.

Eimas et al claimed that their results showed that the ability to categorise sounds so as to discriminate between /b/ and /p/ was innate, but several studies have shown differences between listeners whose native languages use differing phonemic distinctions. Abramson & Lisker (1970) and Lisker & Abramson (1970) showed that Thai speakers perceived categorically a prevoiced distinction which was not perceived by American speakers whose English language does not use that distinction. Subsequently Miyawaki, Strange, Verbrugge, Liberman, Jenkins & Fujimura (1975) investigated the difference between American and Japanese perception of the /la/ /ra/ distinction which is cued by the third formant transition. American subjects perceived the continuum categorically, discriminating between stimuli at the point where /ra/ changed to /la/, but Japanese, whose language does not use the phoneme /r/, perceived differences between pairs of stimuli continuously; that is, all pairs were discriminated at a level only slightly better than chance. When the same third formant components of the syllables were heard separately, (ie as non-speech controls) both Japanese and American subjects were highly accurate in discrimination of all the comparison pairs. These findings suggest that, if categorical perception of phonemic distinctions is innate as Eimas claims, the ability to discriminate categorically between some of the distinctions can also be lost if it is not exercised. Later work by Werker & Tees (1984) showed infants losing the ability to make phonemic discriminations which are not used in their own language by the age of ten months and retaining and developing the ability to make those which are necessary in their native language. Burnham, Earnshaw & Clark (1991), using an infant speech identification procedure to test development of categorisation, showed that perception of a native contrast becomes more categorical with age, while perception of the non-native contrast becomes less categorical.

In contrast to the above findings, however, Pisoni, Aslin, Perey & Hennessy (1982) demonstrated a category boundary and a corresponding, though small, peak of discrimination between stimuli representing a non-native phonemic contrast for adult, monolingual English speakers. The same subjects showed a sharper category boundary with a corresponding, larger peak of discrimination between phonemic contrasts which are critical in English. This suggests that the ability to perceive non-native contrasts is not lost but only reduced through lack of practice.

Eimas et al (1971) also claimed that their findings indicated the possession of a special capacity for phonemic perception in humans. However, Kuhl & Miller (1975) found that chinchillas showed a similar categorical boundary at 30ms VOT in a task designed to test their identification of synthesised speech syllables which differed along a continuum in which the consonant /d/ changed to /t/. In a later study (Kuhl & Miller 1978) they found that the animals showed a peak of discrimination between stimuli of varying VOT which corresponded to the category boundary established in the earlier experiment. Kuhl & Padden (1982) demonstrated a similar phoneme boundary effect in macaque monkeys. Dooling, Okanoya, & Brown (1989) showed categorical perception by budgerigars of the VOT distinctions between stimuli on continua of synthetic speech sounds from /ba/ to /pa/, /da/ to /ta/ and /ga/ to /ka/. The peaks of discrimination occurred at different steps along the continua for the birds than for humans and other mammals. Dooling & Brown (1990) also showed budgerigars discriminating categorically between spoken vowels where the difference is cued by formant frequencies.

Categorical perception of non-speech sounds and visual stimuli.

Other studies showed that categorical perception occurred for non-speech sounds. Miller, Wier, Pastore, Kelly & Dooling (1976) showed categorical perception of stimuli along a noise/buzz continuum. More recently Pastore, Li & Layer (1990) have shown a categorical effect for non-speech chirps and bleats which was similar to that obtained when sounds with the same acoustic properties were heard in synthesised speech. Locke & Kellar (1973), using musical stimuli, showed a change in key perceived categorically by musicians but not by non-musicians and Burns & Ward (1978) showed a similar difference in perception of melodic intervals by musicians and non-musicians. Those two studies give evidence that categorical perception may develop as a result of experience and learning.

In the visual modality Pastore, Ahroon, Baffuto, Friedman, Puleo & Fink (1977) demonstrated categorical perception of flashing light stimuli whose rate of flashing varied along a continuum centred around the flicker-fusion threshold. Etcoff & Magee (1992) have obtained categorical effects in identification and discrimination tasks using line drawings of faces whose expression changed from one to another (eg happy - sad) along a continuum. From the illustrations of their stimuli which have been reduced, it is not possible to see what the differences between the drawings used in the experiment actually were. An earlier experiment by Brown (1990) using line drawings of animal heads which changed along a continuum from a cat to a dog was confounded by the fact that subjects discriminated between drawings on the basis of details in the lines. For example, in one drawing two lines touched and in its neighbour on the continuum they did not. It was not, therefore, possible to claim that the discrimination had been made on the basis of the drawing's identity as a cat or a dog. That could also be true for the results of the Etcoff & Magee experiment. A further attempt (Brown 1990) to demonstrate categorical perception in the visual modality used a continuum of silhouettes changing from a cross to a sword. The boundary between the two categories was not very sharp but the discrimination task revealed a peak of discrimination between the two stimuli which differed in the curvature or angularity of the shaft of the silhouette. Pairs of silhouettes in which both shafts were smoothly curved and pairs in which both were slightly angular were discriminated from each other at the level of chance.

The pair of silhouettes in which one shaft was smoothly curved and the other slightly angular was discriminated much more accurately. The peak of discrimination did not correspond to the category boundary demonstrated in the identification task; that is, the categorisation was not made on the basis of the curvature/angularity of the shaft, but on some other basis.

Interpretation of categorical perception of speech.

The phenomenon of categorical perception was considered at one time to indicate a special process for decoding speech (Liberman 1970), and there are several interpretations of categorical perception of speech sounds. One explanation championed by Liberman et al (1967) and known as the motor theory of speech perception claimed that speech sounds are categorised according to the way they are articulated and that the mechanisms for perception and production of speech are interlinked. That theory is not supported by the fact that non-human species such as rodents and monkeys perceive speech sounds categorically although they cannot produce speech. Dooling & Brown (1990) point out that budgerigars can learn to imitate speech and therefore must have the capacity to perceive it as well as to produce sounds which are similar to human speech. The converse does not follow that because they can perceive speech in a categorical manner they have an interlinked mechanism for producing it.

Another explanation advocated by Pastore (1976) and Stevens (1981) is that there are innate natural regions of sensitivity in the auditory system where discrimination is good and others where it is poor, thus enabling sounds to be heard categorically. That theory is supported by the fact that some non-speech sounds are perceived categorically and that animals have been found to hear some speech sounds in the same categorical manner as humans with a category boundary usually occurring at a similar point on the continuum of sounds.

A third explanation suggested by Lane (1965) is that the categories are learned with experience and are a way of simplifying perception by grouping a set of sounds which carry the same meaning under one perceptual heading. That theory is difficult to reconcile with the evidence that infants perceive phonemes categorically as early as one month of age (Eimas et al, 1971), but lose the ability to distinguish phonetic categories not used in their own language by the age of ten months (Werker & Tees, 1984). If they were learning to categorise the sounds they should learn the discriminations which are relevant to the language they hear spoken by their compatriots rather than lose the ones which are not relevant. However the results of Burnham et al (1991) which demonstrate that perception of a native contrast becomes more categorical with age, while perception of a non-native contrast becomes less categorical, show that there is an element of learning involved in the development of categorical perception of speech sounds. The fact that musicians perceive some sound differences which are significant in music categorically whereas non-musicians perceive them continuously (Locke & Kellar(1973); Burns & Ward (1978)) shows that categorical perception of some types of stimuli may be developed as a consequence of interest and experience. It does not mean that non-musicians do not possess the necessary sensitivity but may simply mean that they have not chosen to exercise it. It is possible that Pastore's and Steven's theory and Lane's are both correct. There could be innate regions of sensitivity in the perceptual systems which are given more or less attention according to need and experience and learning.

<u>Reasons for investigating face perception by the methods used to investigate speech perception.</u>

The phenomenon of categorical perception has been shown to occur in perception of speech sounds by both animals and humans and in perception of some non-speech sounds as well as some visual stimuli. This suggests that it may be a type of organisation which is common to several perceptual processes. Investigation of categorical perception of speech sounds has been useful in identifying which auditory cues are used to discriminate one phoneme from another. It was reasoned therefore that, if visual perception was also sometimes categorical, it might be possible to use similar methods to illuminate how different types of visual cues are used to interpret the meaning of visual stimuli. Because several visual stimuli from a continuum can be viewed side by side for as long as is required it should be easy to ascertain the nature of the differences between them at the point on a continuum at which discrimination was best. That might also throw further light on the processes involved in categorical perception itself.

If a similar paradigm to that used in studies of speech perception were to be used to investigate visual perception it would be necessary to find categories of visual stimuli which were similar enough but different enough to be represented along a continuum which changed from one category to another in a manageable series of small steps. Subjects would also need to be familiar with the type of stimuli used, and to be as expert in processing them as they are with phonemes. Faces are similar to each other in many respects and yet they come in very great variety, no two being exactly the same. Synthesised faces such as photographs or drawings could be used as experimental stimuli and changed in a controlled way along a continuum. Faces also have many visual properties which could be manipulated for experimental purposes. All normal adult human beings are experts in face perception having perceived and discriminated between many hundreds of faces during the course of their lives. Most of them, however, are only expert in identification of faces of their own race (Galper, 1973; Chance, Goldstein & McBride, 1975; Brigham & Bankowitz, 1978) although this varies depending on experience. These factors all have analogies with speech sounds which are made by an enormous variety of voices, have many auditory properties, but also have similar meanings no matter who utters them. Speech can be synthesised electronically and changed in a controlled way for the purposes of experimentation. All normal adult listeners are experts in perception of the sounds of their own language. Faces would therefore seem to be ideally suitable stimuli to use for the investigation of similarities and differences between auditory and visual perceptual processes using the paradigms employed in investigation of speech perception.

Gender as a facial category for investigation.

Since 'male' and 'female' are categories of face with which everyone is familiar it seemed reasonable to suppose that the two genders might be perceived categorically. If a continuum of faces could be created which varied in small stages from one which was definitely female to one which was definitely male there might be a point on the continuum (the category boundary) at which a sharp distinction occurred between those faces classified by subjects as female and those classified as male. If perception of gender is categorical a second prediction could be made. That is that subjects should discriminate more accurately between two faces which straddled that category boundary than between two which were both classified either as female or as male, even though the degree of physical difference between each pair of faces was the same. If that could be established it might then be possible to manipulate various aspects of the two faces to find out which properties of a face give it its gender and whether some properties have more gender value than others.

To satisfy the criterion that two faces should be used, one of which was definitely male and the other definitely female, it was decided to use averaged prototype faces. Using the logic of Galton (1878, 1883) it was reasoned that, if there were properties of a face which were perceived as male or as female, those properties should be embodied in one face which was an average of several male faces and in another face which was an average of several female faces. The averaged male face could then be changed gradually to become the averaged female face.

A first experiment was devised which took advantage of the fact that the facilities to create the averaged prototype male and female faces and the necessary continuum of faces had already been developed (Benson & Perrett, 1992; Benson & Perrett, 1991). A computer program which transforms an image of one person, in measured stages, into an image of another was used to transform a prototype male face into a prototype female one in various numbers of steps. The images were photographic which meant that the problem previously encountered with line drawings of animal heads (Brown, 1990) could be avoided. There would not be sharply noticeable differences between details in the lines of a drawing but rather a smoothly changing, soft, grey-scale, photographic image.

Experiment 1a.

Method.

Stimuli.

The four stimuli used in the experiment are shown in Figure I and were created in the following way. Photographs were obtained of the faces of sixteen male undergraduates and sixteen female undergraduates from Aberdeen University. Each student adopted a neutral expression, wore no eyeglasses or adornments and the males were clean-shaven. The photographs were blended into a prototypical male face and a prototypical female face respectively, using the technique for creating composite faces described by Benson & Perrett (1992). Each facial image was first frame-grabbed on a video camera, and digitized to 512 square pixels at 256 grey levels per pixel,

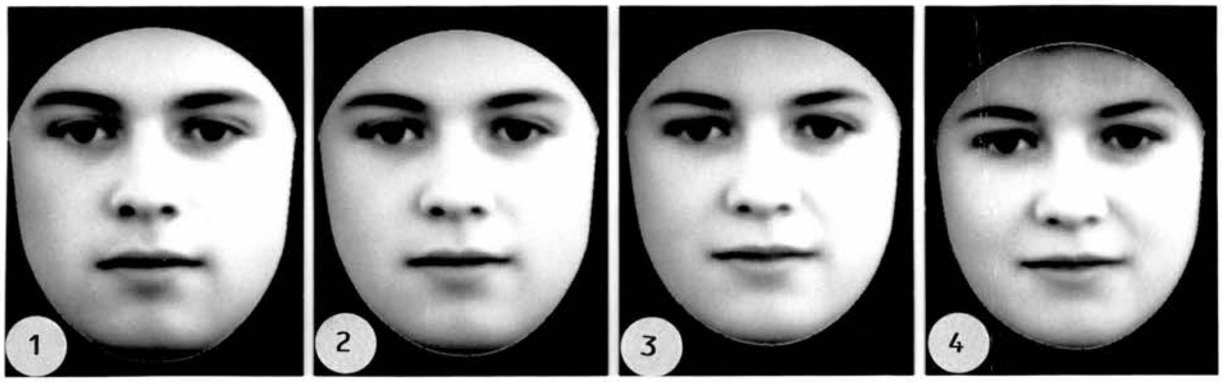


Figure I. The four stimuli used in Experiment 1.
Face 1 is the male prototype and face 4 the female prototype.

inter-pupillary distance for each face being standardised. Images were then transferred to a Silicon Graphics IRIS 3130 graphics computer for processing. For each image, the x and y co-ordinates of 224 feature points were defined by hand using the computer's mouse, and then the average co-ordinates of feature positions for the sixteen male and for the sixteen female faces were obtained. This defined the average configuration of a male face and the average configuration of a female face. The original image of each male face was then distorted into the average male configuration and the sixteen resulting distorted male images were blended digitally into the average male shape to form the male prototype face (face 1). The process was repeated for the sixteen female faces which were blended digitally into the average female shape to form the female prototype face (face 4). (The resulting images are not unlike those illustrated in Grusser & Landis (1992) and obtained by Katz (1953) who used sophisticated photographic superposition techniques to make average male and female faces). Each face was masked to exclude ears, neck, hair and hairline by a border area which was rendered black. The border was defined by the jawline and ear positions and an arc from the point where the top of the right ear joins the head, centred half way down the bridge of the nose to the point where the top of the left ear joins the head. The largest dimensions of the male face measured 12.3cm x 10.5cm and those of the female face measured 11.7cm x 10.2cm. Inter-pupillary distance for both faces was 4.1cm. These two faces were used to synthesize a number of faces along a continuum changing gradually from the male prototype face to the female prototype face.

