DIGITAL MANIPULATION OF FACES AND ITS CONSEQUENCE UPON IDENTIFICATION AND ATTRACTIVENESS

Kieran J. Lee

A Thesis Submitted for the Degree of PhD
at the
University of St Andrews

1999

Digital Manipulation of Faces and its Consequence Upon Identification and Attractiveness

Kieran J. Lee

School of Psychology
University of St. Andrews

Ph.D. Thesis
September 1997

Declaration

i) I, Kieran Lee, hereby certify that this thesis, which is approximately 55000 words in length, has been written by me, that it is the record of work carried out by me and that it has not been submitted in any previous application for a higher degree.

Date: 30/9/97 Signature of candidate:

ii) I was admitted as a research student in December 1994 and as a candidate for the degree of Ph.D. in December 1994; the higher study for which this is a record was carried out in the University of St Andrews between December 1994 and August 1997.

Date: 30/9/97 Signature of candidate:

iii) I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of Ph.D. in the University of St Andrews and that the candidate is qualified to submit this thesis in application for that degree.

Date: 30/9/97 Signature of supervisor:

Copyright

Unrestricted

Date: 30/9/97 Signature of Candidate:

Acknowledgements

Many thanks to my supervisor, Dave Perrett, for his enthusiasm, criticism and encouragement. Thanks also to the other member of the perception lab - Duncan, Ian, Lindsey and Michael - and the various honours students who have studied there over the years for helping make the lab such a stimulating environment to work in.

I would like to acknowledge Duncan Rowland for the writing of the C program used to present the stimuli to subjects in the experiments described here and for the development of software tools enabling the manipulation of faces in colour space, Mike Oram for his statistical advice and Ian Penton-Voak for a crash-course in SPSS.

Thank you to Peter Henzi and Sakiko Yoshikawa for their assistance in the testing of subjects on a cross-cultural basis.

Huge thanks to my good friends - Jack K., Chris K., Liz M., Penny G., Johnny C., Jonathan L., Ben D., Heather P., Dave O., Sarah A., James C., Charles H., Kitty D-H, Sarah M., Dave T., Amy K. and Megan - my brother and parents and whoever else has just been there and put up with my moments of insanity; taking some of the dread out of my dreaded last year. And a special mention to f9314sk for the highest and lowest 30 months of my existence.

Table of Contents

Abstract

The thesis investigated perception of facial identity and attractiveness with digital manipulation of facial traits. A presentation time paradigm was developed by which stimuli could be presented for a range of brief display periods. Using this paradigm subjects recognised photo-realistic target faces caricatured in shape with greater accuracy than veridical images consistent with previous findings using reaction time as a measure. Subjects were further required to identify colour representations of famous faces which were either veridical, caricatured in colour space or had enhanced colour saturation and intensity contrast (as contrast controls). Recognition accuracy was greater when viewing the colour-caricatured stimuli than either the veridical images or the contrast controls. The removal of colour to produce grey-scale images also decreased accuracy of face recognition indicating that colour information aids facial identification. Caricaturing of faces, therefore, can be extended to the colour domain and, as with shape caricaturing, enhancement of distinctive information can produce a recognition advantage for famous faces. Subjects were also asked to identify the best-likeness for individuals using photo-realistic stimuli and an interactive paradigm with shape-caricature, colour-caricature and contrast-control varied by the user in real-time. The best-likeness with shape manipulation was a slight anti-caricature while with colour-caricature and contrast-control images a mildly exaggerated image was selected as the best-likeness. Thus although images caricatured substantially in colour or shape (+40%) induce superior recognition compared to veridical images such substantial exaggerations are not seen as best-likenesses under prolonged exposure. The gender typicality for white British and Japanese composite faces was manipulated and subjects presented with the images using both a static forced-choice paradigm and an interactive paradigm. Male and female face shapes with enhanced feminine features were consistently found to be more attractive than average. Preference for enhanced femininity for female faces was greater for subjects making within- than cross-cultural judgements.

Chapter 1 - Caricaturing and Recognition of Faces (an introduction)

The ability to distinguish between human faces is of particular importance for us. We are required to discriminate between dozens of faces every day and we are all face experts. Faces are a homogeneous class in that features always occur in the same basic configuration. This makes identification of faces an extremely difficult cognitive task. To discriminate one face from another we have somehow to access the features that differentiate it from a multitude of others.

In recent investigations into the cognitive processes underlying the discrimination and identification of faces, caricatures have proved to be an increasingly useful and informative class of stimuli. A caricature can be defined as a representation of a face in which the differences of that face from the norm (or average) have been exaggerated (Brennan 1982, 1985). A *caricature advantage* is the situation in which recognition for the face of an individual is facilitated relative to an accurate (veridical) or uncaricatured depiction by caricaturing that face to accentuate the differences it has from a relevant norm.

Background to studies using caricatures

There are several reasons one might expect a caricature advantage for faces. First, when presented with a face which has a distinctive feature and given instructions that draw attention to the feature, memory is facilitated for photographs of that face (Winograd, 1981). Caricatures may benefit recognition by drawing attention to distinctive features either during training or testing. Second, atypical or distinctive faces are remembered better than typical faces (Bartlett, Hurry and Thorley, 1984; Cohen and Carr, 1975; Davies, 1978; Going and Read, 1974; Light, Kayra-Stuart and Hollander, 1979; Nash, 1969; Shapiro and Penrod, 1986; Valentine and Bruce, 1986). This advantage presumably arises because distinctive faces are less easy to confuse with other faces and give rise to fewer false positives in recognition paradigms. Third, typical category members are often falsely reported as having been seen before in old/new matching tasks (Franks and Bransford, 1971; Solso and McCarthy, 1981). Thus typical faces without distinctiveness enhanced should be the most difficult to recognise.

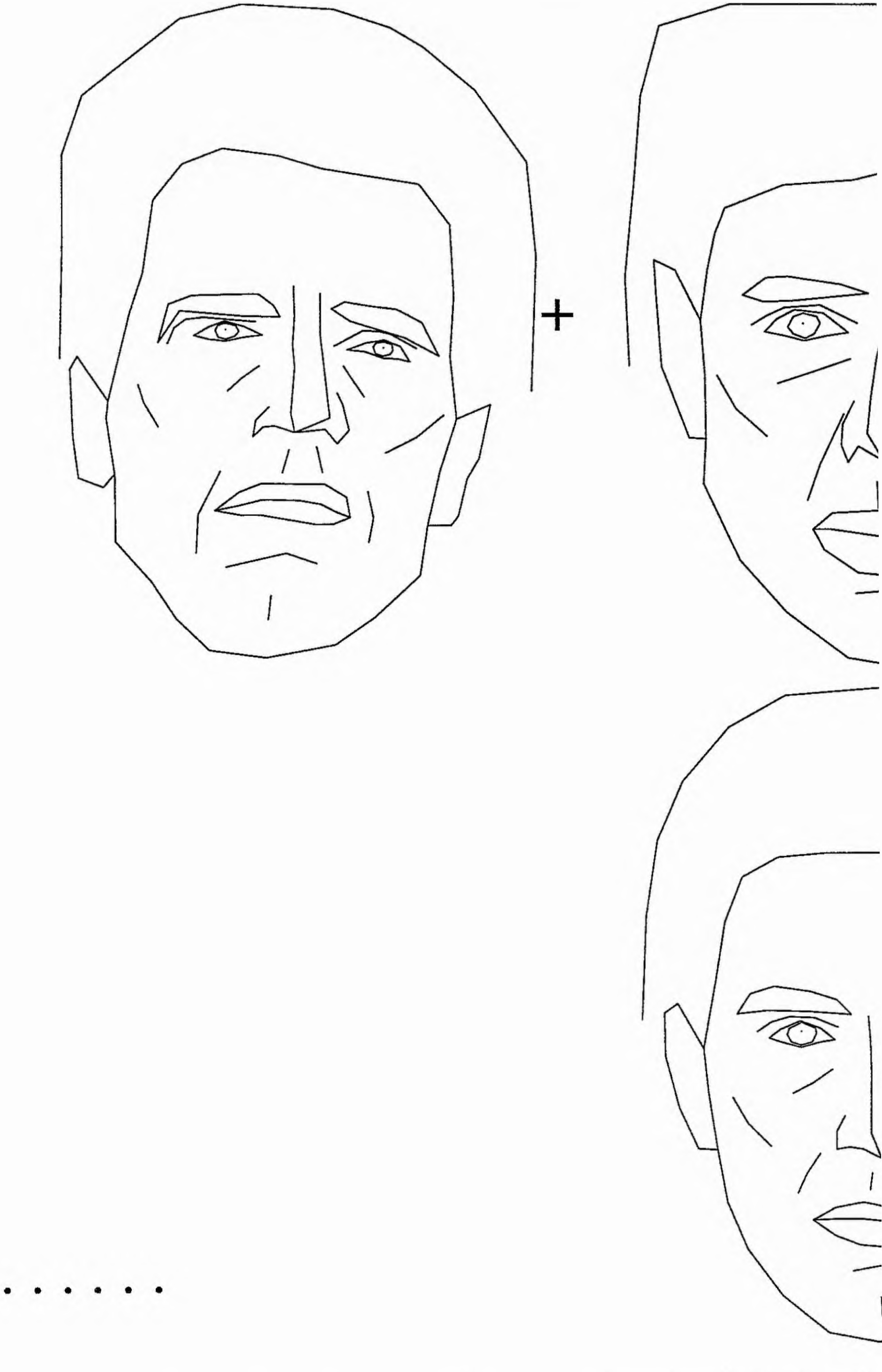
Light *et al* (1979) proposed a two-component theory of memory in order to account for the effect of distinctiveness. This theory is based on inter-stimulus similarity. They suggested that a distinctive face is more likely to access a specific memory because it is less similar to specific memories of other faces. Thus access to specific memory gives rise to an advantage in hit rate for distinctive faces. In the absence of a specific memory, Light *et al* suggested that subjects base their recognition judgement on schematic memory of category structure (similar to a prototype). Use of schematic memory gives rise to the greater false positive rate found for typical faces. In contrast, Bartlett *et al* (1984) interpreted the effects of distinctiveness in terms of familiarity information alone. Bartlett *et al* suggested that presentation of a distinctive face resulted in a greater increment in familiarity than that resulting from presentation of a typical face. Valentine and Bruce (1986a) argued that if the latency to recognise familiar faces arises from the role of a facial prototype, then distinctive faces should take longer than typical faces to be classified as a face in a task in which faces must be distinguished from jumbled faces. Subsequent experiments showed that distinctive faces are recognised faster in a familiarity decision task but classed as faces more slowly than are typical faces.

These findings give rise to the prototype hypothesis (Valentine and Bruce, 1986, 1986a) or distinctiveness hypothesis (Rhodes, Brennan and Carey, 1987). These hypothesese are conceptually identical and assert that the distinctive aspects noted in a face can be differentiated from those in other faces and can be identified or recognised when seen again. Therefore, caricatures should facilitate recognition better than representations in which distinctive aspects are minimised (anticaricatures). The caricature hypothesis (Rhodes *et al*, 1987) can be interpreted as an extension of this distinctiveness hypothesis and implies some form of long-term encoding of the information which is atypical for a particular, familiar face. Hence, this hypothesis implies that caricatures of familiar faces should be recognised better than veridical (undistorted) images, giving a *caricature advantage*.

The skill of the caricature artist begins with realising which features are characteristic. Goldman and Hagen (1978) studied the caricaturing of Richard Nixon by 17 artists and found a high degree of concordance as to what was caricatured. Prior investigations into the effects of caricaturing have involved the caricaturing of shape information pertinent to unfamiliar or highly familiar faces or, in some cases, animals or objects (Benson and Perrett, 1991, 1994; Carey, 1992; Hagen and Perkins, 1983; Rhodes *et al*, 1987; Rhodes and McLean, 1990; Tversky and Baratz, 1985). Most of these investigations have been carried out using line-drawing representations of faces produced through a computer implemented caricature-generating process developed by Brennan (1982, 1985).

Computer-implemented caricaturing

Brennan's system involves forming a line-drawing representation of a digitally captured face. This involves an experienced operator delineating a face by marking a number of 'feature points' - 169 in the original Rhodes, Brennan and Carey (1987) study - and overlaying them on a particular face in a manner uniform to all images to be used (one point for the left corner of the mouth, etc). These points can then be connected in particular ways using spline curves in order to form the outline of a face (*Plate 1.1*). A norm or average face is formed by averaging the positions of corresponding points delineated for several images which have been scaled by mapping the pupil centres for all the constituent faces to the same co-ordinates (*Figure 1.1*). Thus the average shape of a group of faces can be calculated as the mean position of corresponding feature points across the group. The positions of points on a new face can then be compared with the