Choice of the number of steps in the continuum.

Pilot studies were conducted to determine the number of steps along a continuum necessary to allow discrimination above the level of chance. This was to avoid the possibility of making the steps too small for discrimination (in either a categorical or a non-categorical manner). A first pilot study using seven stimuli showed that five subjects could not discriminate between any pair of stimuli which were adjacent to each other on the continuum at an above chance rate.

With a sequence of an odd number of stimuli the stimulus at the centre of the continuum was an amalgam of 50% of the female prototype face and 50% of the male prototype face thus producing an androgynous face which was neither male nor female. It was reasoned that a continuum including the androgynous face, in theory at least, could make identification of a sharp boundary between male and female unlikely.

It was decided that the differences between the stimuli were too small and that an even number of stimuli should be used to avoid having a theoretically androgynous face in the centre of the continuum. A further pilot study using six stimuli showed subjects' performance at the level of chance across all pairs of stimuli, suggesting that the differences between the stimuli were still too small to be detected, so the experiment proper was conducted using four stimuli.

Faces 2 and 3 were formed by deriving a configuration that was intermediate to faces 1 and 4, blending them in the following proportions:-

face 2 = 33.3% of face 1 + 66.7% of face 4; face 3 = 66.7% of face 1 + 33.3% face 4; see Figure I.

Subjects.

Subjects were 9 female and 5 male undergraduates (aged 19-22 years).

Design.

A within subjects design was used, each subject performing two tasks:-

1. Discrimination task.
2. Identification task.

Procedure.

The stimuli were displayed on a Silicon Graphics Personal IRIS computer monitor at a viewing distance of approximately one metre.

<u>Discrimination task.</u>

For the discrimination part of the experiment four different randomisations of twelve pairs of faces were programmed. Each random set contained six pairs which were alike and six pairs consisting of two different faces which were adjacent to each other on the scale from male 'prototype' to female 'prototype'. In three of these different pairs the most female of the two was presented first and in the other three the most male was presented first. This distribution of pairs meant that each of the four matching pairs was presented six times for discrimination and each of the three different pairs was presented eight times.

Subjects were told that they would see two faces on the screen one after the other and that sometimes the two would be the same and sometimes they would be different. Their task was to press one key if the two were the same and another if they were different. The first of a pair of faces was displayed for one second, an interstimulus interval of one second followed, and then the second face was displayed for one second. As soon as the subject had responded the next pair was presented. Trials were self-paced with no requirement placed on speed of response. The left and right positions of same and different response keys were counterbalanced across subjects.

Identification task.

This task was always performed after discrimination to avoid the possibility that it would influence the strategy used for discriminating. Subjects were told they would now see the faces singly and their task was to classify them as either male or female and respond by pressing one of two keys. Each face was shown for one second and the next trial began after response to the first. The position of the response keys was counterbalanced as before. Each of the stimuli was presented four times in random order.

Results.

Identification.

Responses are illustrated as mean percent 'male' response for each single face in Figure II. A clear boundary between those faces categorised as male and those categorised as female is demonstrated by the sharp fall in the graph between faces 2 and 3. A oneway ANOVA of the data showed a highly significant effect of the position of the face on the scale from male to female $F(3,39)=73.98$, $p<0.001$. Post hoc Tukey's HSD tests showed that responses to faces 1 and 2 were significantly different from those to faces 3 and 4 ($p< 0.001$). Differences between responses to faces 1 and 2 and between responses to faces 3 and 4 were not significant.

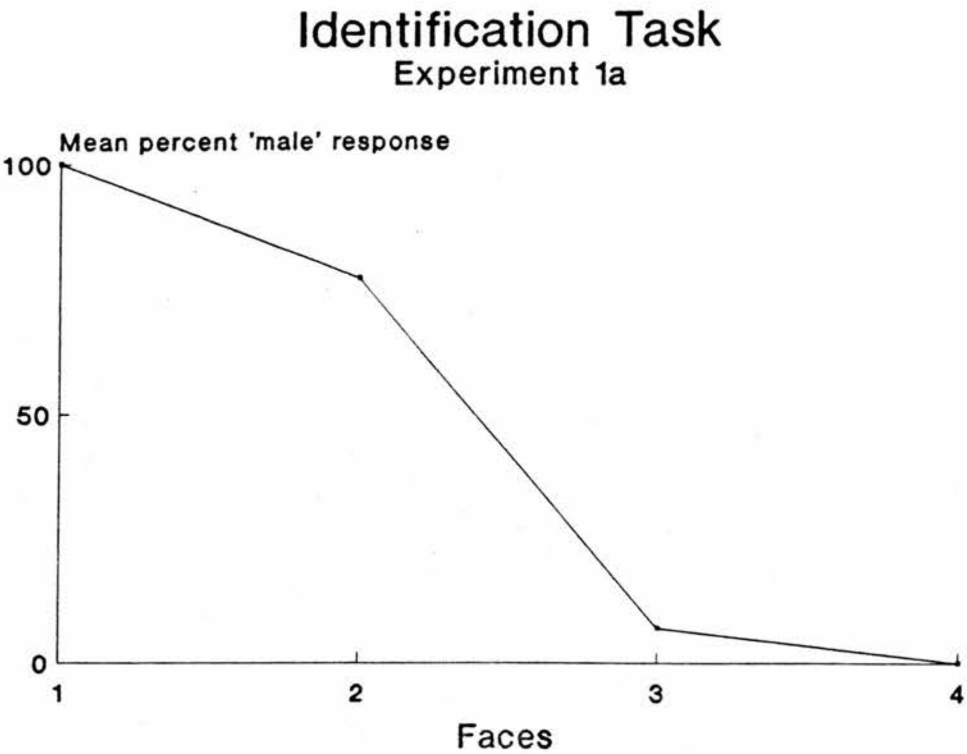


Figure II. Mean percent 'male' response as a function of the position of faces on the scale from male to female in the identification task Experiment 1a.

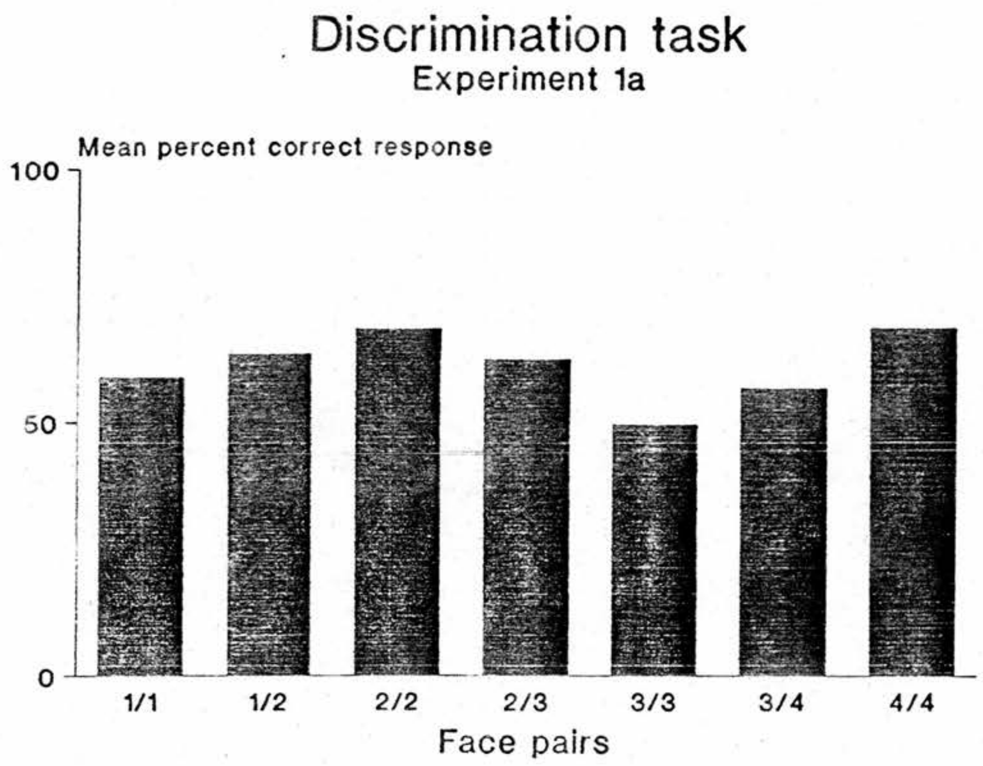


Figure III. Mean percent correct response for same and different pairs of faces in the discrimination task Experiment 1a.

<u>Discrimination.</u>

The results of the discrimination task are plotted as mean percent correct response for each pair of faces in Figure III. There is no peak of discrimination between faces 2 and 3 corresponding to the mean category boundary. A oneway ANOVA of the data showed no significant effect of pair of faces, $F(6,78)=1.5$, $p=0.19$).

Since some subjects did show a peak of discrimination which corresponded to their category boundary and a few had spontaneously said they noticed that some of the stimuli were girls or boys all subjects were debriefed as to whether they had made their discrimination on the basis of gender or not. Some said they had done so part of the time but they were not the subjects whose peak of discrimination corresponded to the category boundary. Most said they had used a variety of criteria, concentrating on a detail such as shape of lips when making their decision, but changing from one detail to another during the course of the experiment.

It was decided to test another group of subjects giving them instructions which were intended to encourage them to see the faces as a whole. It was hoped that that would induce them to discriminate on the basis of gender.

Experiment 1b.

Eleven new subjects, 4 male and 7 female aged 19-24 years, were instructed as before but with the additional information that because the task was difficult there was a tendency to concentrate on a small detail of the face. They were requested to try to resist that tendency and to see the face as a whole as if it was a person.

Results.

Identification.

Results of the identification task for this second group of subjects are plotted as mean percent 'male' response for each face in Figure IV. A sharp category boundary between faces 2 and 3 is again clearly demonstrated. A oneway ANOVA of the data again showed a highly significant effect of the position of the face on the male to female scale $F(3,30)=139.49$, $p<0.001$. Tukey's HSD tests showed that responses to faces 1 and 2 were significantly different from responses to faces 3 and 4 ($p=0.001$), but the differences between responses to face 1 and face 2 and between responses to face 3 and face 4 were not significant.

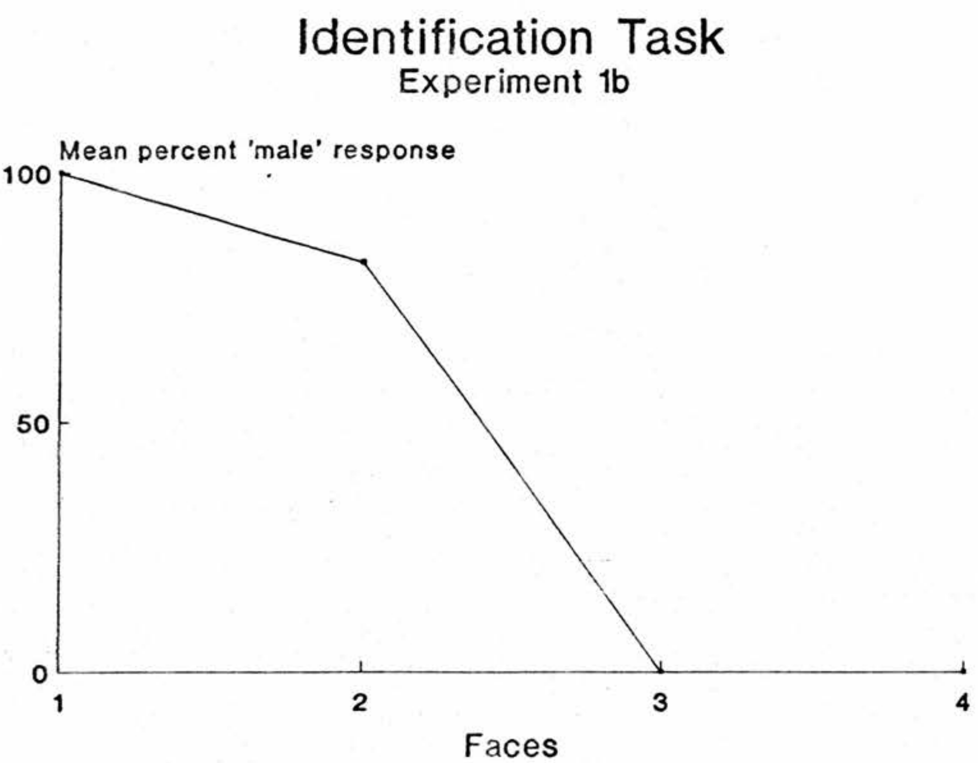


Figure IV. Mean percent 'male' response as a function of the position of faces on the scale from male to female in the identification task Experiment 1b.