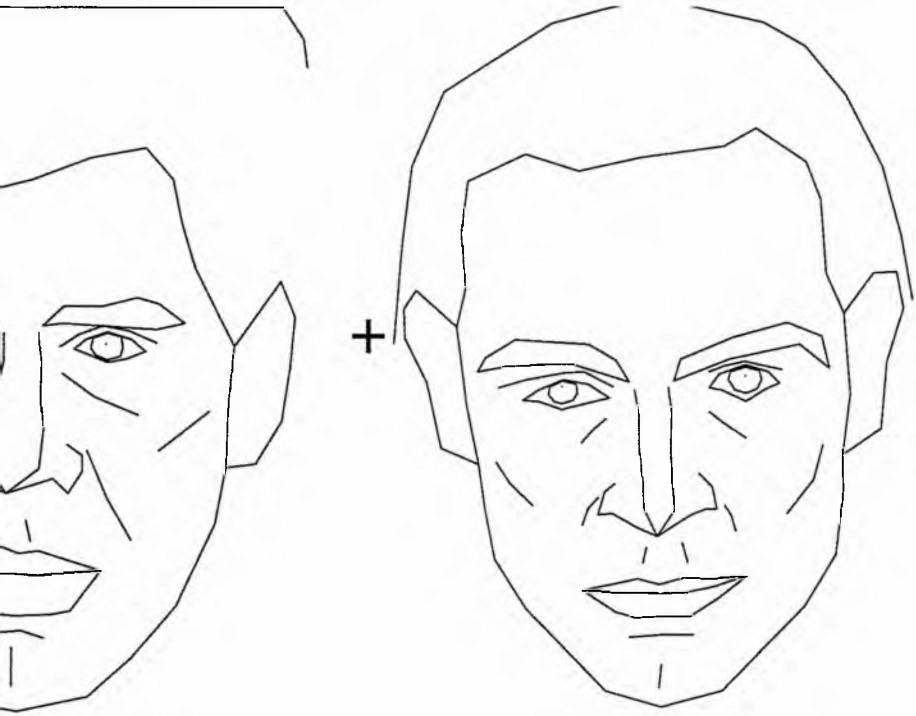


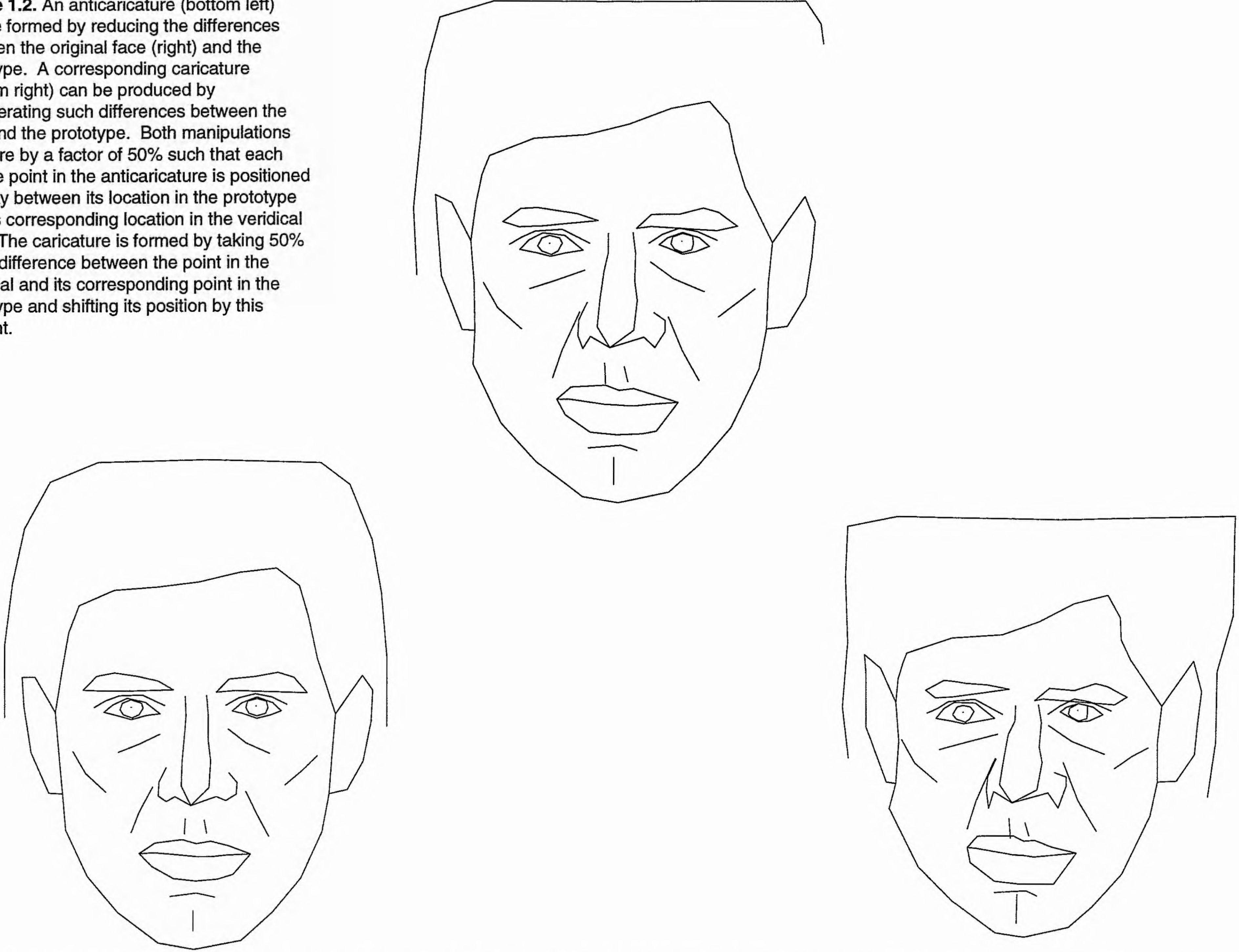

Figure 1.1. A composite line-drawing representation can be created by averaging the x,y coordinates for corresponding feature points across a set of faces (actors, in this case) which have been normalised for inter-ocular position.

norm by scaling the face and aligning the pupil centres. All the metric differences between a norm and the new face can then be exaggerated linearly by an amount specified by the user to produce a caricature of the target (*Figure 1.2*). Similarly, differences can be diminished to form an image which more closely resembles the norm - an anticaricature.

This technique is closely related to a trick used by animators, known as in-betweening, and its increasingly popular digital descendant "morphing". If a line-drawing of a familiar face formed in the way described and another drawing of the norm derived from several faces is produced, each point on the familiar face can then be paired with a respective point on the norm. If the paired points are connected by spline curves, the mid-points on the curves depict a new face that splits the difference between the norm face and the familiar face (an anticaricature). Repeating this process several times and animating between the images which have been produced provides the basis for a "morph" between the norm and target familiar face. If each curve, or line, is now extended beyond the familiar target face by half its original length and points are placed at the end of the lines, a caricature of the target face is produced when the points are connected. The degree of caricature is calculated by multiplying the differences between the horizontal co-ordinates and the vertical co-ordinates of the average face and target face by the desired factor. This caricature process exaggerates differences in feature positions relative to the interocular separation (which is standardised for all faces). Because the process is relative, all dimensions including interocular separation are evenly caricatured.

This caricaturing technique has allowed experimenters to generate stimuli that vary in a continuous and controlled way and maintain holistic cues to the face. This enables experimental procedure to be carried out more rigorously as many previous perceptual studies have often been constrained by the need to compare caricatured line-drawings with photographs or other kinds of pictorially dissimilar images (Hagen and Perkins, 1983; Tversky and Baratz, 1985) and were therefore confounded by representational effects. This is also important as the

Figure 1.2. An anticaricature (bottom left) can be formed by reducing the differences between the original face (right) and the prototype. A corresponding caricature (bottom right) can be produced by exaggerating such differences between the face and the prototype. Both manipulations here are by a factor of 50% such that each feature point in the anticaricature is positioned midway between its location in the prototype and its corresponding location in the veridical face. The caricature is formed by taking 50% of the difference between the point in the veridical and its corresponding point in the prototype and shifting its position by this amount.



configuration of features has a profound impact on facial appearance (Haig, 1984, 1986; Rhodes, 1988; Sergent, 1984; Shepherd, Ellis and Davies, 1977). However, the dimensions of faces which go to make up these configuration cues have not been established (Yamane, Kaji and Komatsu, 1988).

The caricature advantage - a function of expertise

Recognition experiments which use reaction time as a measure of response time to identify a familiar individual from these line drawings have supported the caricature hypothesis (Rhodes, Brennan and Carey, 1987; Carey, 1992; Benson and Perrett, 1994). Caricatures of highly familiar faces were recognised more quickly than veridical images.

Rhodes and McLean (1990) propose a configural encoding mechanism in which individual members of a class of objects with which a subject is highly familiar are recognised in terms of deviation from a norm which contains all the information which is particular to that class. This norm or prototype is envisaged as being derived from some process of averaging all the members of such a class (e.g. faces) which one has encountered in every day life. It is conceivable that each face that one encounters would update this prototype, envisaged as some mental representation of the central tendency for faces, and as a very large number of faces are seen, the prototype becomes increasing stable and represents the mean of the population of faces experienced. Note that such a prototype need not be stored in a literal or explicit manner within the nervous system (Benson and Perrett, 1991; Valentine, 1991). It follows that a caricature should be more likely to improve reaction time for recognition of a more typical individual in an object class which deviates little from the norm than for a distinctive object which already deviates highly from the norm (as found by Benson and Perrett, 1994). Rhodes and McLean reasoned that this form of encoding facial information, which they termed norm-based coding, leads to a caricature advantage over anticaricature or veridical images and may not be specific to face recognition. They suggest that this form of coding may provide a general model for representing spatial variations in any basic configuration and may be used whenever a homogeneous class (e.g. faces, dogs, birds) of objects with a high degree of structural similarity must be discriminated. Supporting this, Dodd and Perrett (unpublished) find a recognition advantage for caricatures of cars.

Recognition and caricature advantage as a function of expertise

Biederman's (1987) model of recognition proposes that objects are mentally represented in terms of their geometric components or parts. However, such analysis is inadequate for distinguishing stimuli that share a common configuration, because they all have the same parts in the same basic arrangement. Members of homogeneous classes present a different problem to the visual system than do objects that can be distinguished on the basis of differences in parts or the relations between parts, and may therefore require different representational structures (Pinker, 1985).

Following this reasoning, Rhodes and McLean (1990) argued that birds form a class whose members share a common configuration and should therefore be well represented by coding variations from a norm. Indeed, they found advantages for caricatures compared with anticaricatures for recognition of line drawings of birds by both novice and expert subjects and a caricature advantage over veridical line drawings only when subjects were experts (ornithologists). Recognition benefited from caricatures over veridical images only when the birds were taken from a highly homogeneous group, e.g. passerines. The size of this caricature advantage showed a strong positive correlation with typicality of the bird, supporting their proposition that exaggeration of distinctive information is most useful when the bird is a relatively typical member of its class. No advantage was found when highly distinctive members of the class were caricatured or when an inappropriate norm (a kiwi) was used as a basis for caricaturing.

Hence the distinctiveness hypothesis has consistently been supported for unfamiliar faces (or birds) and familiar faces, whereas a caricaturing advantage against veridical has been demonstrated exclusively when faces or other homogeneous objects are highly familiar to subjects (Benson and Perrett, 1994; May *et al*, submitted; Rhodes *et al*, 1987; Rhodes and McLean, 1990; Rhodes and Tremewan, 1994).

Ellis *et al* (1979) and Young *et al* (1985) have found familiar face recognition to depend more on internal features (e.g. eyes, nose, mouth) while for non-familiar faces it depends just as much upon external features such as hair and ears. Ellis *et al* thus speculated that different strategies are associated with the two classes of stimulus as the expressive elements of the faces of actors and politicians are frequently attended to, whereas the photographs of unknown faces do not assume the same communicative function. This may be reflected in the findings of a caricature advantage for famous faces but not for unfamiliar faces (Rhodes *et al*, 1987; Benson and Perrett, 1994; May *et al*, submitted) as the precise spatial relationship of features for the familiar actors/politicians is more likely to be attended to.

Stevenage (1995) found a caricature advantage with artist-drawn sketches of unfamiliar faces. Unlike Rhodes et al (1987)'s stimuli, however, differences between veridical and caricatured drawings in Stevenage's (1995) study often included variations in thickness between lines outlining the face or highlighting particular features, in addition to exaggeration of the size and distances between features. Mauro and Kubovy (1992) used identikit faces which differed from the norm in only one feature. Caricaturing these faces involved producing identikit faces which exaggerated this single feature further. They found a caricature advantage for these unfamiliar faces using this technique. They interpreted their results to indicate an attentional aspect to this processing by which caricatures draw attention to features of the face that are critical for recognition. They suggest that this may account for difficulties found in recognising individuals from races different to one's own with viewers commenting that presented faces of other races have a greater degree of resemblance to one another (Chance, Goldstein and McBride, 1975; Cross, Cross and Daley, 1971; Goldstein and Chance, 1979; Malpass and Kravitz, 1969; Shepherd *et al*, 1974; Shepherd, 1981; Valentine, 1991) with which there is a lower degree of familiarity. Investigation of facial morphology within racial groups reveals that feature variability does not differ considerably across Japanese, Black or Caucasian faces although the mean distribution of facial features differs for each type of face (Goldstein, 1979; *Plate 1.2*). The norms for one's own race with which one is highly familiar may not be appropriate to use when attempting to distinguish members of another race with which one is unfamiliar. Age, sex and socio-economic class of faces may also be factors in their recognition (Cross, Cross and Daley, 1971; Dion, Berscheid and Walster, 1972; Goldstein, 1979a; Klatzky, Martin and Kane, 1982, 1982a; Klatzky and Forrest, 1984). In caricaturing and recognising a face one may compare it to a norm for the appropriate sex, race and perhaps perceived socio-economic status or occupation.

In line with Rhodes and McLean's (1990) finding that the caricature hypothesis is supported for birds with bird experts, Diamond and Carey (1986) have also found that dog experts showed a similar inversion effect when asked to discriminate between pictures of individual dogs to that shown by people when discriminating faces. Yin (1969) found that all mono-oriented objects are more difficult to recognise when upside down, but that faces are disproportionately affected by inversion. He thus postulated a "unique" system supporting recognition of the human face by human beings. Yin arrived at this conclusion by presenting subjects with upright pictures of faces, houses, aeroplanes and other mono-oriented stimuli. After studying these pictures, subjects were then tested with the stimuli in the same orientation using a forced-choice recognition paradigm. Through this technique Yin found that when the stimuli were first viewed and then tested in the upright orientation, photographs of faces were better recognised than those of other classes of object. However, when the same stimuli were initially presented and tested in the inverted orientation, recognition for faces fell below the recognition levels for the other object classes. This finding that upside-down faces appear to be disproportionately more difficult to recognise than other inverted mono-oriented objects has been termed the face inversion effect and this effect has proved to be extremely robust (cf Valentine, 1988) both for famous and novel faces (Scapinello and Yarmey, 1970; Yarmey, 1971) and for simple line-drawn faces (Yin, 1969). It is possible, however, that faces are special in this sense only because compared to other mono-oriented stimulus there is more inter-stimulus similarity and they are more frequently seen, learned, memorised, and discriminated and this phenomenon could be extended to other complex stimuli with a high degree of similarity which one is experienced at discriminating between. Diamond and Carey (1986) tried to explain this phenomenon by attributing the inversion effect to the use of second-order relational properties. They claimed that these properties were important for, but not unique to, face recognition.