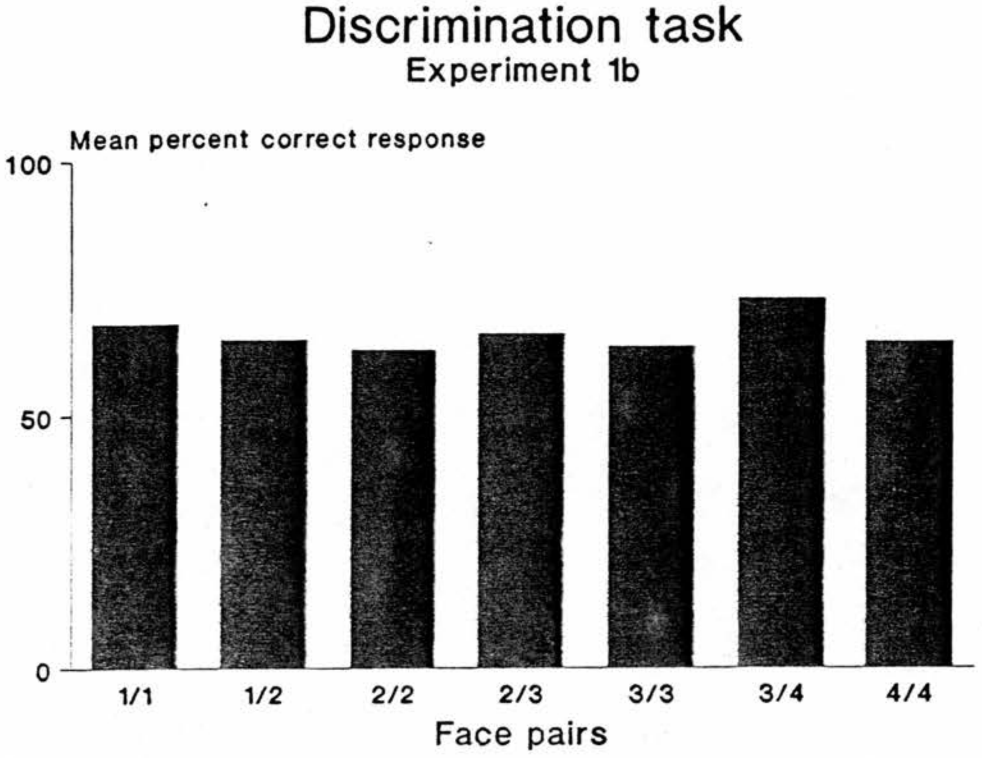


Figure V. Mean percent correct response for same and different pairs of faces in the discrimination task Experiment 1b.

Discrimination.

Results of the discrimination task plotted in Figure V again show no peak of discrimination. A oneway ANOVA of the data showed no significant effect of pair of faces, $F(6,60)=0.3$, $p=0.94$. On debriefing about the strategy they had used this second group of subjects also said that they had scanned the face for differences in detail and had not thought of the faces as being male or female.

Although for a large majority of subjects in both groups the category boundary between male and female fell between faces 2 and 3 there were a few for whom the boundary fell at a different point on the continuum. Since those whose category boundary was not between faces 2 and 3 could not be expected to show a peak of discrimination at that point their data were excluded. The mean percent correct response in the discrimination task for eighteen subjects (nine from Experiment 1a and nine from Experiment 1b) whose category boundary fell at the 2/3 step on the continuum is plotted in Figure VI.

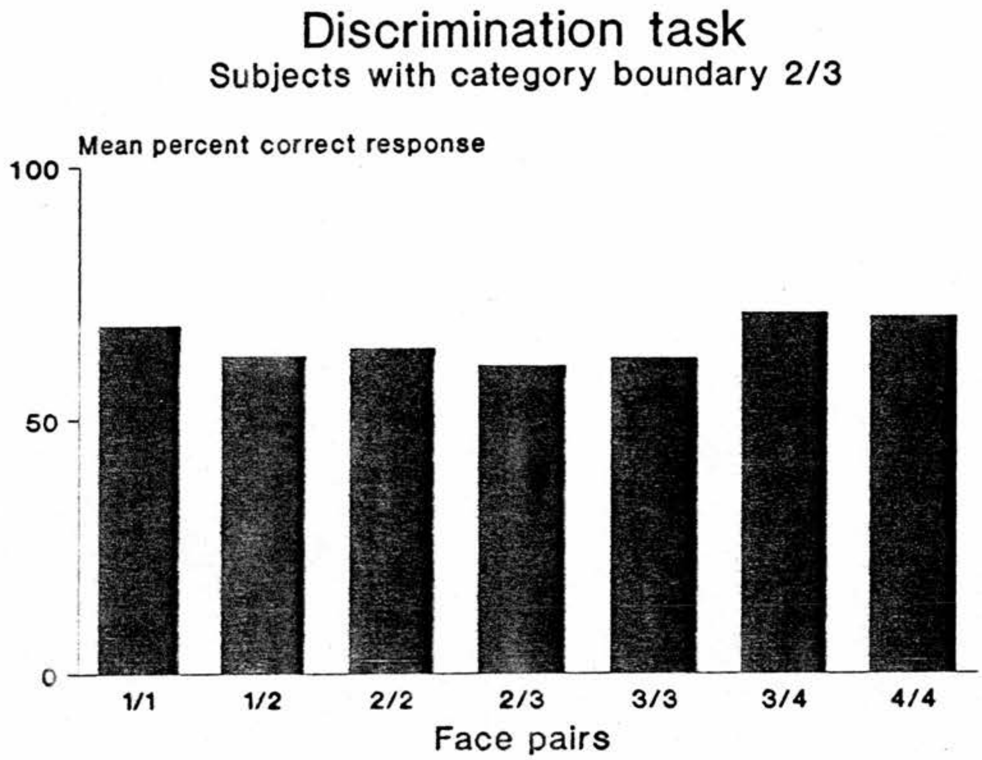


Figure VI. Mean percent correct response for same and different pairs of faces in the discrimination task for all subjects whose category boundary fell between faces 2 and 3.

There is no suggestion of a peak of discrimination at the 2/3 boundary. A oneway ANOVA of these data also showed no significant effect of pair of faces, $F(6,102)=0.70$, $p=0.65$. Students t-tests of the scores for each pair of faces against a hypothesised mean of 50% showed discrimination performance to be significantly above chance ($p<0.01$ in all cases). A Students t-test showed the difference between the results for same pairs and for different pairs to be non-significant ($t=-0.391$, $p=0.712$).

Discussion.

Although a sharp category boundary was demonstrated in the identification task it is clear that the ease of discrimination between stimuli changing from male to female is the same throughout the continuum; that is, discrimination is continuous. There are a number of possible explanations why this happened.

Strategy used for discrimination.

The evidence of spontaneous reports and debriefing of subjects showed that most discriminated between stimuli on the basis of individual detail and ignored the rest of the information available to them. If subjects could detect a difference in the size of a nostril or the thickness of an eyebrow, that was sufficient evidence on which to make a discrimination. There was no need for any assessment of the overall gender of the face. They were not discriminating on the basis of gender, so demonstration of a peak of discrimination corresponding to the category boundary between the two genders could not be expected.

Differences in performance depending on what basis subjects discriminated between cues were demonstrated by Pastore, Schmuckler, Rosenblum & Szczesiul (1983). Using duplex perception when one acoustic signal was presented to one ear and another to the other ear they showed that some subjects perceived the two together as musical chords. The chords were discriminated categorically at a point on the continuum were a change from major to minor key occurred. Two subjects (who, the authors suggest, had perfect pitch) perceived the separate single tones categorically. No subject perceived both chords and single tones categorically. The peak of discrimination occurred at a different point on the continuum depending whether subjects perceived the sounds as chords or as separate tones. This was interpreted to mean that some subjects were listening in the musical 'mode' and some in the acoustic 'mode'. Best, Morrongiello & Robson (1981), investigating trading relations between acoustic cues, also found differences in performance depending on what basis subjects discriminated between synthesised speech sounds. Those subjects who were able to obey instructions to discriminate on a phonetic basis showed a peak of discrimination corresponding to the phonetic category boundary. Those subjects who were unable to hear the sounds as speech or had been instructed to discriminate on the basis of the temporal or spectral information in the stimuli showed very different results from each other and from the listeners discriminating on a phonetic basis. A subject had to be in the phonetic 'mode' for the categorical effects to emerge for those stimuli. According to their own assertions subjects in the present experiment were not usually consciously in the gender 'mode' and categorical effects were not demonstrated.

Memory limitations.

One possible explanation for categorical perception, proposed by Hary & Massaro (1982), is that it occurs because of memory limitations. Pisoni (1973), investigating perception of speech sounds, found that at short interstimulus intervals (ISIs) subjects were able to differentiate stimuli on the basis of acoustic cues and to discriminate two versions of the same vowel sound (ie intra-category discrimination). At longer ISIs subjects had to rely on phonetic codes and consequently could discriminate different vowels but not different versions of the same vowel. Thus at these longer ISIs subjects could perform inter-category discrimination but not intra-category discrimination. Extra memory load may favour stimulus coding at a higher level and perhaps if a longer ISI had been used in this experiment subjects might have been forced to use a different strategy from the one they employed here. It might have pushed them into the gender 'mode'.

However, if subjects had classified the faces by gender in order to discriminate one from another, there is no reason why discrimination between two faces which had been categorised as having the same gender should be at the level of chance. Chance performance in a discrimination task must mean that the subject is unsure of both similarity and difference, in other words s/he must have detected a suspicion of dissimilarity about which s/he is not certain. Subjects had no difficulty in classifying the gender of faces which were presented for only one second and should not have had difficulty remembering that classification (had they made it) even if they did have difficulty remembering the individual physical details of a face.

Multiplicity of cues.

The number of acoustic properties which can change in the contrast between one phoneme and another is large. Lisker (1978) named 16 which contribute to the distinction between /b/ and /p/ and the integration of the information thus conveyed is clearly important. A second experiment to be described later showed that there are many cues to gender in a face, each feature carrying some relevant information. That means that if subjects had discriminated on the basis of gender they could have used a number of different cues each of which would have given them some information relevant to that judgment. For each of those cues, however, the change from male to female might occur at a different point on the continuum. For example, the transition from female eyes to male eyes might take place between faces 3 and 4 and the transition from female chin to male chin between faces 2 and 3. A given subject's judgment of the gender of a given face could vary depending whether that judgment relied more on the evidence from the eyes or on the evidence from the chin. That would mean that demonstration of a clear peak of discrimination at one point on the continuum would be less likely than discrimination spread at an above baseline rate over all stimulus pairs. This constitutes an argument for varying stimuli in only one aspect when attempting to demonstrate categorical perception in both identification and discrimination tasks.

Summary.

The rationale is explained for attempting to demonstrate categorical perception of the gender of faces using similar methods to those used in investigation of speech perception. An experiment is described in which a continuum of synthesised faces changing gradually from male to female in four equal steps was used in identification and discrimination tasks. A discrete categorical boundary was demonstrated in the identification task but discrimination of faces across that boundary was no more accurate than discrimination between two faces which had been identified as being of the same gender. Reasons for this are discussed.

Chapter Two.

It is not known which property of a face carries the critical information for perception of its gender. Most of us think we have a clear concept of what makes a face female or male and yet we cannot say exactly how we make that judgment. Apart from male facial hair and different hairstyles we cannot say exactly when a property of a face is male and when it is female. We do not know whether individual features (in the common sense of eyes, nose, mouth) carry information about gender or whether the gender of a face is expressed in the arrangement or configuration of the features within the whole or in the surface contours of the face or in the shape of its external border or in some combination of these or other factors.

Liggett (1974) notes that the female face is only about four-fifths the size of the male and that the female nose is also proportionately smaller and proportionately wider than the male. Its shape is more concave and its bridge more depressed. He also makes the point that these characteristics of the female nose - smallness, wideness and concavity - are also pronounced in children. He says that the female mouth is relatively smaller and the upper lip often shorter than the male and that the female jaw and brow-ridges are less pronounced. The female eyes are larger with darker shading surrounding them, particularly in the young, and the scantier female eyebrows become thinner with increase in age. By contrast the male brows are generally thinner in youth but grow thicker, longer and coarser with age.

Nakdimen (1984) quotes Birdwhistell (1970), "a man's eyebrows are heavier, straighter and closer to the eyes. A woman's......are usually more arched". He also says that high cheekbones in men are a rather severe, bony feature whereas in women they are covered by a pad of soft tissue and appear more rounded.

Enlow(1982) maintains that "the male nose is proportionately larger than the female nose" and says this is because the male needs to fill larger lungs to support the oxygen requirements of his larger body. This larger size of the nose is said to lead to differences in the shape of the forehead and cheekbones. Linney and Coombes (personal communication), using a laser scanning technique which is described in Moss et al (1987) have produced 3-D images of an average male and an average female head. They too have found that the male has a more protruberant nose and that he has more protruberant brows and chin/neck than the female, while she has slightly more protruberant cheeks. These differences were found largely in the 3-D structure of the face as a whole rather than in the detail of the individual features. Indeed, the shapes of individual features which are normally made visible by their pigmentation are not very clearly portrayed in these images of the average 3-D head models. Although the differences between the two genders assessed by this method of measuring 3-D distances are objective that does not mean that subjects can perceive gender on the basis of those differences. Nor does it mean that feature shape, defined by pigmentation and visible in 2-D images such as photographs, will be capable of supporting gender perception.

Harmon (1973) showed that information about the gender of an unfamiliar face can be extracted from a coarsely quantized 2-D image of it. Sergent (1986) using the same technique notes that in coarsely quantized images detailed information about individual features is indiscernible so that assessment of gender cannot be based on that. It could instead be based on the available configurational cues; that is, the relative positions of the features in the face/head as a whole. Haig (1984) has shown that we are sensitive to very small displacements of features in judgments of identity and that may also be true for judgments of gender. These findings beg the question of whether individual features play any part in assessment of gender at all.