In Diamond and Carey's view, while first-order relational properties code the positional spatial relationships among parts of a stimulus (e.g. eyes above and to either side of nose), second-order relational properties of homogeneous objects code the precise spatial configuration between the presented stimulus and the central tendency or prototypical spatial configurations of its parts. They argue that in order to represent a stimulus with respect to its second-order relational properties one must be equipped with the knowledge of the prototype or norm so that the second-order relations can be detected. Thus, of the stimuli whose spatial structures would allow for the encoding of second-order relational properties, the only class of stimuli for which most people have sufficient expertise to use second-order relational properties is faces. However, the use of second-order relational properties is not necessarily limited to faces and the consequent inversion effect is not necessarily restricted to facial stimuli. A study by Tanaka and Farah (1991) which analysed the effect of inversion of dot patterns that differed in the extent to which they required the encoding of second order relational properties failed to substantiate this hypothesis. Tanaka and Farah interpret their findings to indicate that featural information is important to face encoding and subsequent recognition rather than configural information.

Tanaka and Farah's assertion does not contradict the suggestion that inversion will affect other highly homogeneous stimuli in a similar fashion to faces. It merely disputes Diamond and Carey's reasoning that their so-called second-order relational properties are what contributes to the inversion effect. Bartlett and Searcy (1993) counter Tanaka and Farah's (1991) claim; they take the Thatcher illusion (Thompson, 1980) as further indication that face-processing occurs using configural, rather than feature, encoding. Under the Thatcher illusion the mouth and eyes of a face are rotated through 180 degrees without altering their respective positions in the face. When in upright position the face appears quite grotesque, but upon facial inversion the face appears quite normal albeit upside-down.

The hypotheses stipulated by Rhodes *et al* (1987) and Valentine and Bruce (1986, 1986a) and described here in line with the inversion effect would also suggest that due to the sensitivity of the face recognition mechanism to inversion and to faces which are distant from the norm, inversion should cause a greater disruption in recognition for facial configurations with which one is more familiar. Evidence that Chinese and European subjects show a larger loss in accuracy of recognition of inverted faces for own-race rather than other-race faces (Rhodes *et al*, 1989) is further taken as support for the notion that only with protracted experience of faces of one ethnic background do subjects become sensitive to the mean and variation of facial features incumbent with such ethnicity.

It is interesting to note here that Goldstein and Chance (1980) arrived at the notion of a face schema: an idea that faces are encoded relative to some norm which is similar conceptually to that arrived at independently by Valentine and Bruce (1986, 1986a) and Rhodes *et al* (1987). However, rather than arrive at their schema theory through the effects of distinctiveness and typicality on facial recognition, Goldstein and Chance's (1980) schema theory was derived from the effects found for inversion and race. Schema theory was developed from their finding that adults recognise other-race faces less accurately than own-race faces, while children recognise other-race faces and own-race faces and typical and distinctive faces with equal accuracy (Chance, Turner and Goldstein, 1982; Goldstein and Chance, 1980). They also found analogous results for the effect of inversion on face recognition whereby they noted little effect of face inversion on young children (Goldstein, 1975). Goldstein and Chance accounted for the development of the effects of race and inversion by arguing that with increasing age children become more efficient in their use of a face schema, leading to better recognition performance. However, this increase in efficiency is accompanied by an increase in schema rigidity such that as the schema develops, it becomes relatively less efficient at processing unusual stimuli such as inverted or other-race faces. Thus, in common with the distinctiveness hypothesis and Rhodes *et al*'s (1987) norm-based coding model, face schema theory uses the idea that knowledge of the population of faces is acquired and used in face processing. Evidence that accuracy of recognition for typical individuals improves, but the ability to discriminate between distinctive or other-race faces decreases as children age, would also support the idea of the development of an increasingly stable face prototype with age and that varying perceptual histories might effect recognition judgements.

Other studies (Cross, Cross and Daley, 1971; Feinman and Entwisle, 1976; Kagan and Klein, 1973) have failed to find an increase in the effect of race with age and the ability to recognise inverted faces has also been found to improve with age (Flin, 1985; Valentine, 1988). If the effects of race, inversion and distinctiveness are all of consequence to encoding faces by reference to a schema or prototype, then in accord with Sternberg's (1969) additive factors logic these factors should interact. As noted above, interaction has been found between inversion and race (Rhodes *et al*, 1989; Valentine and Bruce, 1986b). However, the findings here are inconsistent as Valentine and Bruce (1986b) found recognition of other-race faces was more disrupted by inversion than recognition of own-race faces while Rhodes *et al* (1989) noted the converse. One reason for this discrepancy may be that only one race of subjects (Whites) were used in the Valentine and Bruce (1986b) study, identifying both White and Black faces, while both Chinese and White subjects took part in the Rhodes *et al* (1989) study and were required to identify both own- and other-race faces. This further suggests that greater expertise with a particular sets of faces results in a disproportionately higher recognition deficit when such faces are inverted. This, in conjunction with the interaction between expertise with individuated stimuli and orientation discussed (Carey and Diamond, 1977; Diamond and Carey, 1986 Goldstein, 1975), suggests a disruption to processing of configural information in inverted faces. Valentine (1991) notes that recognition of distinctive faces may be less affected by inversion because distinctive faces can be more easily encoded in terms of component information (i. e. information based on isolated features rather than global properties). This, in conjunction with Pinker's (1985) proposition that distinct faces may be accessed using a different strategy to typical faces and Rhodes *et al*'s (1989) finding that other-race faces are affected differentially to own-race faces through inversion, suggests that a different heuristic may be used to discriminate stimuli that are more structurally similar and closer to a norm in face space (typical faces) than more distinct faces which are positioned further away from a norm in face-space.

Rhodes *et al* (1987) and Ellis (1981) propose that multiple norms might exist in order to represent different categories of faces. For example, separate norms might be stored for male and female faces and for own-race and other-race faces. If a single norm is used as proposed by Valentine and Bruce (1986a) then other-race faces may be encoded by reference to an own-race norm, only because it is assumed that other-race faces have rarely been experienced. With increased familiarity with faces of other races, there would be a point at which the stored norm would be sufficiently modified such that other-race faces could be encoded and discriminated on the basis of this norm. It has been documented that with increased experience of other-race faces recognition of individuals who are members of this race becomes more accurate (Carroo, 1986, 1987; Chiroro and Valentine, 1995; Elliott, Wills and Goldstein, 1973). Similarly, judgements of race could also be so effected on this basis. A well-travelled European may be able to distinguish a French individual, or composite (description to follow), from that of a Scot; someone with sufficient experience of oriental faces may be able to distinguish between a Chinese, Japanese or Korean or, similarly, someone with sufficient experience in different North American reservations may be able to distinguish members of various Native American tribes. Should multiple norms be accessed in some manner, the question is raised as to how the appropriate norm is selected to encode and access the identity of a particular face. This could be done through compromise between the norm-based models described and an exemplar based model (Valentine, 1991) of face recognition through which a face is recognised via the face which it most closely resembles. The exemplar based model hypothesises differs from the norm-based models in that faces are encoded as points, rather than vectors directed from a norm, such that the norm plays no part in encoding stimuli. In this model, identification of a face is assumed to depend on: 1) an estimate of error associated with encoding the stimulus, 2) the distance between the locations of the stimulus and the nearest-known face, and 3) the distance between the stimulus and the next nearest neighbour (Valentine, 1991). Another possibility is that with increased exposure to faces of a given other-race, an adequate norm for that race could be developed by which one can efficiently code members of that race. If this is the case, perhaps if a face is established to be of sufficiently close proximity in "face-space" to a particular norm relative to another, encoding and subsequent recognition of that face is triggered through a description of the differences of that face from this norm. In support of this possibility, there is evidence that across races the saliency of particular features for recognition differs. Ellis, Deregowski and Shepherd (1975) report that different facial features are used to describe black and white faces. However, the use of a single norm as a basis for face recognition rather than multiple norms is also supported through some evidence that as our experience of other races grows, with the subsequent benefit in recognition, this is accompanied by a detriment to own-race recognition (Chiroro and Valentine, 1995).

Production of photographic quality stimuli

Photographic quality prototypes or composites have been developed from techniques first employed by Galton (1878, 1879). Galton averaged faces by creating prototypes photographically using multiple exposure. In order to minimise blur in these composites, Galton aligned the eye positions of each of the individual faces to be incorporated into the prototype. More recently this technique has been reproduced digitally by Burson *et al* (1986) and Langlois and Roggman (1990) who aligned individual faces by pupil position and averaged faces on a computer. This process has been further refined in recent studies of gender differences (Benson and Perrett, 1991a; Brown and Perrett, 1993), age differences (Burt and Perrett, 1995) and attractiveness (Perrett *et al*, 1994). In these studies computer manipulations of component faces have enabled the position and shape of all the features in a set of images to be matched more precisely (*Plate 1.3*). Following Brennan's (1982, 1985) techniques 179 pre-defined points are positioned to describe the shape of major features for each face image (e.g. one point for the centre of the upper lip). The average shape of the group of faces can then be formed from the mean of the x,y co-ordinates for each corresponding point across the group. A composite image with the average appearance of the group can then be formed by "warping" (stretching as if on a thin sheet of rubber) each face into the average shape and then blending the images together digitally (Benson and Perrett, 1991a, 1991b; Benson and Perrett, 1993; Perrett *et al*, 1994). As a result such composites are far less blurred (Benson and Perrett, 1993) than those generated by Galton's technique or its digital equivalent (Langlois and Roggman, 1990). Photographic prototype images of this form were used in the experiments described in the following thesis and define the average shape and colour information from which caricaturing can be performed. *Plate 1.4* shows prototypes for a set of male and female actors.

By use of this technique photographic images can also be caricatured in shape by warping an image to the shape of the caricatured line drawing. This is performed by dividing the face up into a mesh of triangular tessellations (*Figure 1.3*), the vertices of each formed by three adjacent feature points from the 179 originally delineated (Rowland, 1997; see *Figures 1.1 and 1.2, Plate 1.1* and *appendix*). Sets of corresponding tessellations are produced for the veridical image and for the caricatured line drawing.

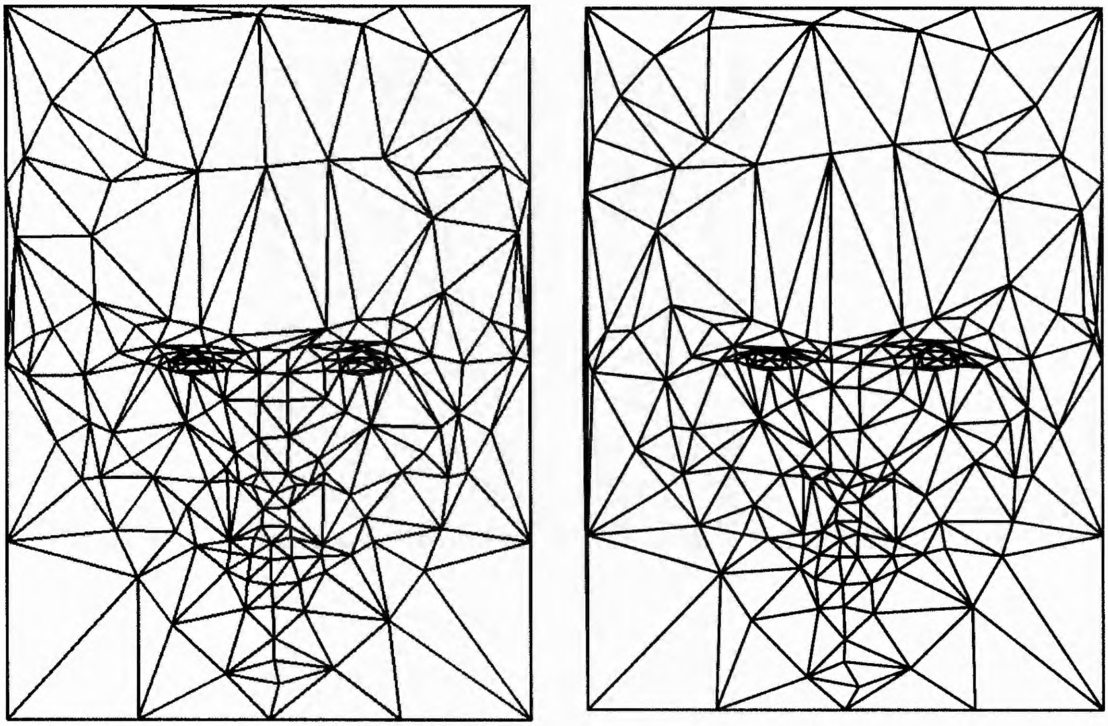


Figure 1.3. A tessellation structure is created from the feature point information of the veridical face and the caricature. The triangles which make up this structure are formed by three adjacent points from the 179 originally delineated. Caricatured photographic stimuli are created by warping the information from the veridical stimuli to the corresponding positions on the caricatured face.

The pixel values for each tessellation of the veridical image can then be remapped into the corresponding tessellation in the caricatured or anticaricatured image (Benson and Perrett, 1991b; Rowland, 1997; *Figure 1.4*). Davies, Ellis and Shepherd (1978) found much greater recognition accuracy for photographs than for detailed line-drawings. This shows that more of the information required for recognition of an individual is present in a photographic quality image than in a line-drawing.