Roberts & Bruce (1988) attempted to assess the role of individual features by masking with tape either the eyes, nose, mouth or none of the features of photographs of unfamiliar faces and then recording response times taken to judge their gender. They do not specify whether 'eyes' refers to the whole eye region including the brows or the eyes alone without brows. Results showed response times increasing when eyes, mouth or nose were masked suggesting that some information about gender is given by the eyes, a little more by the mouth and more again by the nose. When the features were presented in isolation (ie the rest of the face was masked) it was found that the eyes provided the most reliable information for gender judgment and the nose the least. The difference between the effect of masking the nose and presenting it in isolation was interpreted as meaning that the nose did not carry information about gender in itself but its position in the overall configuration of the face did. This again suggests an important role for configuration in perception of gender in a face. A second experiment was therefore devised to investigate further the role of individual features in the perception of gender.

Experiment 2.

Using a different method from that of Roberts & Bruce (1988) this experiment examined the extent to which features in isolation may allow discrimination of gender. A second method of assessing the gender value of a facial feature tested the effect on perceived gender of interchanging features between faces of opposite genders. Davies, Ellis & Shepherd (1977) and Haig (1986) had shown that interchanging different features had varying effects on perception of face identity. Davies et al manipulated the features contained in Photofit Kit and found that hair and forehead substitutions were noticed most accurately, then eye or mouth substitutions followed by nose substitutions. Chin substitutions were more or less noticeable depending how different the substituted chin was from the original. Haig found a similar pattern with changes in the head outline being most noticeable, then changes in the eye/eyebrow combination, then the mouth and last the nose. A similar order of importance might also hold to be the case in perception of gender.

Method.

For the purposes of the experiment the outline shape of the face and the eyebrows and chin were counted as features as well as the eyes, nose and mouth used by Roberts & Bruce (1988). Blank faces with no features were used as stimuli representing outline shape. Features were presented in pairs as well as singly because it is possible that gender only becomes apparent when the distance between two features is seen (eg brows - eyes, nose - mouth).

To test the effect of feature substitutions on the perceived gender of a face, male features were grafted into a female face and female features into a male face. The resulting composite faces were then presented for classification as male or female. The number of times the perceived gender changed from that of the original face was taken as a measure of the influence which that feature had on the perceived gender of the face.

Stimuli.

Stimuli (illustrated in Figure VII) were created by isolating and manipulating parts of the male prototype face and the female prototype face which had been used in the first experiment.

Eighteen images of parts of the prototype male and female faces were made by masking them horizontally to form a slice of face which contained one or two features. Single features isolated in this way were the brows, eyes, nose, mouth, chin and pairs of features isolated were brows &

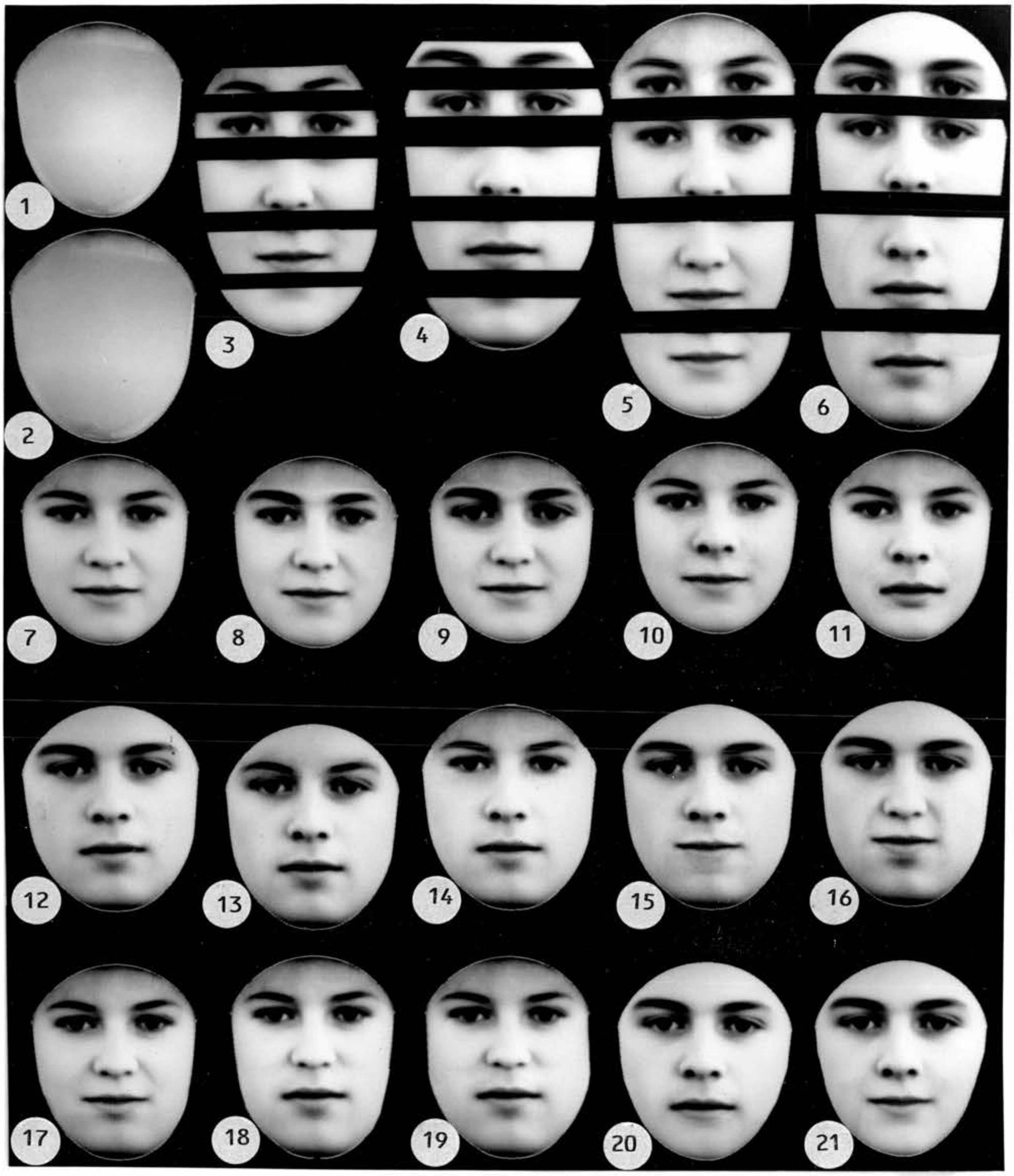


Figure VII. Stimuli used in Experiment 2.
1.Blank female face shape. 2.Blank male face shape. 3.Single female face parts. 4.Single male face parts. 5.Pairs of female face parts. 6.Pairs of male face parts. 7.Female prototype face. 8.Female face with male brows. 9.Female face with male brows & eyes. 10.Female face with male nose. 11.Female face with male nose & mouth. 12.Male prototype face. 13.Male face with female brows. 14.Male face with female brows & eyes. 15.Male face with female mouth. 16.Male face with female nose & mouth. 17.Female face with male chin. 18.Female face with male jaw (version 2). 19.Female face with male jaw (version 1). 20.Male face with female chin. 21.Male face with female jaw.

eyes, eyes & nose, nose & mouth, mouth & chin (hereafter referred to as the jaw because that name better describes the area of face contained in the slice) of each face. Two further images were created of the overall shapes of the two prototype faces with the features blanked out.

Nineteen new images were formed by grafting a 'part' from the prototype face of one gender onto the face of the other gender, keeping the rest of the recipient face as it was. This was done using the 'Designer Paint' software of the Pluto 24:i graphics system. This allows part of an image to be copied in any other desired position using a 'cut and paste' function. The edges of the grafted part were then blended into the rest of the recipient image by grading pixel intensities across the join. In this way nine images of the male face with female features and ten images of the female face with male features were made. The features changed were brows, eyes. nose, mouth, chin, brows & eyes, eyes & nose, nose & mouth, and jaw. Two versions of the female face with male jaw were made because the male jaw was wider than the appropriate part of the female face. In version 1 the width of the male jaw was reduced and in version 2 the width of the female cheeks was increased. The effect of both types of alteration were tested in the experiment. When joining the male face and female jaw only one version was made by reducing the width of the male cheeks slightly.

Each continuous tone black and white image was photographed from the monitor screen using a Minolta 7000 camera, FP4 black and white film, and a one second exposure.

Subjects.

Subjects were ten females aged from 22 to 74 (average 45), and ten males aged from 14 to 52 (average 39).

Procedure.

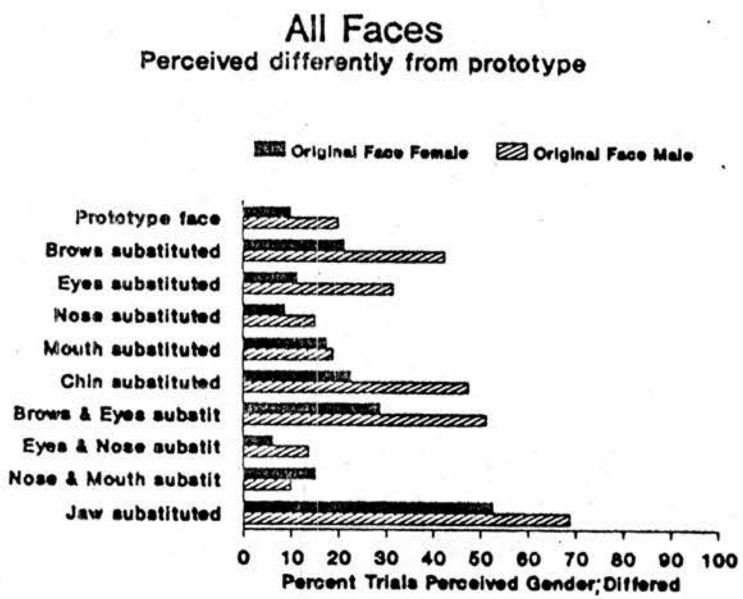
The 21 photographs of whole faces (2 prototypes and 19 faces with feature substitutions) were mixed at random and subjects were asked to place them in two piles, one for faces perceived as female and one for faces perceived as male. They were allowed as much time as they needed but were asked to give their first impression of the gender of the face rather than a considered opinion. The experimenter recorded which images had been perceived as male. The photographs were mixed at random again and the subject was asked to repeat the task until each photograph had been classified four times. Exactly the same procedure was then followed with the twenty photographs of parts of faces. Parts of faces were always presented after whole faces to avoid encouraging any tendency for subjects to concentrate on details of a face instead of processing the whole of it.

Results.

Whole faces.

The results for classification of whole faces are illustrated in the first graph in Figure VIII. They were analysed as percent of the total number of trials when a prototype face was assigned the wrong gender and when a face with feature/s substituted was assigned a gender which was different from that of the original prototype face. The difference in the responses for each of the two versions of the female face with a male jaw was subjected to a Wilcoxon test and proved not to be significant, so the data from version 2 were used and those from version 1 were not. Because the stimuli had been classified by a forced choice a fixed number of times the data were transformed using an inverse sine transformation.

A 2 (gender of subject) x 2 (original gender of facial image) x 10 (feature/s substituted or no substitution) mixed ANOVA showed a significant main effect of the original gender of the face, male faces being assigned a



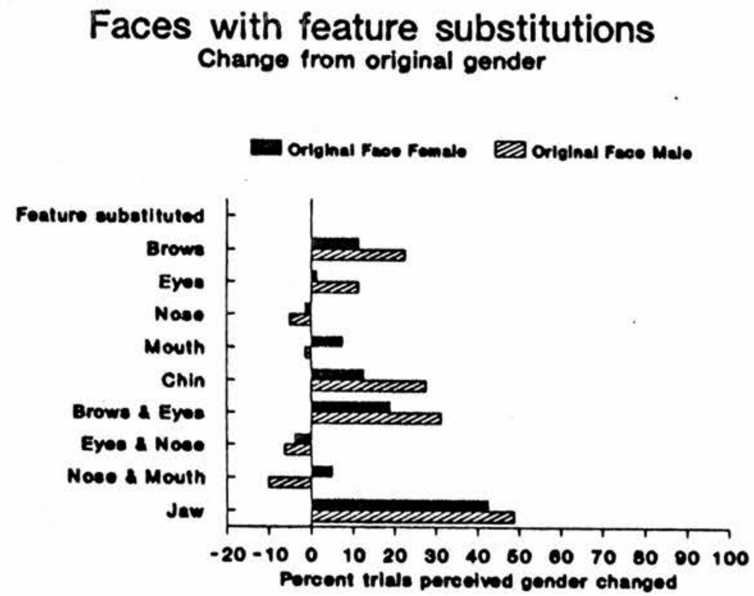


Figure VIII. Left. The percentage total number of trials when each face with substituted feature/s was classified as having a gender different from that of the original face compared with the number of times the prototype faces were assigned the wrong gender.
Right. The same data with each count reduced by the number of times the corresponding prototype face was assigned the wrong gender.

different gender more frequently than female faces, $F(1,18)=14.65$, $p=0.0012$. The main effect of feature substituted was significant, $F(9,162)=15.56$, $p<0,0001$. Post hoc Protected Least Significant Difference (PLSD) tests (Snedecor & Cochran, 1980) showed that substitutions of eyes, brows, chin, brows & eyes and jaw (in ascending order) resulted in significantly increased assignation of the opposite gender to a face compared with that to the appropriate prototype face.

The interaction between feature substituted and original gender of the face was significant, $F(9,162)=3.59$, $p=0.0004$. Post hoc PLSD analysis showed that when substitutions of male features into the female prototype face were made, substitution of the mouth, mouth & nose, brows, chin, brows & eyes and jaw effected a significant change in the perceived gender of the face. When substitutions of female features into the male prototype face were made, however, it was substitution of eyes, brows, chin, brows & eyes and jaw which was effective in changing the perceived gender of the face. The main effect of subject gender was not significant, but gender of subject interacted significantly with gender of the original face, $F(1,18)=10.41$, $p=0.0047$.

The prototype male face had been classified as female more often than the female prototype was classified as male (see Figure XI) suggesting that the male prototype was less male than the female was female. The figures were therefore adjusted by subtracting the individual subject's scores for the prototype faces from each of their scores for faces with feature substitutions. These normalised results are shown in the second graph in Figure VIII. A 2 (gender of subject) x 2 (original gender of facial image) x 9 (feature substituted) ANOVA of the normalised data showed a significant main effect of feature substituted, $F(8,144)=15.06$, $p<0.0001$. Main effects of gender of subject and original gender of the facial image were not significant. The interaction between the original gender of the facial image and the particular feature substituted was significant, $F(8,144)=3.85$, $p=0.0004$. Other interactions were not significant.

Post hoc PLSD comparisons ($p<0.05$) showed that substitutions of brows, chin, brows & eyes or jaw (in ascending order) effected a significant change in the perceived gender of the prototype faces. Jaw substitutions were significantly more effective than all other substitutions. There was no significant difference between the effects of brows and of brows & eyes substitutions showing that the brows were the most important part of the eye region in bringing about the change in perceived gender.

Post hoc PLSD analysis of the interaction between the feature substituted and the gender of the original prototype face indicated that substitutions of male jaw and male brows & eyes into the female prototype face effected a significant change in perceived gender while other substitutions did not. For substitutions of female features into the male prototype face brows, chin, brows & eyes and jaw were effective in significantly changing its perceived gender. Of these four feature substitutions only brows and jaw were significantly different from each other in their effect.

The interaction between gender of subject and original gender of facial image which was significant before normalisation of the data is not significant after normalisation. The significant interaction before normalisation is nevertheless of interest because it shows differences between male and female subjects' assignment of gender to the same faces. The interaction is illustrated in Figure IX. Male faces with female feature substitutions were perceived as female more often by male than by female subjects. Female faces with male feature substitutions were perceived as male less often by male than by female subjects.

The female subjects' responses were very similar for the various feature substitutions whether the original face was male or female. When

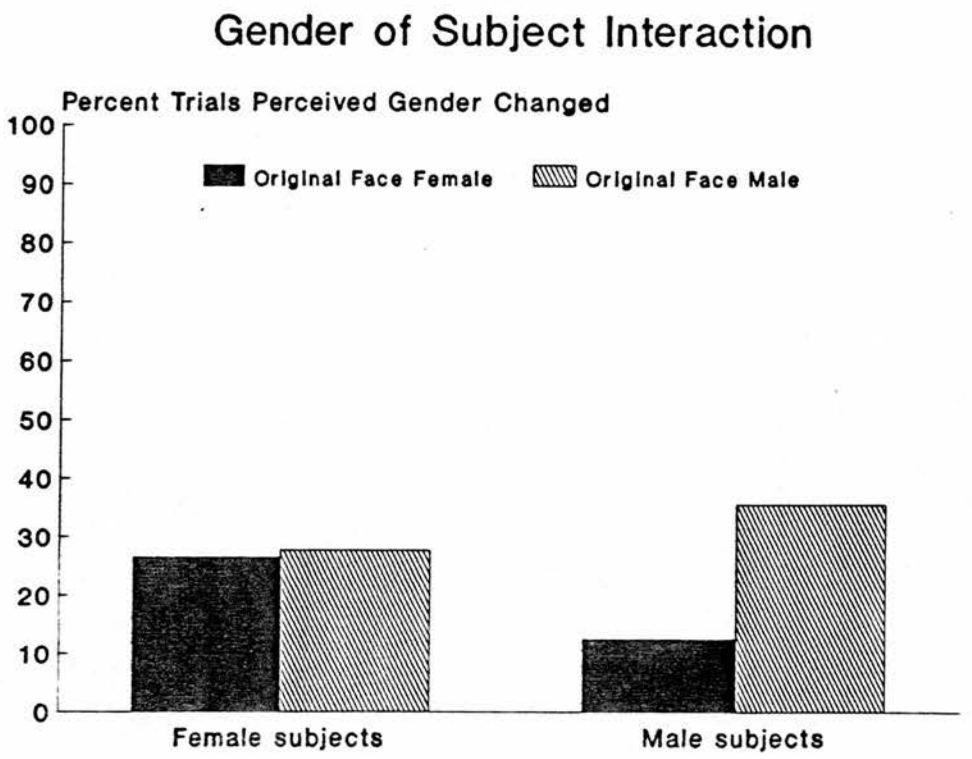


Figure IX. The percentage total number of trials, for male and female subjects, when the perceived gender of the faces changed following feature substitutions.

subjects were females correlation of the number of times the perceived gender of male and female faces changed after substitution of a feature of the opposite gender gave a correlation coefficient which was highly significant (Spearman's $r=0.83$, $p<0.005$). When subjects were males the correlation was not significant (Spearman's $r=0.54$, $p>0.05$). This difference between male and female subjects' responses is illustrated in more detail in the two graphs in Figure X.

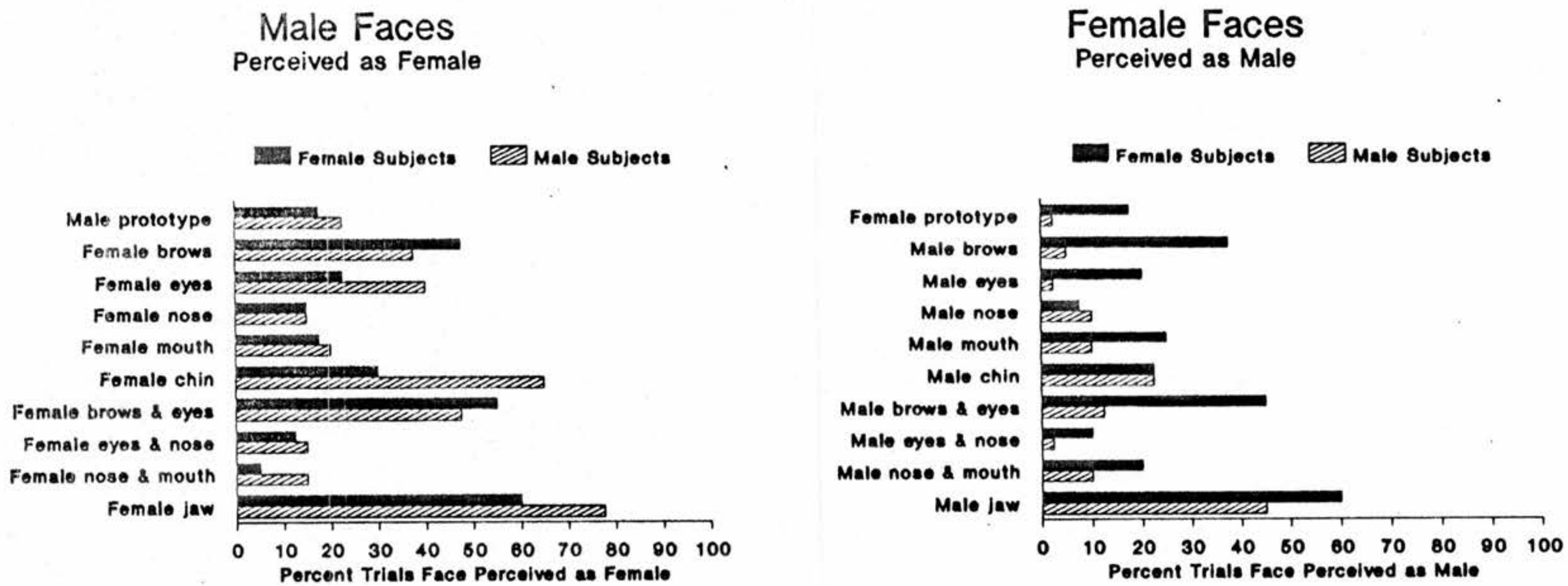


Figure X. The percentage total number of trials when male and female subjects:-
Left. classified the male face with female feature substitutions as female.
Right. classified the female face with male feature substitutions as male.

Parts of faces.

The percentage of the total number of trials when each of the photographs of parts of a face were classified as male is shown in Figure XI.

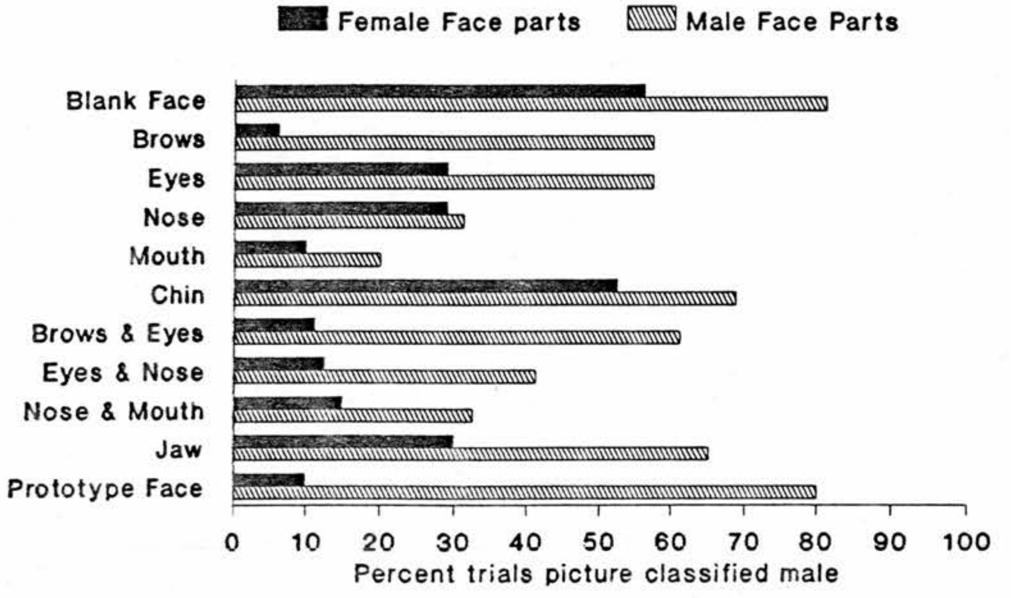


Figure XI. The percentage total number of trials when each photograph of part of a face was classified as male compared with the number of times the prototype faces were classified as male.

Data for classification of the prototype faces are also included in the graph for comparison. The data were transformed using an inverse sine transformation as for whole faces and submitted to a 2 (gender of subject) x 2 (gender of facial image) x 10 (part of face) mixed ANOVA. This showed a significant difference between male and female face parts, $F(1,18)=85.49$, $p<0.0001$, parts from the male prototype face being classified as male more frequently than those from the female prototype face. There was a highly significant main effect of part of face, $F(9,162)=12.09$, $p<0.0001$, some parts being classified male more often than others and there was a highly significant interaction between the part of the face represented by a picture and its gender, $F(9,162)=3.79$, $p=0.0002$. There was no significant main effect of gender of subject. Other interactions were not significant.

The main effect of part of face was not of interest for the purposes of the experiment but the interaction between part of face and its gender was; that is, the difference between the way a part taken from a male face was classified and the way the equivalent part taken from a female face was classified.

Post hoc PLSD tests showed that the difference in the way male and female versions of parts of faces were classified was significant ($p<0.05$) for all parts except the nose. The other parts, ranked in ascending order of magnitude of the difference between classifications of the male and female versions, read mouth, nose & mouth, chin, jaw, blank face, eyes, brows, brows & eyes. There was a significant difference between mouth and nose & mouth, between nose & mouth and eyes, between eyes and brows, and between brows and brows & eyes.

Discussion.

The significant main effect of gender of face parts presented in isolation, shows clearly that individual features do carry information about gender, the male and female versions of each part, except for the nose, being classified differently. The results confirm those of Roberts & Bruce (1988) who showed that when features are viewed in isolation the eyes carry more information about gender than the mouth and the mouth more than the nose. This contrasts with the statements by Liggett (1974), Nakdimen (1984) and Enlow (1982) about differences between the male and female nose. That could mean that the physical differences between male and female noses are not of psychological significance. Alternatively, the lack of importance of the nose in the current experiment and that of Roberts and Bruce may be because the nose shape is not clearly visible unless it is viewed in profile or three-quarter view. The latter seems intuitively to be the more likely explanation.

Responses to the blank faces were quite illuminating. Subjects seemed to be unable to see blank shapes in face terms and wanted to know what they were supposed to do with the photographs. It seems that a face without a configuration of features is not perceived as a face at all.

One thing is very clear from the results of this experiment. That is that the smallest area of face (ie the eyebrows) carries more information about gender than the largest (ie the eyes & nose together) when they are seen in isolation. It is also apparent when the features are substituted for each other that the particular parts changed matter more than the total area of the face changed, brows again having more effect on perceived gender than larger areas such as nose & mouth or eyes & nose.

The results for feature substitutions are roughly in agreement with those of Davies et al (1977) and Haig (1986) in their investigations of the effects of feature substitutions on identity. Results cannot be compared precisely because the substitutions are a little different but Davies et al and Haig found alterations to the outline of the face and to the eye area more noticeable than alterations to the mouth and nose. In the present experiment alterations to the jaw area (which change the outline) and to the eye area were most effective in changing perceived gender.

There is a difference in the order of magnitude of gender information carried by the parts of faces depending on whether they are viewed in isolation or grafted into a face of the opposite gender. Ascending order for parts seen in isolation is mouth, nose & mouth, chin, jaw, eyes, eyes & nose, brows and brows & eyes. For feature substitutions it is brows, chin, brows & eyes, and jaw. This may be because, when features are interchanged in whole faces, configurational information comes into play whereas it does not when they are viewed in isolation. Brows substitutions shift the balance of the pattern of features within the shape of the face (the male brows being heavier than the female). There is also a difference in the gap between the two brows, which is proportionately smaller in the male face than in the female, and that too would change the pattern or configuration of the face as a whole. Chin and jaw substitutions alter the vertical dimension between base of nose and tip of chin. The fact that the two versions of the female prototype face with a male jaw were not classified differently suggests that the vertical dimension is more important than the horizontal in the jaw area. This experiment has only partially disentangled the influences of individual features by themselves and of their contribution to the overall pattern or configuration of the face.