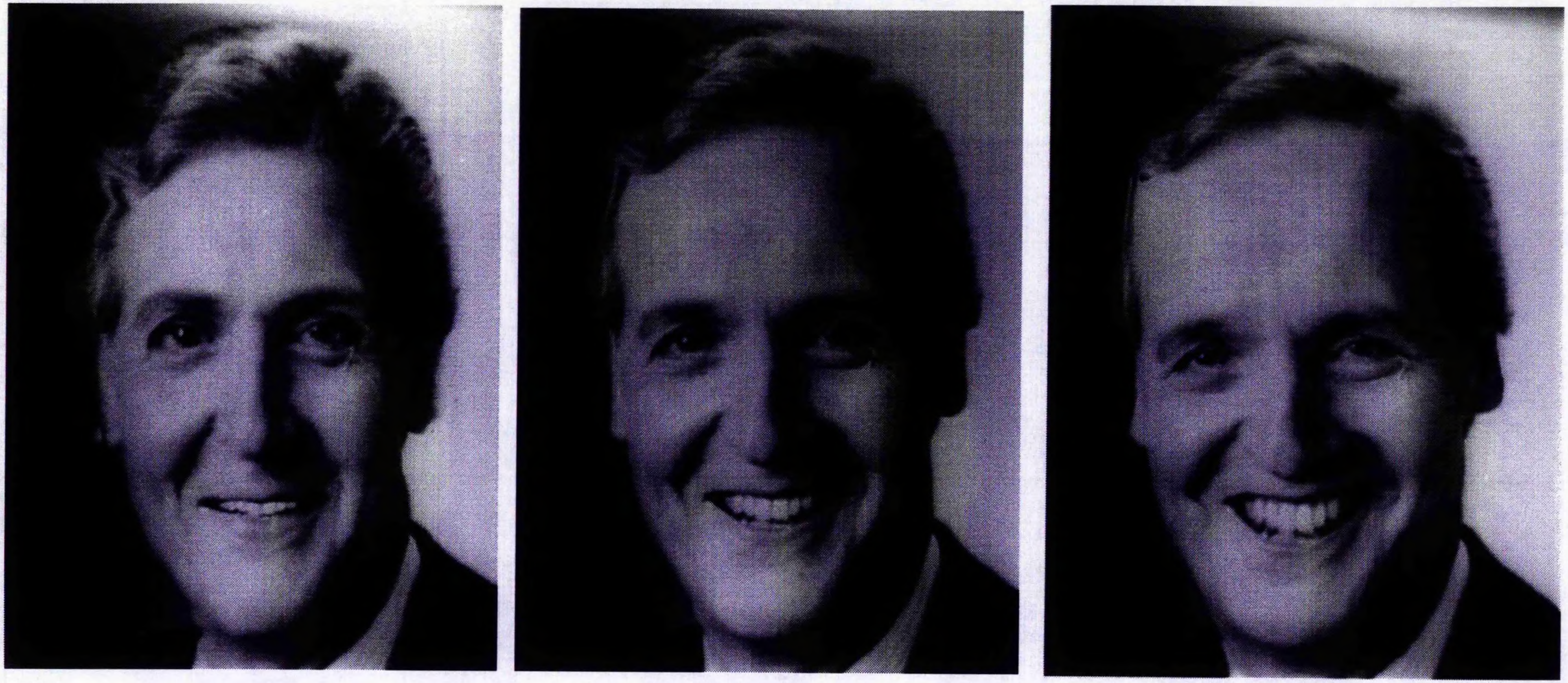


Figure 1.4. Adapted from Benson and Perrett (1991, 1991b). Photorealistic anticaricatures and caricatures. A photorealistic anticaricature (50%, left) or caricature (50%, right) is formed by mapping the grey-scale information from the veridical face (centre) from the veridical tessellation to the anticaricatured or caricatured tessellation structure.

Photorealistic images and the caricature effect

Benson and Perrett (1993) suggested that improved recognition with photographic caricatures may not be expected because such images are still photographic images of those people and already include texture details, which are used to form the basis of any recognition and accentuated in caricature, thus limiting the effect of shape change in photographic images. Certainly the advantage with photographic stimuli is more modest than with line drawings (Benson and Perrett, 1994) though by excluding the external details of the face the advantage can be stronger (Benson and Perrett, 1992). Benson and Perrett (1991), however, did find a reaction time advantage for matching names for caricatured 256 grey-scale photographs of faces over veridical grey-scale photographs when presenting these images on a computer. These advantages were subtle and seen in a small improvement in reaction time when a subject rejected non-target faces. Caricature advantages and faster processing of prototype configurations can be demonstrated in artificial network simulations using photorealistic stimuli with landmark feature points. Again the caricature advantage was found only for images with which the network was familiar and this advantage was superior for typical over atypical faces (Tanaka and Simon, 1996).

As previously stated, one of the findings leading to the distinctiveness hypothesis is the result that distinctive faces are remembered better than typical faces. This, taken in conjunction with the findings in the caricature studies described, suggests that any extreme value of anything that can be measured or described in a face should contribute to this distinctiveness or memorability (Valentine and Endo, 1992). On the basis of this assumption, it has been proposed that it might be possible to correlate memorability of a face with physical deviations based on measurements of individual parts of the face (Ellis and Shepherd, 1987; Bruce *et al,* 1994). Bruce *et al* (1994) found a strong correlation between measures of deviation from the norm and subjective ratings of distinctiveness for the faces and concluded that it is possible to account largely for the psychological variable of distinctiveness in terms of physical deviation from the norm. The pattern of correlations they found was consistent with the possibility that variability of some dimensions contributes to making a face memorable while variability in others contributes to the apparent familiarity of a face; Bruce and her colleagues argue that rated distinctiveness of a face results from deviation on the combined set of dimensions. Although the correlations found between physical variability for faces and memorability/distinctiveness were significant they did not account for a large part of the variance, suggesting that distinguishing between faces might rely on many other cues such as skin texture, freckles, texture of hair at eye-brows and length of eye-lashes. The experiments which follow take advantage of the latest technology developments which allow the manipulation of texture and shape information for photographic facial stimuli. This technology is exploited to build upon previous exploration into the caricature advantage using photographic stimuli (Benson and Perrett, 1991, 1992, 1994) and to explore the effect of such digital manipulation on various aspects of face perception.

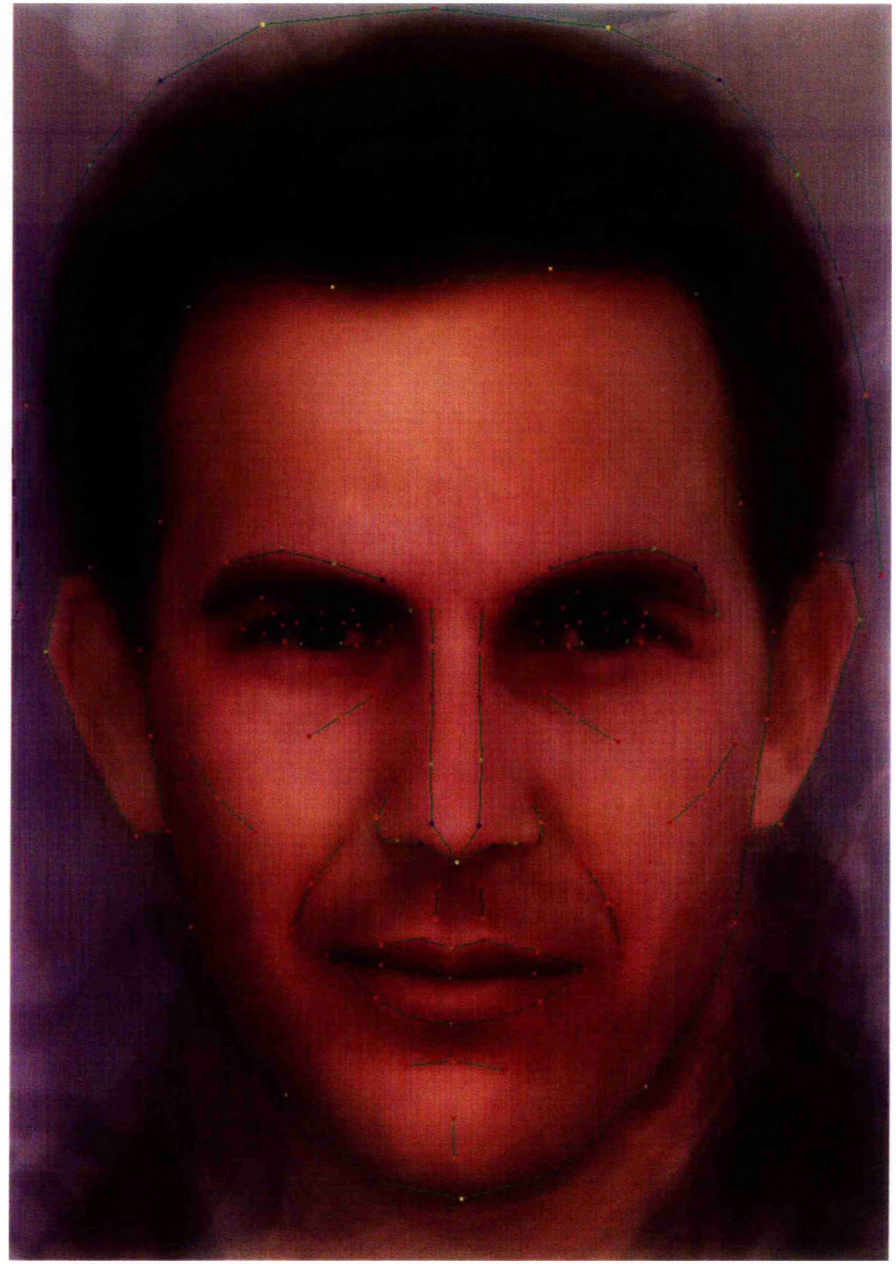


Plate 1.1. Points can be placed on an individual's face, here the actor Kevin Costner (a composite is used for purposes of illustration), in a uniform manner for all faces to be used. These points can then connected in particular ways using spline curves in order to form the outline (internal and external) of a face. Here, 175 points have been used (179 when the mouth is open).

Plate 1.2. These line-drawings represent an average male Japanese (left) and average male Caucasian (right) face. Each of these averages is formed by taking the shape data for 10 individuals of the relevant race. These faces are quite different in structure. Goldstein (1979) has argued that, although the mean shape of these face is very different, variation around the mean shape would be similar for both race. Individuals of a given background may use something analogous to this mean shape when judging faces. This may make recognition of other-race faces, with features lying far outside this individual mean, more difficult than discriminating faces of one's own race.

Plate 1.3. Earlier methods used to produce blends have normalised constituent images on inter-ocular distance. This technique results in a degree of blurring on the composite face (top).

Warping each face shape to the average shape and blending the colour information following this process ensures that all the facial features are aligned and this reduces the degree of facial blurring (bottom).

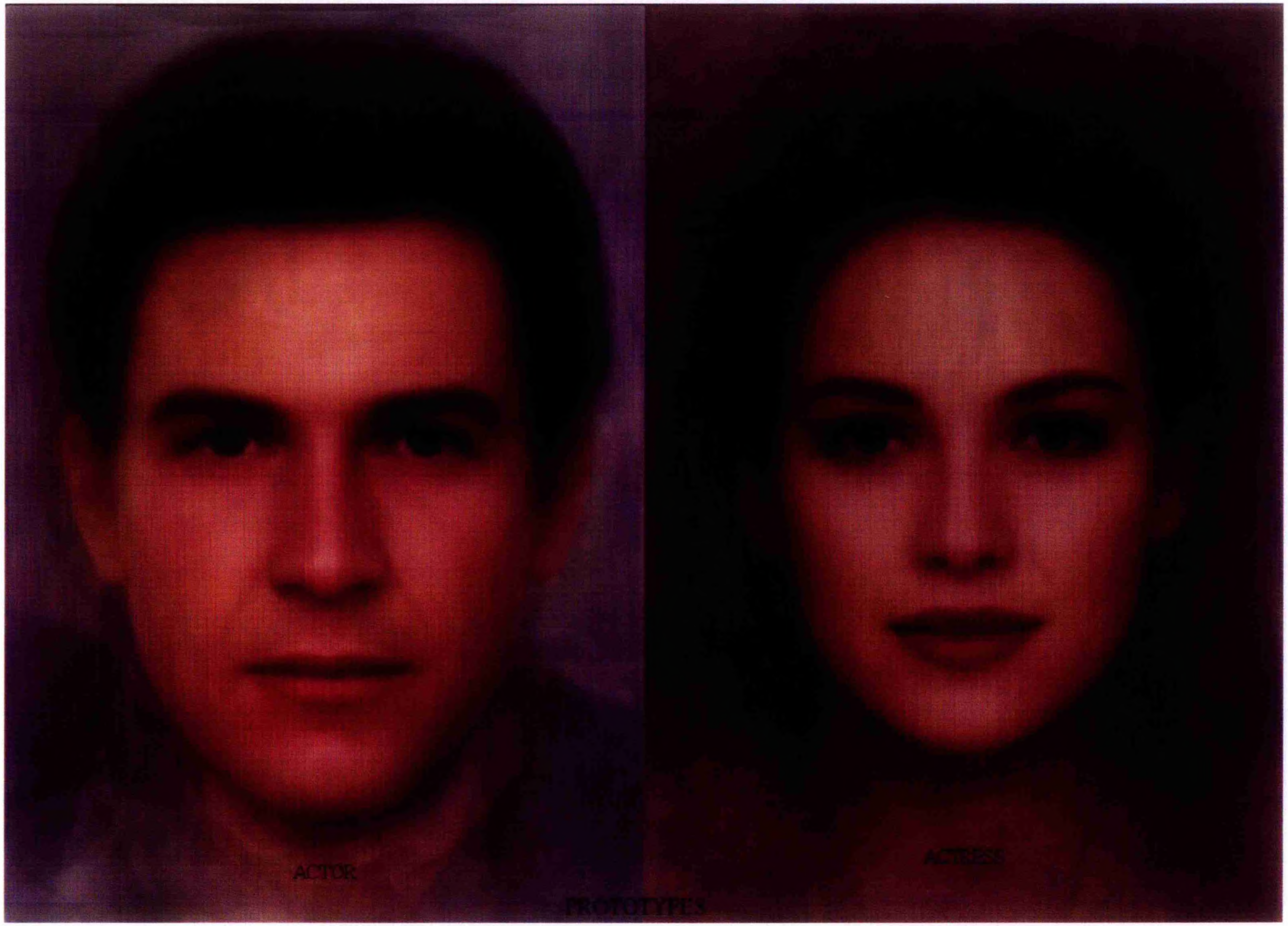


Plate 1.4. Photorealistic composites, or prototypes, can be created for various types of faces, including those for different gender and race. Here, separate composites have been formed from a collection of 22 male actors and 22 female actresses. Each prototype was produced by warping the shape of each component face image to the average face shape derived from the relevant set of actors or actresses and then averaging the RGB colour values for each pixel.

Plate 1.5. The picture on the left is a veridical image of the actor Kevin Costner. The picture on the right is a 50% caricature. This was created by warping the colour information for Kevin Costner to the caricatured shape using the new caricatured tessellation structure as described in Figure 1.3.

Chapter 2 - Experiment One: Perception Time and Shape Caricature

Introduction

The experiment which follows was designed in order to examine the possibility of a benefit to recognition with photographic caricatures. There is some evidence that photographic caricatures benefit recognition in reaction time studies (Benson and Perrett, 1992, 1994). However, this advantage is not consistent despite consistent reduction in reaction times for caricatures in these studies. This may be because such great detail is available from photographic images that any advantage for caricatures is lost with the display times used in reaction-time studies (Benson and Perrett, 1993). By showing faces at brief presentation times so that the amount of information which the viewer can glean from the photographic images is reduced, it may be that a more consistent caricature advantage would occur with caricatured stimuli.