Figure XI shows that most of the male parts were classified as being less strongly male than the female parts were female. Liggett's statements about differences in developmental changes between males and females, particularly in the eye region, could explain some of that. The original faces used in the creation of the prototype faces were those of young undergraduate students so the male eyebrows would be thinner than those of a mature male while the female ones would be thicker than those of a mature female. If childlike features are perceived as being more female, then immature faces, being more childlike than mature ones, would appear more female. If mature adult faces had been used instead of youthful ones the faces and parts of faces might have been perceived as being more masculine.

Several subjects remarked that some of the whole faces looked like a nun, one that some were like a North African woman and another like an Arab woman. The masking of hair, hairline and ears was intended to exclude obvious clues about gender from hairstyles or hairlines, but it seems to have given the illusion of a veil and possibly biased perception of the faces towards the female. To separate that effect from the possible effect of the immaturity of the faces used here it would be necessary to repeat the experiment using more mature faces. If a bias towards perception of the faces as female was still demonstrated when mature faces had been used to create the prototypes it could not then be attributed to immaturity in the faces. It could, however, still be attributed to the illusory 'veil' effect caused by masking the faces to exclude the evidence of hairstyles and hairlines.

A surprise finding in the experiment was the interaction between gender of the original face and gender of subject. Almost all the variation between the alteration in perceived gender of male faces and of female faces following feature substitution was due to the male subjects' responses. They required more masculinity to classify a face as male than females did and consequently less femininity to classify it as female. The two sexes gave different values to the evidence about gender in a face. For example, female subjects took more account of brows than male subjects (see Figure X). The raw data also showed large variation between subjects of the same sex indicating that the perception of gender is subjective rather than absolute. Any attempt to define in absolute terms what is male and what is female would therefore be doomed to failure. It would only be possible to find average concepts of gender characteristics.

Summary.

This experiment assessed the capacity of individual features of a face viewed in isolation to convey information about its gender. It also tested the effect on perceived gender of grafting male features into a female face and female features into a male face. The results showed that, for the faces used, all features except the nose carried information about gender when they were seen in isolation. The jaw, brows & eyes, chin and brows (in descending order of magnitude of effect) caused significant change in perceived gender when they were grafted into a face of the opposite gender. A probable role for configuration in perception of gender is suggested. Differences in perception of gender by male and female subjects were also revealed.

Chapter Three.

It has now been established that a part of a face carries information about gender by itself and that substituting a part from a face of the opposite gender can alter the perceived gender of a recipient face. It was argued in Chapter One that only one aspect of a face should be changed along a continuum if categorical perception was possibly to be demonstrated. With information as to which parts of a face had the most significant influence on its perceived gender it was now possible to devise such an experiment. A continuum of faces could be created in which one part of the face was changed from a male version to a female version while the rest of the face remained the same.

Experiment 3.

Experiment 2 showed that jaw substitutions had the most influence in altering the perceived gender of a face. It was decided, therefore, to vary the jaw in a continuum of faces to be categorised by their gender and to be discriminated one from another.

Returning for the moment to auditory studies of categorical perception, Bailey and Summerfield (1980) demonstrated that the different acoustic properties representing a phoneme 'trade' with each other. The distinction between the /s/-vowel sound such as occurs in the word 'sir' and the /s/-stop consonant-vowel sound such as occurs in the word 'spur' depends on the integration of temporal information (the length of the closure gap when there is an absence of sound following /s/) and spectral information (the change in frequency at the onset of the voiced part of the 'ir/'ur' syllable which is known as the first formant transition). A shorter closure gap is needed in order to hear the /p/ when the rise in frequency of the first formant transition is larger.

To test whether cues to the gender of a face also 'trade' with each other two versions of the continuum of faces, one with male brows and one with female brows, were used. If 'trading relations' do operate then a categorisation task would reveal a shift in the boundary between male and female categories depending whether faces had male brows or female ones. Brows have been shown to exert a significant influence on perceived gender (Experiment 2) so it was reasoned that they should effect a significant shift in the category boundary. A face which had male brows and which fell at one point on the continuum from male jaw to female jaw should be more readily categorised as male than a face which fell at the same point on the continuum but which had female brows.

Method.

If the effects of jaw and brows as cues to gender were to be ascertained it would be necessary to keep the rest of the face as neutral as possible with regard to its gender. New stimuli were made in which all other features of the face except the jaw and the brows (ie eyes, nose, mouth, outline shape and size) were of indeterminate gender. The male and female prototype faces were averaged to create an androgynous face which was then used as the basic face in which male and female versions of the jaw and brows were manipulated.

Another four stimuli were employed as controls. These were the blank shapes of the four jaws (the area of the face below a horizontal line just under the base of the nose) in which no evidence of facial contours (eg mouth, chin or cheek shape) remained, but only overall dimensions and outline shape. This meant that evidence such as ratio of upper lip depth to chin depth or shading of the chin, which was present in the whole faces, was absent from the control stimuli and could not be used by subjects when classifying stimuli or discriminating between them. By examining the way subjects classified and discriminated between these stimuli and comparing that with the way they dealt with the whole faces it might be possible to better understand processes involved in face perception.

Stimuli.

The twelve stimuli used in the experiment are shown in Figure XII and were created by altering the male and female prototype faces which were used in Experiments 1 and 2. First an androgynous face was formed by averaging the two prototypes in the same way that the prototypes themselves had been created. Then, by means of the Pluto Graphics system, the male face was reduced and the female face was enlarged so that they were the same width at the level of the base of the nose as the androgynous face. Two versions of that face were then made, one by substituting the eyebrows from the reduced male prototype face, and the other by substituting the eyebrows from the enlarged female prototype.

Further manipulations of those composite faces were made by substituting a male jaw for the androgynous one, but retaining the androgynous mouth. Face 1m, thus created, had a male jaw and male brows and face 1f had a male jaw and female brows. Faces 4m and 4f were formed in a similar way by substituting a female jaw but retaining the androgynous mouth, so that face 4m had a female jaw and male brows and face 4f had a female jaw and female brows.

Using the same method as in Experiment 1 face 2m was made by blending 66.7% of face 1m with 33.3% of face 4m, and face 3m by blending 33.3% of face 1m with 66.7% of face 4m. Face 2f was made by blending 66.7% of face 1f with 33.3% of face 4f, and face 3f by blending 33.3% of face 1f with 66.7% of face 4f.

Faces 1m-4m were now used to create stimuli 1j-4j. The portion of

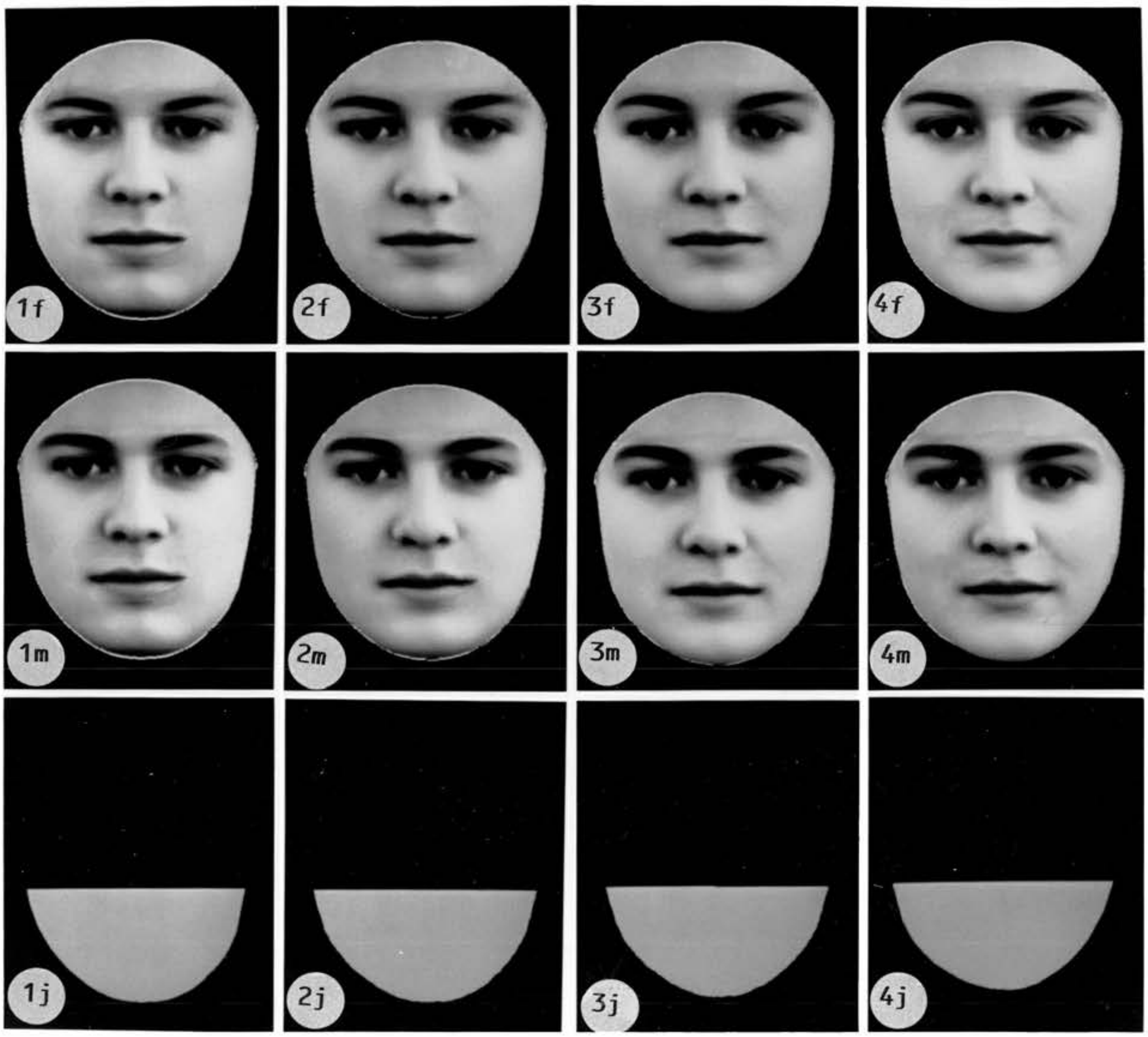


Figure XII. Stimuli used in Experiment 3.
Each continuum of stimuli varies from one with the most male jaw (stimulus 1) to one with the most female jaw (stimulus 4). Faces 1m-4m have male brows. Faces 1f-4f have female brows.

each of the four faces from the base of the nose upwards was masked and the remaining image of the jaw was rendered a uniform grey colour by means of the Pluto Graphics system. Only the shapes of the four jaws with no evidence of the mouth or of any shading of contours remained, to be used as control stimuli.

There were now three continua of four stimuli each.

1. The androgynous face with male brows changing from face 1m with a male jaw to face 4m with a female jaw.

2. The androgynous face with female brows changing from face 1f with a male jaw to face 4f with a female jaw.

3. The jaw-shapes changing from stimulus 1j the male jaw-shape to stimulus 4j the female jaw-shape.

Subjects.

Subjects were 8 male and 8 female students aged 19-26.

Design.

A mixed design was used. All subjects saw the stimuli with male and with female brows and each subject performed four tasks:

1. Discrimination between pairs of faces.
2. Discrimination between pairs of jaw shapes.
3. Categorisation of faces as male or female.
4. Categorisation of jaw shapes as male or female.

Half the male and half the female subjects viewed faces before jaw-shapes (ie order of tasks 1,2,3,4) and the other half viewed the jaw-shapes before the faces (ie order of tasks 2,1,4,3).

Procedure.

The same test procedure as was used in Experiment 1 was repeated to find out what difference, if any, the manipulations of the salient cues to gender had effected, but each subject performed four tasks instead of two. The two discrimination tasks were always performed before the categorisation tasks so that the strategy chosen for discrimination could not be influenced by that used for categorisation. After the two discrimination tasks and again after the categorisation tasks subjects were asked what strategy they had employed and notes were made of their answers.

Discrimination tasks.

For discrimination of pairs of faces four different blocks of twelve pairs were programmed, including in each block some pairs with male brows and some pairs with female brows. Each of the blocks contained six pairs which were the same and six which were different. In half of the different pairs the most male of the two appeared first and in the other half the most female appeared first. Each of the eight same pairs (1m/1m,1f/1f,2m/2m,3f/3f,etc) was included three times and each of the six different pairs (1m/2m,2m/3m,3m/4m, 1f/2f,2f/3f,3f/4f) was included four times altogether.

For discrimination of pairs of jaw-shapes, four different randomisations of six same pairs and six different pairs were programmed. Each of the same pairs appeared six times and each of the different pairs appeared eight times so that the total number of times each jaw-shape was seen in isolation was the same as the number of times it was seen as part of a face.

Categorisation tasks

For the categorisation of faces two different randomisations of the eight faces (four with male brows and four with female brows) were programmed. For the categorisation of jaw-shapes four different randomisations of the four shapes were programmed. This ensured that the number of times a particular jaw-shape was categorised when it was viewed as part of a whole face and the number of times it was categorised when it was viewed in isolation was the same.

Results.

Discrimination.