In Shepard's (1984) view, psychometry of reduced input to the visual system by means of impoverished representations of natural objects may be a better and more controlled means for determining the properties of internal representations for classes of objects. Although not actually impoverishing the internal detail of a face, presenting photographs for extremely brief periods of time does degrade the total information available from a photographic image while still allowing the investigation of the effect that caricaturing (enhancing the difference from the norm) a photograph has on the recognisability of a face. Presenting faces for brief periods of time in this manner does not allow the subjects the opportunity to make realism judgements as to whether or not a caricatured image is a realistic representation of the target subject. It simply provides an index of the effect of manipulation on processing of presented images.

Therefore, an alternative approach to the reaction time paradigm described in the preceding chapter is the use of a presentation time technique. This allows the experimenter to vary the amount of time for which an image is displayed while measuring recognition accuracy. Repeated presentation of different stimuli over varying display epochs may be used to measure the time necessary to discriminate stimuli. This paradigm (inspection time) has been used most commonly to investigate the processing time required to discriminate between simple line stimuli (Nettlebeck, 1982; Vickers and Smith, 1986) and between face stimuli (Lee, 1992). Presenting faces for brief periods of time in this manner may also prevent subjects from judging whether or not a manipulated image is a realistic representation of the target subject. For instance, Benson and Perrett (1991) found the most common best-likeness for famous faces was chosen at veridical rather than a caricature. This may be important in the study of caricature effects as several studies reveal a 'reality principle'. In their studies of shape caricatures with line drawings of birds, Rhodes and McLean (1990) found that although 50% caricatures - stimuli in which the metric differences between an individual bird and a prototype are exaggerated by a factor of 50% - were generally recognised quicker than veridical drawings in reaction time paradigms, the same caricatures were judged to be poor likenesses of the subject they represented. If a caricature is seen as non-real, then this may present subjects with a Stroop-like conflict of responses to the question 'is this the face of Margaret Thatcher?' It is valid to respond "no" as this stimulus is a distortion - no one looks like that in reality; it is also valid to respond "yes" as the depiction looks more like Thatcher than anyone else. Thus in reaction time paradigms, faster processing of caricatures can be counteracted by a reduction of correct identifications. Rhodes and McLean (1990) have noted that because of this, reduction of proportion of correct identifications of caricatures should not necessarily be seen as a speed-accuracy trade off.

A simplified version of the inspection time paradigm was designed in order to determine whether photographic quality caricatures of famous faces could be recognised accurately at briefer presentation rates than veridical images. The study described here sought to confirm the effect of caricaturing in shape for familiar faces found by Benson and Perrett (1991) using this new perception time technique.

Photorealistic images of famous faces were selected for use in this study. Although employment of faces personally familiar to subjects allows standard photographic conditions previous studies have found recognition advantages for shape caricatures of celebrity photographs (Benson and Perrett, 1991; Calder *et al*, 1996) but not for other personally familiar individuals such as classmates (Benson, 1992; Benson, Perrett, Dyson and Calder, unpublished). This may be because such faces may not be consistently familiar to all subjects or that celebrities, particularly actors as selected here, may have been seen with a wider range of hairstyles and facial expressions. Since the level of familiarity affects the benefit of caricaturing (Rhodes and McLean, 1990) the faces of celebrities which were chosen were selected for high levels of familiarity across the subject pool. The external features of the faces and background information were removed so as not to produce additional cues to the identity of the individual being shown. Further, it has been illustrated that recognition of famous faces is determined largely by internal features (Ellis *et al*, 1979; Young *et al*, 1985). A stronger caricature advantage has also been found after removing such external features in a reaction time paradigm (Benson and Perrett, 1992). It is likely that without being able to rely on such external information, subjects utilise the additional information from the caricatured features to recognise the individual. These stimuli differed to those used in the previous studies of photographic caricatures (Benson and Perrett, 1991, 1992) as the images were captured and presented in full colour rather than 256 shade grey-scale.

Methods

Apparatus. Silicon Graphics 4D-25 Personal Iris computer with a non-interlaced monitor synchronised at 60 Hz, Epson GT-6500 scanner connected to an Escom 486 DX2/66 IBM-compatible personal computer

Subjects. 11 undergraduates at the University of St Andrews; 9 females and 2 males

Stimuli. The faces of 41 actors and 22 actresses were scanned at a resolution of 517 pixels in width by 720 pixels in height at 150 dpi in 24-bit colour. Each of the individuals were clean-shaven and wore no spectacles. The features on the faces were defined using 179 feature points placed manually in standardised positions (see *appendix* and *Plate 1.1*). A separate composite was created for each of a set of 22 actors and the set of actresses (*Plate 1.4*). These composites were formed by using the position data from all the faces in the appropriate set to compute an average for shape. Average feature colour information was then calculated by digitally averaging corresponding pixels of individual faces from each face in the set after they were warped into the group average shape. These methods are described in greater detail elsewhere (*Chapter 1*; Rowland and Perrett, 1995).

A questionnaire was given to seven undergraduate students at the University of St Andrews. From a list of 24 well-known male and female celebrities, these students were asked to rate on a scale of one to five how familiar they believed themselves to be with the face of each of these celebrities. Each student was then shown several images of each of the five male and five female celebrities that they had felt to be the most recognisable. For each celebrity, the student was asked to put into rank order how well they thought each image of the individual celebrity represented their personal impression.

From this survey, 10 images - each of an individual celebrity - were chosen for use in the study (see *Plate 2.1* and *appendix*). In this collection of ten images were five males and five females.

In each image, the internal features from the face were cropped from the hair, ears and background so that these could not provide additional cues to identity (*Plate 2.2*). As noted, facial hair and glasses were also not evident in any of the photographs so that this could not have further discriminating power. A black background was then substituted for the former background and external facial features of each image.

A mask stimulus was also created in the shape of several professional models (male and female). The faces from which this shape was derived were neither individual test stimuli nor were they incorporated in the prototypical blends used for caricaturing the test stimuli. The colour information apparent in this mask was produced by blending the RGB colour values for each pixel from each of the individual test stimuli. Shape information for the mask stimulus was derived from faces not included in the test in order to minimise the possibility that beneficial recognition for caricatures might result from caricature appearing more different to the mask stimulus. Again, this mask was cropped around the internal face area (see *Plate 2.3*). Such a mask stimulus was used in order that individual stimuli would not be easily recognised at such a brief presentation time.

Each of the target images was caricatured against a gender relevant prototype of actors or actresses. This caricaturing process exaggerated the metric differences between the individual and the prototype by 40% (*Plate 2.4*). The colour of the images was left unaltered. Each of the 20 images - 10 veridical images and 10 caricatured images - was normalised to the same inter-ocular distance as the mask so that all the faces would be presented in the same position on the screen and would appear in a similar orientation. The mask and stimuli all had the dimensions 517 x 720 pixels.

Design. The independent variables in this experiment were the two levels of caricature, veridical images and caricatured images, and the durations over which the images were presented (33 milliseconds, 67 milliseconds, 100 milliseconds, and 133 milliseconds). Images were presented in blocks with the display duration increasing from block to block. This was done in order to prevent extended exposure to the stimuli to influence recognition accuracy. This experiment was trying to determine difference in recognition accuracy between stimuli at exposures just sufficient to produce recognition greater than chance. If the images had been presented initially at 133 milliseconds then preference may have been at ceiling. The dependent variable investigated was the accuracy of recognition for individual faces when presented (in either the caricatured or veridical form and at any of the durations of presentation). A within-subject design was employed such that the same subjects saw all the stimuli in both veridical and caricatured form and displayed at all four durations.

Procedure

Introductory phase. Subjects were initially instructed as to the nature of the task in which they were to participate. This involved displaying two target images on the screen in uncaricatured form that were not to appear in the test phase. These images were shown with the display rates in reverse order to that in the test phase. Subjects were first shown the images at 8 frames, then at 6 frames, 4 frames and finally 2 frames, where the duration of a frame was 16.7 milliseconds. In each instance, a mask was first presented for 500 milliseconds followed by the test face image, followed again by the mask for a further 500 milliseconds. Here the mask acted as a signal to indicate the onset of a target image and as a forward and backward masking device. This phase was used in order to alert subjects to the nature of the task, particularly that the image that they were to name was the individual target image and not the mask.

Test phase. Subjects were given a list of the names of the ten target faces to be presented in the task. They were asked to rate each of the ten celebrities on a scale of 1 to 5 to indicate how familiar they believed themselves to be with a given individual's face, where 1 indicated completely unfamiliar and 5 indicated a high degree of familiarity.

Subjects were informed that a mask would be presented for 500 milliseconds followed briefly by a target face followed again by the mask. Subjects were told that the nature of the task was to determine the nature of the relationship between recognition and presentation time. They were not informed of the nature of the stimuli (caricatured or otherwise) to be presented. They were asked to write down the name of the celebrity which the target face represented following each trial from the list given. Trials were given with veridical images and shape-caricatured images presented in random order. Trials were given in blocks of 20, with 10 images from each of the two image conditions presented. Images were presented on a monitor using the Silicon Graphics computer in a dimly lit windowless room. During the first block of 20 trials, images were presented for a duration of two frames. In each subsequent block of trials the images were presented for a further two frames. Four blocks of trials were shown in total, the last in which images were displayed for eight frames.

After the task was complete, each of the veridical images was presented in turn and subjects were asked to rate each on a scale of 1 to 5 (where 1 indicated highly unrepresentative and 5 indicated highly representative) how close to their personal representation of the relevant celebrity the believed the images to be.

Results. Medians were found for the post-task questionnaire to indicate how good the likenesses were seen by the subject population used. Three of the celebrities were then eliminated from further analysis because they had medians of 1 or 2 from the 5 point scale of recognisability. The results for the caricatures and veridical images of the remaining seven celebrities were used collectively in further analysis .

Subjects were scored for correctly identifying an image representing an individual celebrity on any given trial. The results were tallied separately for veridical and caricatured images. These scores were also marked separately for each of the blocks of four trials.

A within-subject analysis of variance [with accuracy transformed by $\arcsin\sqrt{(x/y)}$ with Bartlett's correction[1] where x is the number of correct responses given and y is the number of correct responses possible; Snedecor and

[1]**Footnote.** This transformation was performed as the data could be described as a fraction (number of faces correctly identified out of the total possible). Bartlett's correction was used as there were many instances at the briefest display duration whereby overall accuracy was at 0% and at the longest display duration whereby overall accuracy was at 100%.

Cochran, 1980, pg. 290] was carried out on the data across duration (4 levels) and level of caricature (2 levels) as main factors. This analysis found a significant main effect for duration (*Figure 2.1*, $F_{3,30} = 46.93$, $p < 0.0001$, mse (mean square error) = 0.04) but not for level of caricature ($F_{1,10} = 0.002$, p=9.66, mse = 0.02). An interaction was found between caricature and duration (*Figure 2.2*, $F_{3,30} = 3.45$, $p < 0.05$, mse = 0.03). A posteriori PSLD (protected least significant difference; Snedecor and Cochran, 1980, pg. 234) analysis of this interaction found significantly greater correct responses given to caricatured stimuli than to veridical stimuli at 33 ms (*Figure 2.3*, $p < 0.05$).

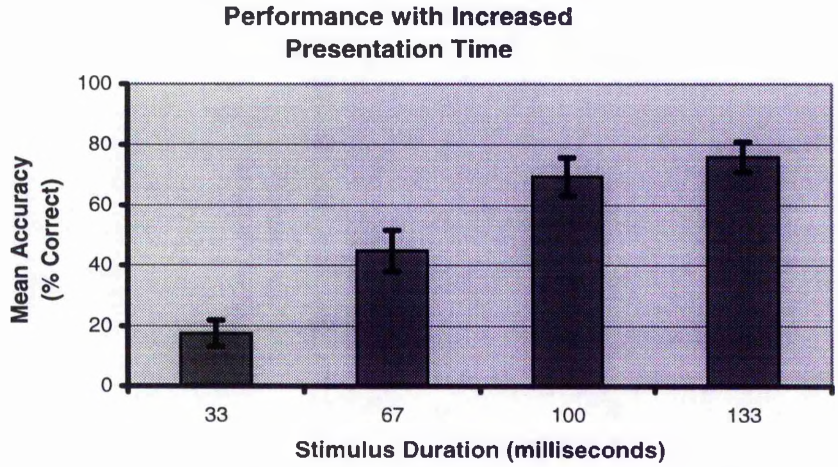


Figure 2.1. Performance with increased presentation time. Accuracy discriminating faces improves with increased duration of display ($p < 0.0001$, F_{30}=46.93, mse=0.04).

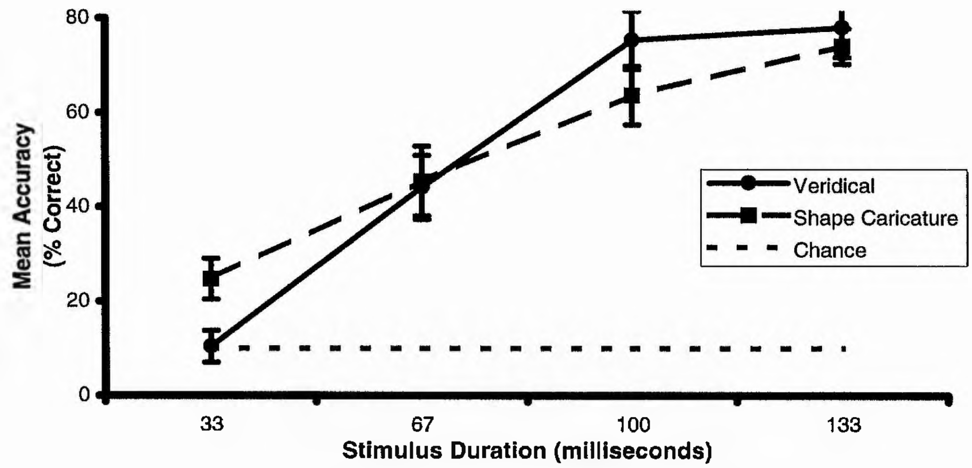


Figure 2.2. Recognition accuracy between conditions with increased presentation. Face recognition at short durations is superior for shape-caricatured photorealistic images than to veridical images ($p < 0.05$, $F=3.45$, $df=30$, $mse=0.03$). Caricaturing produces an advantage in recognition at short durations (33 ms), but not at longer durations.

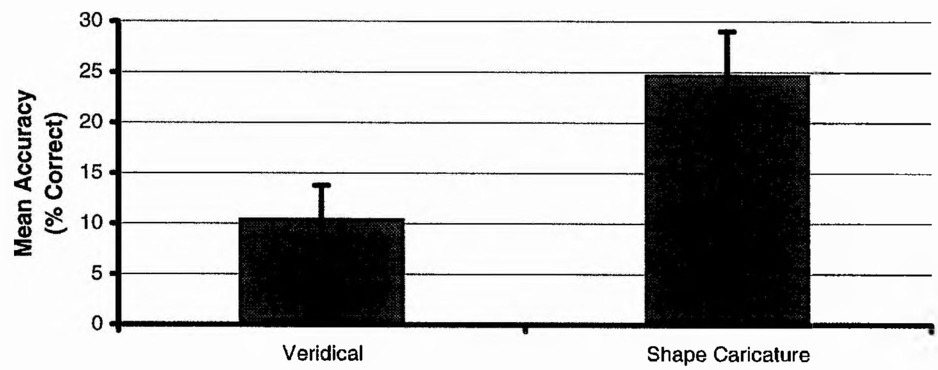


Figure 2.3. Recognition advantage for shape caricatures. There is a significant improvement in recognition with shape caricatures at 33 ms ($p < 0.05$).

Discussion

As one might expect general recognition accuracy improves as images are displayed for longer periods of time. Enhancing the shape information relevant to a given person's face increases the ability to recognise a representation of an individual which is presented for an extremely brief period of time. This effect is lost at longer durations. This is either as a result of a cumulative learning effect such that faces are learned at the first duration of 33 ms so there is no difference between recognition accuracies for veridical or caricatured images at longer display rates or because the importance of shape information for facial discrimination is only observable when the image is sufficiently degraded through showing it for a maximum of 33 ms (two frames). In order to examine this further, this experiment could be repeated by counterbalancing the order of duration at which images are displayed. Thus some subjects might be presented with the images initially at a duration of eight frames, other subjects at six frames, others at four and finally some subjects at two. Later studies using full counterbalancing revealed no order effects with the perception-time paradigm (*Chapters 3 and 4*). It seems that images must be displayed for a sufficiently brief period in time in order to inhibit processing and for the extra information available in the caricatured image to be a benefit to recognition. There is also some possibility that the type of mask used in this experiment may have contributed to the caricature effect. Caricatured images may have appeared more distinctive relative to the mask. Should this benefit recognition, the effect itself might lend some support to the caricature hypothesis. However, in order to avoid the mask being a comfounding factor, attempts were made to improve the mask stimulus in many of the subsequent experiments.

This result[2] confirms previous findings using line-drawings (Benson and Perrett, 1994; Carey, 1992; Rhodes *et al*, 1987; Rhodes and McLean, 1990; Rhodes *et al*, 1996) that caricaturing in shape aids recognition when images are degraded. Unlike reaction time studies. where images are displayed for longer periods of time, this perception time task allows subjects the freedom to respond in their own time while the images are degraded by being presented for very short

---

[2]**Footnote**. The results for many of the experiments described in this thesis were subjected to ANOVA using faces as the random factor. The main effect of image type failed to reach significance. This lack of significance is attributable to the reduction in the power of the analysis, since the degrees of freedom of the random factor were reduced. For example in Experiment 4, there were 35 subjects but only 9 faces (therefore the degrees of freedom reduce from 34 to 8).

periods of time. The findings here also differ from Rhodes *et al*'s (1987) in that greater recognition accuracy with caricatures at brief presentation time was found. This provides further evidence that configural coding is an important pre-requisite to face recognition and that this information is enhanced by caricaturing as it has been asserted that configuration is important if the task is to be done quickly (Bartlett and Searcy, 1993). There is an interaction suggesting that brief exposure is a form of image degradation and favours the right hemisphere configural processing (Sergent, 1984). Rhodes *et al* (1987) show significantly superior performance for their reaction time data but not for their accuracy data (with 33, 28 and 27% performance for caricature, veridical and anti-caricature images respectively). This may, however, be due to the pressure on the subject to respond as rapidly as possible in the Rhodes *et al* study.

In conclusion the presentation time paradigm has been successfully applied to studying the effect of shape caricature. Caricatures appear advantageous at short image durations - when the images is particularly degraded. Such degradation may favour configural processing, allowing the amplification of distinctive cues to benefit recognition while at the same time preventing subjects from realising that caricatures are in fact large distortions from reality.

Plate 2.1. Ten images, each of an individual celebrity, were used as a basis for Experiment 1. Each of these images was later cropped with the external features and background information to be replaced by a uniform black background. Shape caricatures were also produced for each of these images.

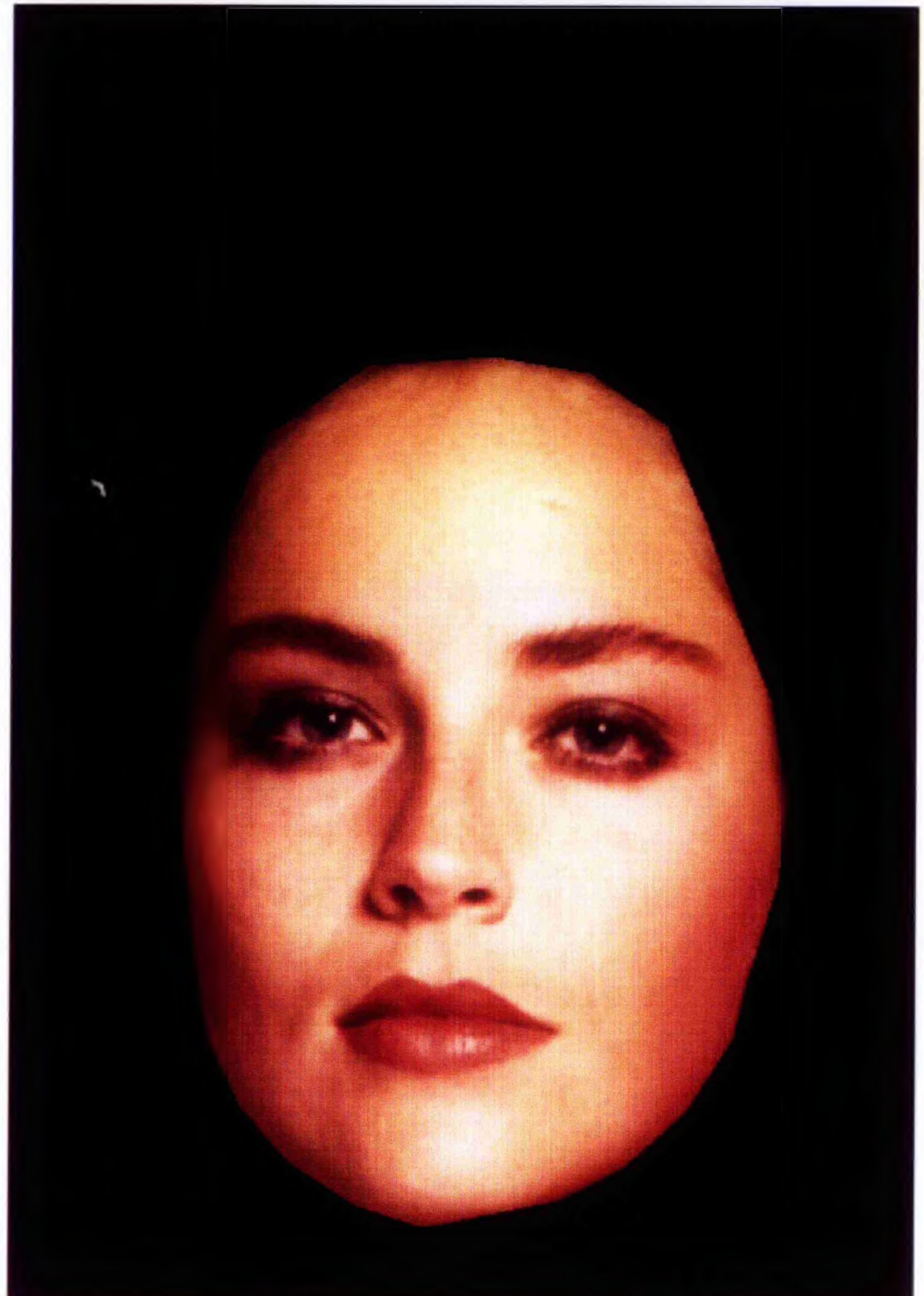


Plate 2.2. Each of the ten images of celebrities used in Experiment 2 was cropped to remove the external features such as ears and any information from the background which might provide cues to the identity of the celebrity. These cropped features were replaced with a uniform black background.

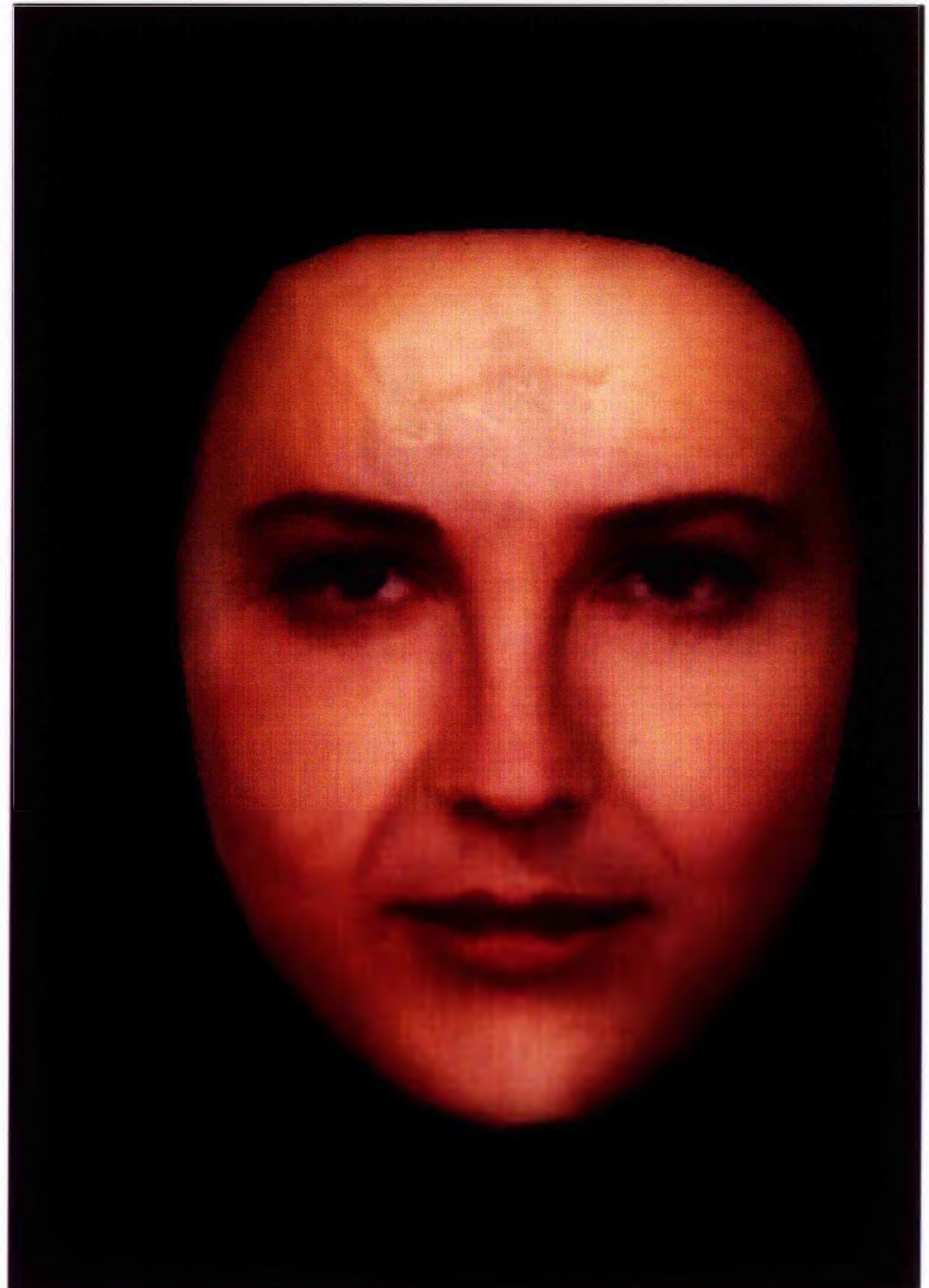


Plate 2.3. A mask stimulus was created. This was used as a forward and backward masking stimulus and as an indicator of stimulus onset. The shape of the mask was derived by averaging the shape information of 60 models, 30 of each gender. The colour of the mask was formed by calculating the average RGB information for each pixel from the ten images of celebrities to be used in Experiment 1.

Plate 2.4. Each image was presented in two forms: the veridical or unaltered image and a caricature. Both images were cropped to remove the background, neck and ears. This was replaced with a black background. The caricature represented the shape differences between the individual - Sharon Stone, here - and the composite (an average actress in this case) exaggerated by a factor of 40%.

Plate 2.5. Seven of the veridical images of celebrities were rated as being highly typical of the individual which they represented by the participants in the experiments. Each of these had been caricatured in shape by 40%. The caricatured images were recognised with greater accuracy at brief presentation times than the veridical images.

Chapter 3 - Experiment Two: The Role of Colour in Recognition

Introduction

Role of colour in recognition

The role of colour in the recognition and discrimination of objects has been the subject of much debate. Biederman's (1987) account of object recognition proposes that the basic input for recognition of objects is an edge-based representation. This representation would be akin to the primal sketch proposed by Marr and Nishihara (1978) and similar to that available in a line drawing. According to Biederman's account it is the geometrical aspects of the edge-based representations that are crucial in object recognition and these can be sufficiently well specified in line drawings. Certainly the sufficiency and importance of edge-based information for face recognition are supported by caricature studies using line drawings (Benson and Perrett, 1994; Rhodes *et al*, 1987; Rhodes *et al*, 1996) which have been described in *Chapter 1* and photographic stimuli (Experiment 1, *Chapter 2*). Biederman and Ju (1988) further state that surface cues such as colour and texture are generally less efficient routes for "accessing the memorial representation of an object category ... we may know that an image has a particular colour, brightness and texture simultaneously with its edge-based description, but it is only the edge-based description that provides efficient access to the mental representation" (page 40) of the object. Supporting their view, Biederman and Ju (1988) demonstrated that colour pictures and line drawings were recognised with reaction times which were not significantly different even for objects with a diagnostic colour such as a banana.