A 2 (gender of subject) x2 (order of presentation) x2 (type of brows) x7 (pair of faces) mixed ANOVA of the mean percent correct responses in discrimination between pairs of faces showed that the main effect of gender of subject was not significant, $F(1,24)=0.16$, $p=0.69$. The main effect of order of presentation of faces and jaw-shapes was not significant, $F(1,24)=0.64$, $p=0.43$, the main effect of type of brows was not significant, $F(1,24)=0.01$, $p=0.92$, and no interaction was significant. Responses under all conditions are therefore combined and shown in Figure XIII and there is clearly no peak of discrimination. Students t-tests of the

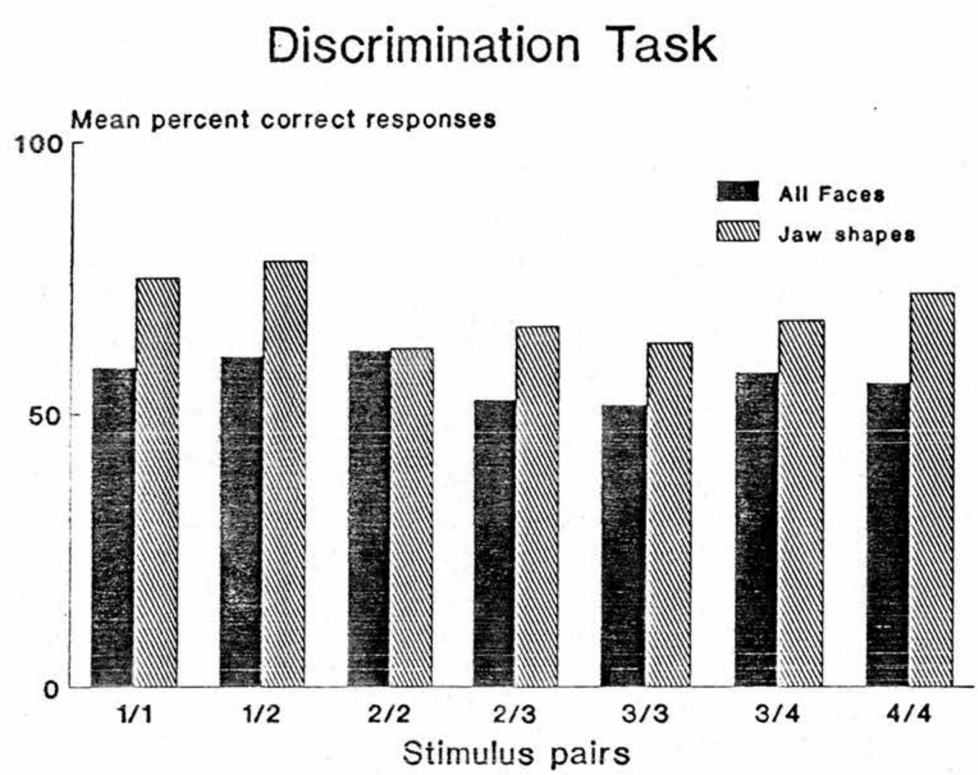


Figure XIII. Mean percent correct response in the discrimination task for pairs of faces and pairs of jaw-shapes.

results for each of the pairs against a hypothesised mean of 50% showed the combined results for pairs 2m/2m and 2f/2f and for pairs 1m/2m and 1f/2f (which have the same jaw but different brows) to have been discriminated at a level significantly above chance ($p<0.05$). The remaining pairs were discriminated at chance level. The difference between discrimination of same pairs and different pairs was not significant, $t=0.02$, $p=0.98$).

Performance in discrimination between pairs of jaw-shapes is also illustrated in Figure XIII. Students t-tests against a hypothesised mean of 50% showed that all pairs were discriminated at a level which was significantly above chance ($p<0.05$ in all cases) except for pairs 2j/2j and 3j/3j which were discriminated at chance level.

A within subjects 2 (faces or jaw-shapes) x7 (pairs) ANOVA showed that jaw-shapes were discriminated significantly better than faces, $F(1,15)=17.07$, $p=0.001$. The main effect of pairs of stimuli was not significant, $F(6,90)=1.34$, $p=0.25$, no pair being discriminated better than any other. The interaction between face or jaw-shape and pair of stimuli was not significant, $F(6,90)=0.63$, $p=0.7$.

When subjects were questioned about the strategy they used to discriminate between pairs of faces none of them said they had discriminated on the basis of the perceived gender of the face. Seven said they had compared the overall shape of the faces (the correct cue to differences), three mentioning the shape of the jaw specifically, but their performance was no better than other subjects'. Other replies included six mentions of eyes, brows or upper face, four mentions of lips or mouth, three 'don't knows', two mentions of expression and one of colour of skin.

When asked how they discriminated between the jaw-shapes there were twelve mentions of the curvature or symmetry of the outline shape, three 'don't knows' and one subject who had fixed his gaze on a speck of dust on the screen and watched for any movement of the stimulus in relation to that. No subject said s/he had discriminated on the basis of size.

Categorisation.

The mean percent 'male' responses in the categorisation of faces are plotted in Figure XIV. A 2 (gender of subject) x2 (order of presentation) x2 (type of brows) x4 (position of face on male/female scale)) ANOVA showed that the main effect of gender of subject was not significant and nor was the main effect of order of presentation of stimuli. The main effect of type of brows was highly significant, $F(1,12)=34.84$, $p=0.0001$, those faces with male brows being categorised as male more often than those with female brows. The main effect of position of a face on the scale from male jaw to female jaw was highly significant, $F(3,36)=63.08$, $p<0.0001$, the faces with the more masculine jaw being categorised as male more frequently than those with the more feminine jaw. No interaction was significant.

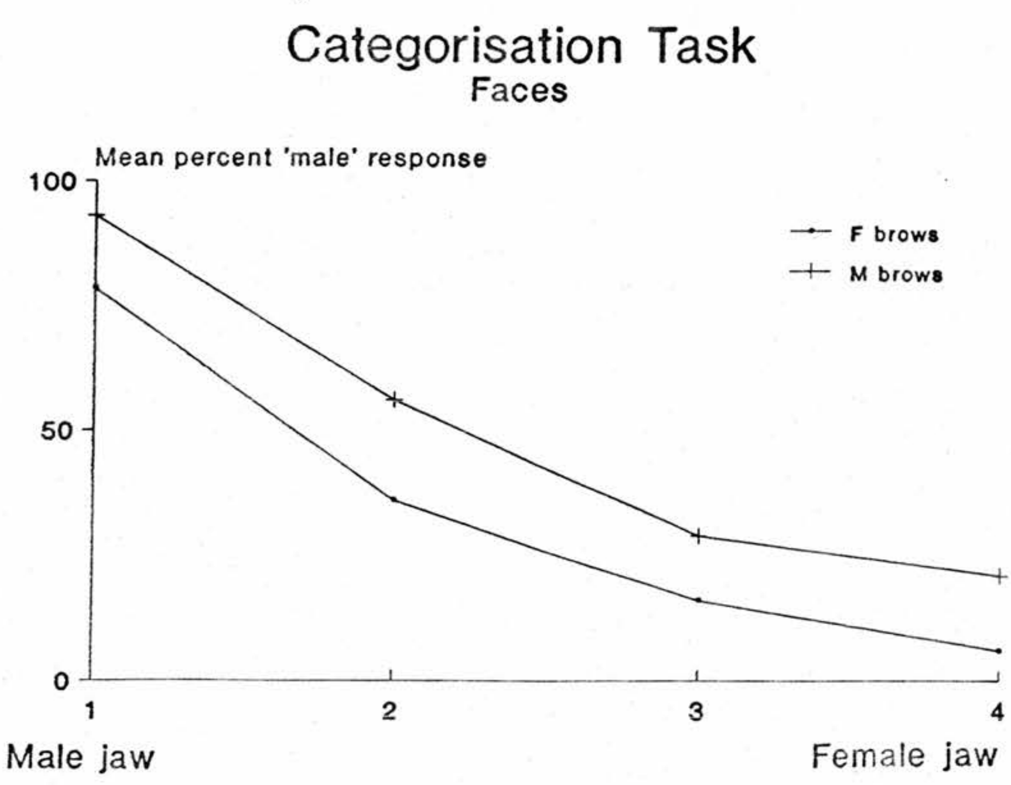


Figure XIV. Mean percent 'male' response in the categorisation task for faces with male brows and faces with female brows as a function of the position of the faces on the scale from male jaw to female jaw.

The difference between categorisation of faces and of jaw-shapes as male is illustrated in Figure XV. A 2 (faces/jaws) x4 (position of stimulus on male/female scale) ANOVA showed a significant difference between categorisation of faces and of jaw-shapes, $F(1,6)=11.8$, $p=0.004$. The main effect of the position of the image on the male/female scale was significant, $F(3,6)=25.6$, $p<0.001$ and the interaction between type of image and its position on the scale was significant, $F(3,6)=9.1$, $p<0.001$. A 1 (jaw-shape) x4(position on male/female scale) ANOVA showed the difference in categorisation of the different jaw-shapes to be approaching significance, $F=2.7$, $p=0.059$.

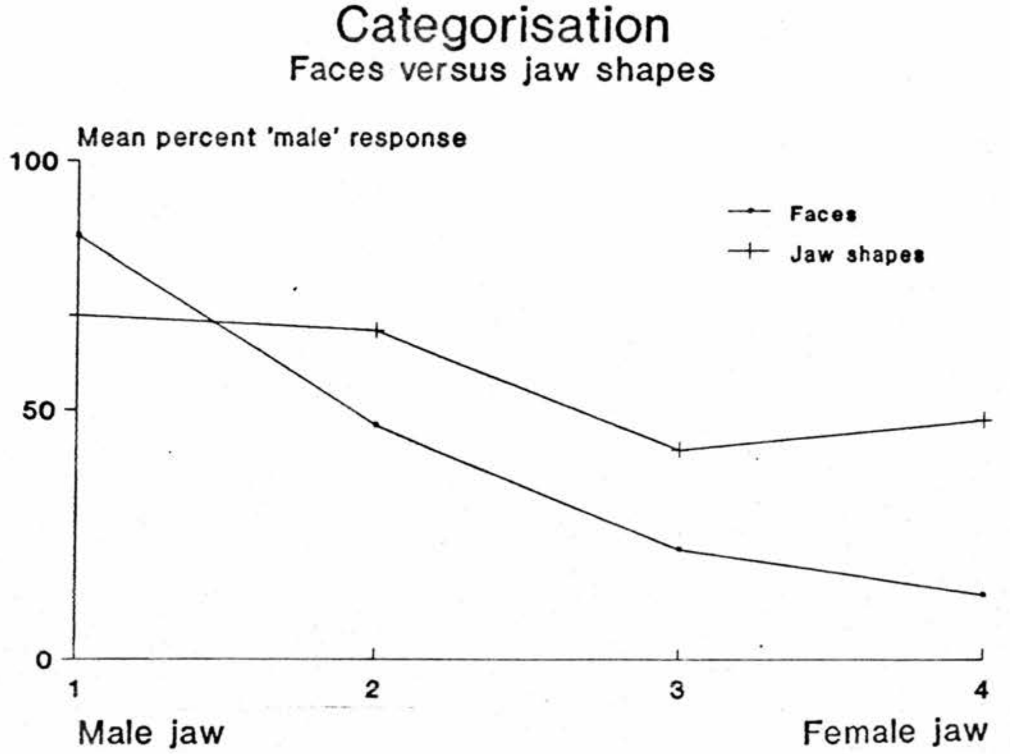


Figure XV. Mean percent 'male' response in the categorisation of faces and jaw-shapes as a function of the position of the stimulus on the scale from male jaw to female jaw.

Answers when subjects were questioned about the strategy they used for categorisation of faces included six mentions of the jaw area, four references to size (eg male faces were classed as broader, heavier, longer or as having a bigger chin). There were four references to shape, female faces being classed as rounder or smoother, five mentions of the eye area with 'narrower' eyes being classed as male. Three subjects referred to lips, one to the nose and there were two 'don't knows'.

When describing their strategy for categorising jaw-shapes subjects referred to size seven times and shape nine times, male jaw-shapes being considered to be larger and more angular or less symmetrical and there were four 'don't know's.

Discussion.

Discrimination.

Again, as in Experiment 1, the discrimination function for face pairs shows no peak at any point along the continuum. Again subjects did not discriminate on the basis of gender but attempted to do so on the basis of individual details or properties of the faces and jaw-shapes. The ease of discrimination between faces along the continuum was continuous both when several cues to gender were varied in Experiment 1 and when only the jaw was varied in Experiment 3. It was reduced to the level of chance when jaw-shape was the only difference between stimuli, in other words, subjects now found it impossible.

Discrimination of jaw-shapes was above the level of chance but there was no peak of discrimination. Subjects said they noticed the asymmetry in the curvature of the jaw shape which they had not reported noticing in the faces although it was present in those too. Changes in the outline of the jaw-shape were the only possible differences to look for because all other evidence had been removed. That also meant there was less visual noise in the stimuli making processing simpler. The area to be examined was also smaller than in the face stimuli so that subjects had time to examine it more thoroughly. Even so, performance was far from perfect, suggesting that the shapes were difficult to encode accurately in memory for comparison with the next shape.

Five of the seven subjects who said they had discriminated between faces on the basis of overall shape had seen the jaw-shapes first suggesting that their attention had been drawn to outline shape. However, performance of subjects who had seen the jaw-shapes first was not significantly better than that of subjects who had seen the faces first. Having attention drawn to outline shape (if that is what happened) does not seem to have helped discrimination. That may be because when the jaw-shape is part of a larger face shape the difference is relatively smaller and therefore more difficult to discriminate. This would accord with Weber's Law.

Categorisation.

As in Experiment 1 the eight faces were categorised by gender without difficulty after presentation for only one second. The categorisation function in Figure XIV shows no sharp division between male and female occurring at one point on the continuum, but rather a gradual change from one to the other indicated by the gently sloping curve of the graph which occurs whether the faces have male or female brows. The difference between this gently sloping curve and the sharp drop in the identification function of Experiment 1 (FigureII) suggests that the evidence which subjects need to make a clear-cut categorisation has been removed from the stimuli used in the present experiment. Experiment 2 showed that each aspect or feature of a face gives some information about gender and the results of the present experiment could mean that all of that evidence is necessary for a clear-cut classification to be made. It is also possible that one aspect of a face - perhaps its overall configuration - is crucially important and that that aspect is not represented adequately in these modified stimuli. Another possibility is that the different subjects used for this experiment use different criteria in their judgment of gender. In Experiment 2 wide variation was shown between individual subjects' classification of faces according to their perceived gender, and the difference was highly significant between male and female subjects. Experiment 1 used 9 male and 16 female subjects for the categorisation task and Experiment 3 used 8 males and 8 females. The imbalance between the numbers of male and female subjects in Experiment 1 could have biased the results of the categorisation task. There seems to be no clear demarcation for the subjects in Experiment 3 between a male jaw and a female one but rather a graded change in perceived masculinity/femininity along the continuum.