Other investigations have contradicted these observations and report superior performance for naming objects in coloured photographs compared to line drawings (Brodie *et al*, 1991; Davidoff and Ostergaard, 1988; Price and Humphreys, 1989), for coloured photographs compared to black-and-white photographs (Ostergaard and Davidoff, 1985) and for coloured computer images compared to grey-scale images (Wurm *et al*, 1993). It has been suggested that these studies may contrast with those of Biederman and Ju (1988) because of differences in the inclusion of organic objects as stimuli (Humphrey *et al*, 1994). Humphrey and his colleagues note that many of the studies of object recognition which demonstrate an advantage for colour representations over line drawings or grey-scale representations have included many organic items such as fruits, vegetables, animals and other objects with characteristic colours (Ostergaard and Davidoff, 1985; Wurm *et al*, 1993; Price and Humphreys, 1989), whereas Biederman and Ju (1988) used only four organic objects - mushroom, apple, banana and fish (out of a total of 29 stimuli). Each of these four objects presents a different shape from the others and the diagnosticity of colour for some of these organic objects is also arguable as apples, fish and mushrooms can vary in colour quite dramatically.

Price and Humphreys (1989) also note that Biederman and Ju used manufactured objects from disparate categories. Price and Humphreys argue that colour and other surface properties are necessary or at least helpful in discriminating among members of a structurally similar class of object (animals, birds, fruit, insects, plants and vegetables). For such objects, Price and Humphreys found naming to be more rapid with coloured photographs and appropriately coloured line drawings than it was with black-and-white line drawings or inappropriately coloured line drawings. The lack of effect of colour found by Biederman and Ju (1988) may have been dependent on the fact that most of the objects in their study were manufactured, not members of a homogeneous class and therefore colour lacked any discriminative function. One could predict that colour should show its largest effects when natural objects with subtle variations in shape and colour are to be recognised. More recent findings (Humphrey *et al*, 1994) support this proposal as colour information was found to improve recognition consistently across a variety of natural objects.

Human faces are one such set of natural objects for which there are a great many members and a strong degree of homogeneity in both shape and colour. It would be expected then that the inclusion of the colour information pertinent to an individual face would improve recognition compared to an achromatic version of that face and that the exaggeration of the relevant colour away from a suitable facial norm or "prototype" would improve individual recognition compared to a veridical colour image. This would be consistent with findings that, although edge-based information is often sufficient for identification (particularly in caricatures), faces are poorly recognised from outline drawings (Davies, Ellis and Shepherd, 1978; Leder, 1996) and laser scanned 3D faces (Bruce *et al*, 1991), which lack information about superficial properties such as pigmentation and texture, compared with photographs. The importance of intensity information for face recognition is highlighted by the finding that recognition accuracy for black and white photographs of famous faces dropped from 95% correct to 55% when the same faces where shown in photographic negative (Bruce and Langton, 1994; Johnston *et al*, 1992) and shape-from-shading information was demonstrated to be an important factor in the detriment to recognition of faces in photographic negative (Kemp *et al*, 1990). Here, the intensity information for the face was found to be important in conveying configural information, asserted to be important for face recognition (Rhodes *et al*, 1987), in addition to feature information. This is indicated by the finding that subjects took longer to identify whether two faces belonged to the same person when one of the pair was subjected to the Thatcher illusion (Thompson, 1980) if faces were inverted and in photographic negative than if they were inverted but kept their natural illumination (Lewis and Johnston, 1997). Colour and intensity may also convey shape, and thereby configural, information to the face not available in two dimensional shape representations alone through shadow.

The presence of colour and intensity information in faces has been found to aid discrimination of sex and race (Hill *et al, 1995),* and to influence perception of age (Burt and Perrett, 1995) and attractiveness (Perrett *et al*, unpublished). The role of colour in face identification is a relatively neglected subject yet eye colour is a characteristic of individuals which is frequently remarked upon.

The importance of colour may occur at an implicit level of processing (Wurm *et al*, 1993) since colour has been found to reduce reaction time in naming objects displayed on a computer screen where subjects had no explicit knowledge of the natural colour of the objects to be named. That is, when subjects were asked to describe the colours specific to a given object prior to testing they were unable to do so yet, nonetheless, colour was beneficial to their recognition. This may be important given that the discrimination of faces involves responding to subtle differences in pigmentation between individuals. We may be unable to describe pigmentation characteristics of an individual but much of our knowledge about face cues may not be accessible to reflection or explicit description.

Definition of colour

There are several ways of representing 'colour space'. One co-ordinate system is HLS space which defines hue, lightness and saturation. Another is RGB space whereby the term colour is used to refer to red, green and blue intensity distributions in images. This specification of colour is analogous to that produced on colour television screens, where images are formed by the firing of electrons at red, green and blue phosphors at varying intensities from a cathode ray tube. Techniques have recently been developed which allow the colour information found in a face image to be varied systematically in HLS or RGB colour-space (Rowland and Perrett, 1995). Manipulations in RGB colour space affect both colour (hue and saturation) and intensity and contrast. The stimuli used in these experiments were created through such RGB manipulations and the term colour is used to refer to this representation.

In the experiment described here the role of colour hue and saturation information in face identification was investigated by contrasting recognition of full colour and grey-scale versions of the same images.

Methods

Apparatus. A Silicon Graphics 4D-25 Personal Iris computer was used with a non-interlaced monitor synchronised at 60 Hz; an Epson GT-6500 scanner was connected to an Escom 486 DX2/66 IBM-compatible personal computer.

Subjects. 15 undergraduates at the University of St Andrews took part in a survey of familiarity and picture quality of celebrities. 12 different undergraduates at the University of St Andrews (6 females and 6 males) were participants in the main experiment.

Stimuli. The facial areas of photographs of 41 actors and 22 actresses were scanned at a resolution of 517 pixels in width by 720 pixels in height at 150 dpi in 24-bit colour. Individual stimuli used in the experiments consisted of a subset of these images. Each of the individuals was clean-shaven and wore no spectacles. The features on the faces were defined using 179 feature points placed manually in standardised positions.

A questionnaire was given to 15 undergraduate students at the University of St Andrews. From a list of 158 well known male and female celebrities, these students were asked to rate on a scale of one to five how familiar they believed themselves to be with the face of each of these celebrities. Each student was then shown several images of each of the five male and five female celebrities that they had felt to be the most recognisable. For each celebrity, the student was asked to put into rank order how well they thought each image of the individual celebrity represented their personal impression. From this survey, 10 images - each of an individual celebrity, five males and five females - were chosen to maximise familiarity and representativeness of the image for use in the study (*Plate 3.1*). A grey-scale version of each of these 10 images was also produced (*Plate 3.2*). The intensity of the pixel values for the grey-scale images was calculated from the values of the red, green and blue plane using the equation:

$$grey = 0.30 * red + 0.59*green + 0.11*blue$$

In each image, the internal features from the face were cropped from the hair, ears and background so that these latter items could not support recognition. The internal features were then placed on a black background. The border linking the outline of the face to the background was also softened through a convolving procedure, blurring over a gradient, so there was not an abrupt change between the facial image and the black background.

A mask stimulus was developed by taking features from the internal facial area of several faces and placing them at random on a face background. Averaging the colour information of the new mask stimulus and a grey-scale version of the same image produced a 'sepia' image (*Plate 3.3*) with the colour saturation half-way between that of the colour and grey-scale mask in RGB space. The resultant mask was therefore equidistant from the grey-scale and the colour images on average. Each image was normalised to constant eye-centre positions. The mask and stimuli had the dimension 517 x 720 pixels.

Design. The independent variables in this experiment were the two types of image, colour images grey-scale images, and the durations over which the images were presented (33 milliseconds, 67 milliseconds and 100 milliseconds). Images were presented in blocks with the order of display duration counterbalanced between subjects. This was done in order to exclude any potential order effects and effects derived from repeated exposure to the stimuli. The dependent variable investigated was the accuracy of recognition for individual faces when presented (in either the colour or grey-scale form and at any of the durations of presentation). A within-subject design was employed such that the same subjects saw all the stimuli in both colour and grey-scale form and displayed at all three durations.

Procedure

Introductory phase. Subjects were initially instructed that they would be presented briefly with faces for recognition. A demonstration was made involving the display of two veridical images that were not to appear in the test phase on the screen in grey-scale form. These images were shown at decreasing display durations. Subjects were first shown the single images at 6 frames, then 4 frames and finally 2 frames, where the duration of a frame was 16.7 milliseconds. In each instance, the mask stimulus was first presented for 500 milliseconds followed by the test face image, followed again by the mask for a further 500 milliseconds. Each of these images - mask and face - was presented in the centre of the screen while the remainder of the screen remained black for the entire set of trials. Here the mask initially acted as a signal to indicate the onset of a target image. It also functioned as a forward and backward masking device. This phase was used in order to alert subjects to the nature of the task and the mask and, particularly, that the image that they were to name was the individual target image rather than the mask stimulus.

Test phase. Subjects were given a list of the names of the ten target faces to be presented in the task. They were asked to rate each of the ten celebrities on a scale of 1 to 5 to indicate how familiar they believed themselves to be with a given individual's face, where 1 indicated completely unfamiliar and 5 indicated a high degree of familiarity.

Subjects were informed that a mask would be presented followed briefly by a target face followed again by the mask. Subjects were told that the purpose of the study was to determine the nature of the relationship between recognition and presentation time. They were asked to write down the name of the celebrity which the target face represented following each trial from the list given. Trials were given with the colour and grey-scale versions of the same images presented in random order. Trials were given in blocks of 20 at the same presentation duration such that in each block each of the 10 celebrity faces was presented at random in the two image conditions. Three presentation durations were used - two, four and six frames. The masks were presented for the same durations as in the introductory phase. After a subject had viewed each image at all three presentation times, they had seen each image a total of six times, three times under each image condition. Order of presentation was counter-balanced such that two subjects were presented with the blocks of trials in each of the six permutations of ordering for the three sets. This meant that for two subjects, during the first block of 20 trials, images were presented for a duration of two frames; in the subsequent block of trials the images were presented for four frames; finally, in the third block of trials images were displayed for six frames. For another two subjects, the order of presentation was altered so images were presented in the reverse order - six frames, four frames and two frames and so on. In this manner any order effects were counterbalanced.

After the task was complete, each of the colour images was presented in turn and subjects were asked to rate each on a scale of 1 to 5 (where 1 indicated highly unrepresentative and 5 indicated highly representative) as to how close to their personal representation of the relevant celebrity they believed the images to be.

Results and analysis

The post-task questionnaire showed that all the faces were regarded as good representations and they were used collectively in further analysis. Moreover, each subject was familiar with each of the celebrity faces. These scores were collated separately for each of the three stimulus durations. Results for all faces were used in further analyses.

A within-subject analysis of variance was carried out on the accuracy data, using the same transform as described for Experiment 1 (*Chapter* 2), with duration (3 levels), type of image (2 levels, colour/grey-scale) and gender of face (2 levels) as main factors. This analysis found a significant main effect for duration ($F_{2,22}$ = 29.86, $p < 0.0001$, mse = 0.13) and type of image (*Figure 3.1*, $F_{1,11}$ = 5.29, $p < 0.05$, mse = 0.04) with colour images being recognised more accurately than grey-scale versions of the same face images. There was no main effect of stimulus gender ($F_{1,11}$ = 0.48, p=0.51, mse = 0.04). No interaction was found between duration and type of image (*Figure 3.2*, $F_{2,22} = 0.21$, $p=0.82$, mse = 0.03). Mean accuracy increased with stimulus duration with 24% mean accuracy at 33 ms, 63% at 67 ms and 74% at 100 ms.

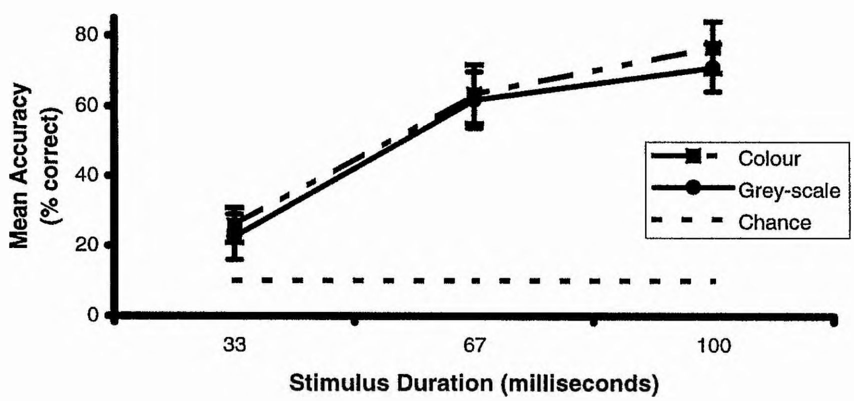


Figure 3.1. Effect of colour presence on accuracy of face recognition. There was a small (~4%) but significant improvement in recognition accuracy for the images with the colour information included. This was found at all lengths of duration tested.

Discussion

Introducing the colour information relevant to a given person's face increases the ability to recognise that face. This effect does not seem to depend on stimuli duration, at least over the range tested (33 to 100 milliseconds). The use of a sepia mask is a potential confounding factor as although the intensity and saturation of the colour in mask is equidistant between grey-scale and colour this may not be the case in the processing done by the visual system. A more rigorous approach would be to use two masks in a counterbalanced stimuli such that colour stimuli were followed by a coloured mask, black and white stimuli followed by a black and white stimuli, and vice-a-versa. Unlike the previous experiment, the order of presentation here was fully-counterbalanced. Despite some of the stimuli being shown first at the longest duration a benefit for recognition with colour stimuli was still found at all durations, indicating that order-effects are not a problem.

The general increase in accuracy with colour stimuli was small at approximately 4% but consistent across all three display durations tested. Therefore, colour information in the form of hue and saturation would appear to have importance for the recognition of familiar faces. The role of colour may only be evident when the process of recognition is challenged by a form of stimulus degradation such as brief stimulus duration.

Plate 3.1. The cropped full-colour image of each celebrity was also converted to grey-scale in order to determine whether colour information aided facial recognition.

Plate 3.2. The ten colour images used in Experiment 2 were normalised and cropped so as to replace the background with uniform black and the images were normalised to the same interocular position.