The difference between the identification curve when the faces had male brows and that when they had female brows indicates that 'trading relations' do operate between the cues to the gender of a face. This suggests that multiple aspects of the available evidence concerning the gender of a face are weighed in the balance when it is categorised as male or female. This is in contrast to discrimination which can be judged on the basis of only one aspect of the available information.

General discussion.

Experiments 1 and 3 have not shown categorical perception of the gender of a face as judged by the definition specified in the introduction to Chapter One. Only the results of the identification task in Experiment 1 are typical of the results obtained when categorical perception is demonstrated and it is clear that gender is not perceived categorically.

There may be some reason that a particular facial dimension is not likely to be perceived in a categorical manner. From facial appearance alone there can be a degree of uncertainty about the gender of some individuals. This may explain the pressure from society for individuals to adopt distinctive hairstyles and clothing to make gender less ambiguous. However, we have many common labels for male and female gender but few labels and none in common use for an ambiguous (androgynous) gender. Thus there is no doubt that we are aware at a cognitive level of two gender categories even if at a perceptual level the masculine/feminine dimension is not wholly categorical.

Medin and Barsalou (1987) make a distinction between sensory perception categories (eg speech sounds) and generic-knowledge categories (eg natural kinds [birds], artifacts [cars] and events [going to restaurants]). They claim that most investigators who study generic knowledge categories implicitly believe that categorical perception occurs only for sensory perception categories. The results of the present experiments support that belief. Gender would seem to belong amongst generic knowledge categories which may not be perceived categorically.

It could be that categorical perception of complex stimuli occurs spontaneously when there is a need to process the perceived signals very rapidly. Categorical perception has been demonstrated in perception of speech sounds and of musical stimuli (when the listeners are musicians). There is some evidence that it may occur in perception of facial expression (Etcoff & Magee, 1992). There is variation in the utterances of the same phoneme by an individual speaker and considerable variation across speakers and yet they are perceived as being the same by all listeners. Listeners have to classify speech sounds rapidly because the sounds are present only momentarily. Similarly musicians who wish to follow music in terms of their specialised knowledge have to do so very rapidly. Different keys and melodic intervals are perceived in the same way although the music may be played on different instruments. Non-musicians are not listening to music with the same kind of understanding and do not need to perceive and interpret its musical categories. Facial expressions are also transient in nature (Eible-Eibesfeldt, 1973) and must be interpreted rapidly during social interchange if they are to have any meaning for the observer. There is ample evidence that expressions are perceived in the same way when emitted by different individuals (Ekman et al 1972). If it is the case that categorical perception occurs when there is a need to perceive changes in a signal very rapidly one would not expect the gender of a face to be perceived categorically. Gender does not alter so there is no pressure for rapid perceptual categorisation. The results of Experiment 2 suggest that the gender of a face is not categorised by all observers in the same way. Different individuals, particularly if they are of opposite gender, have differing concepts of male and female characteristics. In spite of that there must be considerable consensus about gender, when enough evidence is available. Otherwise a sharp category boundary would not have been demonstrated in Experiment 1.

A second purpose of this investigation was to establish whether the paradigms used in investigation of speech perception could be usefully adapted for the investigation of face perception. The identification paradigm has been shown to illustrate differences in categorisation of stimuli after they have been modified. A series of modifications to the stimuli could have been made to investigate perception of gender further. In the same way that jaw and brows were traded with each other in Experiment 3 other features could have been changed in the androgynous face adding evidence of gender little by little and creating new continua of faces. After all the individual features had been changed so that the whole face now changed from male to female the difference in size of male and female faces could then be restored to find out what effect that had. There might be a graduation towards a sharp category boundary or the identification function might suddenly change from a smooth curve to one with a sharp boundary. A graduation would indicate that categorisation of gender was a matter of summing the evidence from all the features. A sudden change would indicate that the most recently changed feature or aspect of the face was the crucial one signifying gender. It would be necessary to have the same subjects categorise all the continua of faces to avoid the possibility of variation from one stage to the next in the criteria used by subjects for categorisation.

Discrimination has been shown to be based on one detail which varies between stimuli. The discrimination paradigm would therefore be useful only when the critical difference between two versions of a detail was to be ascertained. For example, the question to be answered might be, is there a critical difference between thick lips and thin ones or can we discriminate continuously all variations in thickness of lips?

It is concluded from the results of this study that gender is not perceived categorically. It is also concluded that the identification and discrimination paradigms borrowed from investigation of speech perception could be useful in further investigation of perception of faces.

References.

Abramson A S, Lisker L. (1970). Discriminability along the voicing continuum: Cross-language tests. Proceedings of the 6th International Congress of Phonetic Sciences (pp 569-573). Prague: Academia.

Bailey P J, Summerfield Q. (1980) Information in speech: Observations on the perception of /s/-stop clusters. Journal of Experimental Psychology. Human Perception and Performance. 6, 536-563.

Benson P J, Perrett D I. (1991). Computer Averaging and Manipulation of Faces. In P. Wombell (Ed). PhotoVideo: Photography in the Age of the Computer. London: Rivers Oram Press.

Benson P J, Perrett D I. (1992). Extracting prototypical facial images from exemplars. Perception (in press).

Best C T, Morrongiello B, Robson R. (1981). Perceptual equivalence of acoustic cues in speech and nonspeech perception. Perception and Psychophysics, 29, 191-211.

Birdwhistell R L. (1970) Masculinity and femininity as display. In Kinesics and Context: Essays on Body Motion Communication. Philadelphia: University of Pennsylvania Press.

Brigham J C, Barkowitz P. (1978). Do "They all look alike"? The effect of race, sex, experience and attitudes on the ability to recognise faces. Journal of Applied Social Psychology, 8, 306-318.

Brown E. (1990). Trading relations between visual cues in categorisation of drawings. Unpublished BSc Thesis: University of York.

Burnham D K, Earnshaw L J, Clark J E. (1991). Development of categorical identification of native and non-native bilabial stops: infants, children, and adults. Journal of Child language, 18, 231-260.

Burns E M, Ward W D. (1978). Categorical perception - phenomenon or epiphenomenon: Evidence from experiments in the perception of melodic musical intervals. Journal of the Acoustical Society of America, 63, 456-468.

Chance J, Goldstein A G, McBride L. (1975). Differential experience and recognition memory for faces. Journal of Social Psychology, 97, 243-253.

Davies G, Ellis H D, Shepherd J. (1977). Cue saliency in faces as assessed by the 'Photofit' technique. Perception, 6, 263-269.

Dooling R J, Brown S D. (1990). Speech perception by budgerigars (Melopsittacus undulatus): Spoken vowels. Perception & Psychophysics, 47, 568-574.

Dooling R J, Okanoya K, Brown S D. (1989). Speech perception by budgerigars (Melopsittacus undulatus): The Voiced-Voiceless Distinction. Perception & Psychophysics, 46, 65-71.

Eible-Eibesfeldt I. (1972). Similarities and differences between cultures in expressive movements. In R Hinde(Ed), Nonverbal Communication. Cambridge: Cambridge University Press.

Eimas P D, Siqueland E R, Jusczyk P, Vigorito J. (1971). Speech perception in infants. Science, 171, 303-306.

Ekman P, Friesen M V, Ellsworth P. (1972). Emotion in the human face. New York: Pergamon Press.

Enlow D. (1982). Handbook of Facial Growth. Philadelphia: W.H.Saunders.

Etcoff N L, Magee J J. (1992). Categorical perception of facial expressions. Cognition, 44, 227-240.

Galper R E. (1973). 'Functional race membership' and recognition of faces. Perception and Motor Skills, 37, 455-462.

Galton F. (1878). Composite portraits. Journal of the Anthropological Institute of Great Britain and Ireland, 8, 132-142.

Galton F. (1883). Inquiries into Human Faculty and its Development. London: J M Dent & Co.

Grusser O-J, Landis T. (1992). Visual Agnosias and Other Disturbances of Visual Perception and Cognition. Boston: CRC Press.

Haig N D. (1984). The effect of feature displacement on face recognition. Perception, 13, 505-512.

Haig N D. (1986). Exploring recognition with interchanged facial features. Perception, 15, 235-247.

Harmon L D. (1973). The recognition of faces. Scientific American, 227, 71-82.

Harnad S (1987). Categorical Perception. Cambridge: Cambridge University Press.

Hary J M, Massaro D W. (1982). Categorical results do not imply categorical perception. Perception & Psychophysics, 32, 409-418.

Katz D. (Ed.) (1953). Studien zur Experimentellen Psychologie. Basel: Schwabe.

Kuhl P K, Miller J D. (1975). Speech perception in the chinchilla: Voiced-voiceless distinction in alveolar plosive consonants. Science, 190, 69-72.

Kuhl P K, Miller J D. (1978). Speech perception by the chinchilla: Identification functions for thew synthetic VOT stimuli. Journal of the Acoustical Society of America, 63, 905-917.

Kuhl P K, Padden D M. (1982). Enhanced discriminability at the phonetic boundaries for the voicing feature in Macaques. Perception & Psychophysics, 32, 542-550.

Lane H. (1965). Motor theory of speech perception: A critical review. Psychological Review, 72, 275-309.

Liberman A M. (1970). The grammars of speech and language. Cognitive Psychology, 1, 301-323.

Liberman A M, Cooper F S, Shankweiler D S,Studdert-Kennedy M. (1967). Perception of the speech code. Psychological Review, 74, 431-461.

Liberman A M, Harris K S, Hoffman H S, Griffith B C. (1957). The discrimination of speech sounds within and across phoneme boundaries. Journal of Experimental Psychology, 54, 358-368.

Liggett J. (1974). The Human Face. London: Constable.

Lisker L. (1978). Rapid *vs* Rabid: A catalogue of acoustic features that may cue the distinction. Haskins Laboratories Status Report on Speech Research (SR-54,pp. 127-132).

Lisker L, Abramson A S. (1967). Some effects of context on voice onset time in English stops. Language and Speech, 10, 1-28.

Lisker L, Abramson A S. (1970). The voicing dimension: Some experiments in comparative phonetics. Proceedings of the 6th International Congress of Phonetic Sciences (pp 563-567). Prague: Academia.

Locke S, Keller L. (1973). Categorical perception in a non-linguistic mode. Cortex, 9, 355-369.

Medin D L, Barsalou W. (1987). Categorization processes and categorical perception. In S.Harnad (Ed), Categorical perception. New York: Cambridge University Press.

Miller J D, Wier C C, Pastore R E, Kelly W J, Dooling R J. (1976). Discrimination and labeling of noise-buzz sequences with varying noise-lead times: An example of categorical perception. Journal of the Acoustical society of America, 60, 410-417.

Miyawaki K, Strange W, Verbrugge R, Liberman A M, Jenkins J J, Fujimura O. (1975). An effect of linguistic experience: The discrimination of /r/ and /l/ by native speakers of Japanese and English. Perception and Psychophysics, 8, 331-340.

Moss J P, Linney A D, Grindrod S R, Arridge S R, Clifton J S. (1987). Three-dimensional visualisation of the face and skull using computerised tomography and laser scanning techniques. European Journal of Orthodontics, 9, 247-253.

Nakdimen K A. (1984). The physiognomic basis of sexual stereotyping. American Journal of Psychiatry, 141, 499-503.

Pastore R E. (1976). Categorical perception: a critical re-evaluation. In S.K. Hirsh, D.H. Eldredge, I.J. Hirsh, & S.R. Silverman (Eds.). Hearing and Davis: Essays honoring Hallowell Davis (pp. 253-264). St. Louis: Washington University Press.

Pastore R E, Ahroon W A, Baffuto K J, Friedman C, Puleo J S, Fink E A.(1977). Common-factor model of categorical perception. Journal of Experimental Psychology, Human Perception & Performance, 3, 686-696.

Pastore R E, Schmuckler M A, Rosenblum L, Szczesiul R. (1983). Duplex perception with musical stimuli. Perception & Psychophysics, 33, 469-474.

Pastore R E, Li X-F, Layer J K. (1990). Categorical perception of nonspeech chirps and bleats. Perception & Psychophysics, 48, 151-156.

Pisoni D B. (1973). Auditory and phonetic memory codes in the discrimination of consonants and vowels. Perception and Psychophysics, 13, 253-260.

Pisoni D B, Aslin R N, Perey A J, Hennessy B L. (1982). Some effects of laboratory training on identification and discrimination of voicing contrasts in stop consonants. Journal of Experimental Psychology, Human Perception and Performance, 8, 297-314.

Roberts T, Bruce V. (1988). Feature saliency in judging the sex and familiarity of faces. Perception, 17, 475-481.

Sergent J. (1986). Microgenesis of face perception. In Aspects of Face Processing. H.D. Ellis, M.A.Jeeves, F. Newcombe, & A. Young (Eds). Dordrecht. Martinus Nijhoff.

Stevens K N. (1981). Constraints imposed by the auditory system on the properties used to classify speech sounds: Data from phonology, acoustics and psycho-acoustics. In T.F. Myers, J. Laver,& J.Anderson (Eds.), The Cognitive Representation of Speech. Amsterdam: North Holland.

Werker J F, Tees R C. (1984). Cross-language speech perception: Evidence for perceptual reorganisation during the first year of life. Infant Behavior and Development, 7, 49-63.