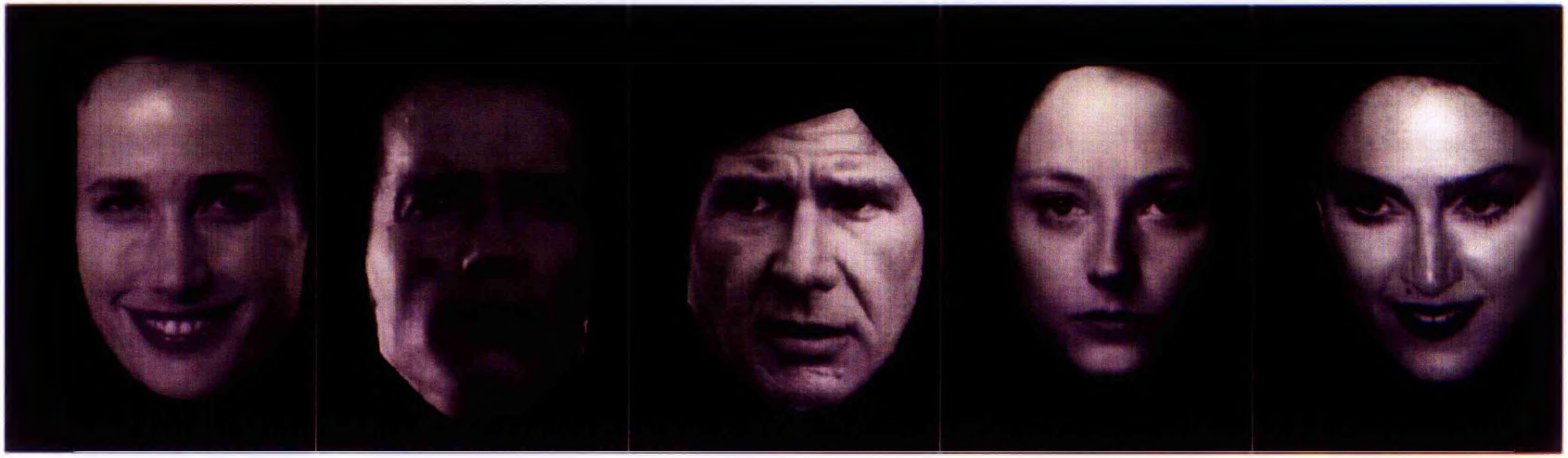
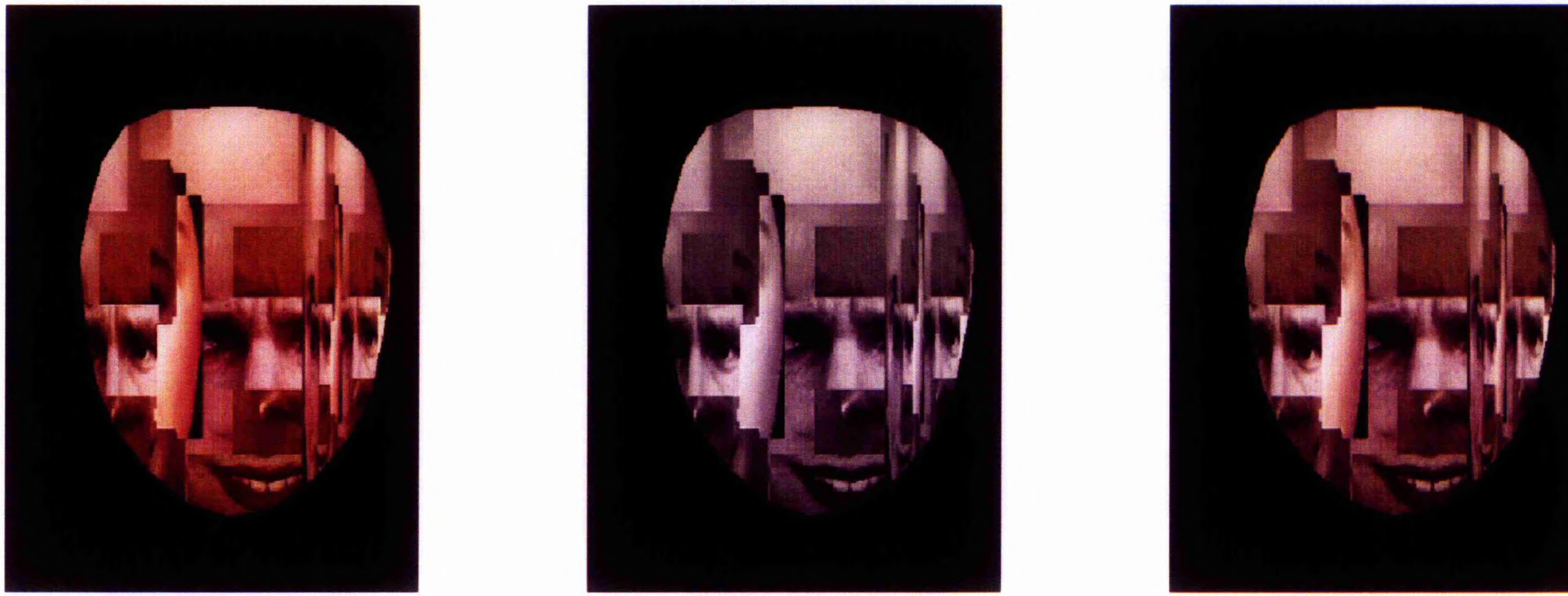


Plate 3.3. Each of the ten colour stimuli used in Experiment 2 was converted into a 256 shade grey-scale image. Subject performance was compared between the colour and grey-scale stimuli.



a b c

Plate 3.4. (a) A new mask was designed by placing features from a number of faces at random on the background of an individual face; (b) this mask was converted to a 256 scale grey-scale image; (c) a sepia mask was created by averaging the colour information of each pixel of the colour and grey-scale versions of the mask so that the RGB value for each pixel of (c) is equidistant from the value for the corresponding pixels of (a) and (b)- thus neither the grey-scale and colour images used in Experiment 2 could benefit from being closer in colour to the mask.

Chapter 4 - Experiments Three and Four: Perception Time and Colour Caricature

Introduction

Caricaturing studies to date have kept the intensity, colour and textual information in the faces to be viewed constant. Valentine (1991) and Valentine and Endo (1992) have proposed the notion of a facial norm which encompasses more than just relative differences in the spatial configuration of faces. They suggest that faces can be represented within a multidimensional "face-space" where the dimensions correspond to different physical dimensions along which faces differ. In these accounts, the assumption is that faces are distributed normally across each of the dimensions but the nature of these dimensions is not specified - "any attribute that can serve to discriminate faces would form a dimension of the space" (page 673; Valentine and Endo, 1992). Thus, faces are represented in this space by reference to a facial norm. This norm need not be restricted to the shape domain; it could convey information regarding any visible facial details including surface texture and colour. Caricaturing shape in two-dimensional (or three-dimensional) space accentuates the distinctive feature shapes and configurations. Caricaturing the colour and intensity differences of an individual enhances the difference between the individual and the prototype and will affect a variety of cues, including texture, such as characteristic freckles, lines, spots, stubble, feature pigmentation and shading, which provides information pertinent to three-dimensional form.

The impact of caricaturing has so far been assessed only in the shape domain. In Experiment 3 it was investigated whether caricaturing in RGB colour space would benefit recognition of familiar faces, since colour may aid differentiation of faces. Experiment 3 was replicated in Experiment 4 using revised masking procedures, control stimuli and different faces. Previous experiments using photo-realistic depictions of faces have predominantly employed black and white images yet, as described in the previous chapter, information related to identity appears to be present in the colour domain.

General Experimental Methods

Apparatus. A Silicon Graphics 4D-25 Personal Iris computer was used with a non-interlaced monitor synchronised at 60 Hz; an Epson GT-6500 scanner was connected to an Escom 486 DX2/66 IBM-compatible personal computer.

Stimuli. The facial areas of photographs of 41 actors and 22 actresses were scanned at a resolution of 517 pixels in width by 720 pixels in height at 150 dpi in 24-bit colour. Individual stimuli used in the experiments consisted of a subset of these images. Each of the individuals was clean-shaven and wore no spectacles. The features on the faces were defined using 179 feature points placed manually in standardised positions.

Experiment 3

Introduction. Experiment 3 sought to investigate whether enhancing colour and contrast of a face relative to an average face improved recognition accuracy.

Methods

Subjects. These were 28 undergraduate students at the University of St Andrews; 15 females and 13 males.

Stimuli. A questionnaire was given to seven undergraduate students at the University of St Andrews. From a list of 24 well known male and female celebrities, these students were asked to rate on a scale of one to five how familiar they believed themselves to be with the face of each of these celebrities. Each student was then shown several images of each of the five male and five female celebrities that they had felt to be the most recognisable. For each celebrity, the student was asked to put into rank order how well they thought each image of the individual celebrity represented their personal impression. From this survey, 10 images - each of an individual celebrity, five males and five females - were chosen to maximise familiarity and representativeness of the image for use in the study.

In each image, the internal features from the face were cropped from the hair, ears and background so that these latter items could not support recognition. The internal features were then placed on a black background.

Each of these veridical images was now "caricatured" in the colour domain against a gender relevant prototype created from a blend of 22 actors or actresses (*Plate 1.4*). This caricaturing process is analogous to the shape caricaturing technique of previous studies and was used to enhance the colour differences between the individual image from the prototypical blend by 40%. A shape caricature is defined as an exaggeration of the differences between an individual face shape and the average face shape. Colour caricatures are defined in the same way. The colour caricaturing process was carried out by warping the shape of the relevant prototype face image to that of the target face image. For each corresponding pixel, 40% of the difference in RGB values between the target image and the prototype was then added to the target image. The result is a colour-caricatured image with the colour and intensity difference between the target image and the prototype exaggerated by 40% (*Plate 4.1*; see Burt and Perrett, 1995; Rowland and Perrett, 1995 for a more detailed description). A caricature level of 40% was chosen based on the levels of caricature found to work in shape and to avoid individual pixels exceeding the available intensity range (0-255) during caricaturing. A contrast control was also created by enhancing by 40% the differences in RGB values between the individual face image and a uniform 'grey box' - a mid-grey rectangle with RGB values (127,127,127). Here, the co-ordinates (R,G,B) represent the intensities of the red, green and blue intensity values where 0 indicates zero intensity and 255 indicates maximum intensity. The contrast control comparison is digitally equivalent to turning up the colour saturation and contrast on a television set (*Plate 4.2*). The shape of each face image was left unaltered. Each of the 30 images - 10 veridical images, 10 colour caricatured images and 10 contrast-controlled images - was normalised (i.e. rotated, scaled and translated) to have the same eye-centre positions.

Caricaturing colour uses a relevant base (e.g. an average male) from which to enhance the information for a distinctive individual. The contrast control enhances colour saturation and lightness contrast but not in a way which accentuates all the information which is distinctive of the individual within its class. Indeed this can be demonstrated numerically by considering changes to particular features of an image in veridical, colour caricatured and contrast control form relative to the prototypical blend. This analysis shows the way in which caricatures differ from the contrast controls.

For instance, Table 1 describes the relative means of the different red, green and blue colour components for an eyebrow of one of the targets (Kim Basinger). Here, the value of each colour component can range from 0 to 255; thereby a pixel with the components (0,0,0) would be black, while one with the components (255,255,255) would be white. Thus, a pixel with a value of 127 for each of the red, green and blue components would appear as grey.

From this Table, it is clear how caricaturing a veridical image away from an average blend alters the colour of the image. RGB values in the veridical are greater than in the prototype but much smaller than in the grey box. In the colour caricature, RGB values increase but in the contrast control the RGB values decrease. Not only does the direction of intensity change differ between caricature and contrast control but also the magnitude of the change is greater with the contrast control. This is because the veridical is much closer to the prototype in colour-space than to the grey box.

Therefore, in the contrast control (caricaturing from a 127,127,127 grey box) the mean RGB values are much smaller as there is a smaller difference between the mean values of the blend and veridical than between the grey box and the veridical.

	GREY BOX	PROTOTYPE	VERIDICAL	CARICATURE	CONTROL
RED	127	91.4	117.7	127.9	107.1
GREEN	127	54.9	67.5	72.2	37.0
BLUE	127	40.8	44.0	45.0	11.0

Table 4.1. Mean of red, green and blue values in the left eyebrow - (46x5 pixel area) of one female target face.

This numerical analysis has been replicated in HLS (hue, lightness and saturation) space. Lightness can be estimated by summing the RGB values. Bright pixels have high RGB values, while darker pixels have low RGB values. In the veridical image, the eyebrow area for this image is lighter than that of the prototype. Caricaturing that area increases the lightness of the area and the eyebrow becomes even lighter. Thereby, the lightness difference of the eyebrow relative to an average face is accentuated. The eyebrow area is, however, darker than mid-grey. Thus the contrast control manipulation of the image reduces lightness in the area and the eyebrow becomes darker than the veridical. In this way, the contrast control and caricaturing process can be seen to produce different directions of lightness change. Therefore the eyebrow region of the contrast control becomes more average and moves toward the prototype and becomes less representative of the target individual. For other faces, colour in the eyebrows and other features would change in different ways.

Two mask stimuli were also created. One was a blend in both colour and shape of 60 professional models (male and female), where the individuals used were neither individual test stimuli nor incorporated in the prototypical blends used for caricaturing the test stimuli (see *Plate 4.3*). A second mask was made in the same shape as the first mask, but the colour information was produced by blending the RGB colour values for each pixel from each of the 10 original individual stimuli (see *Plate 2.3*). Again, these masks were cropped around the internal face area. The mask and target face stimuli all retained the dimensions 517 x 720 pixels.

Design. There were three independent variables in this experiment: image type, duration and mask stimulus. There were the three types of image: veridical images, colour-caricatured caricatured images, and contrast-enhanced images.

Images were presented over four durations (33 milliseconds, 67 milliseconds, 100 milliseconds, and 133 milliseconds). Images were presented in blocks with the display duration increasing from block to block. This was done in order to prevent extended exposure to the stimuli to influence recognition accuracy. The dependent variable investigated was the accuracy of recognition for individual faces when presented (with either image type and at any of the durations of presentation). A within-subject design was employed such that the same subjects saw all the stimuli as a veridical image and under both types of manipulation, displayed at all four durations. Mask stimulus was a between-subject factor. In this manner, all subjects viewed every image in all three of its forms and at all four durations but half the subjects were presented with one masking stimulus while the other half were presented with another.

Procedure

Introductory phase. Subjects were familiarised with task procedures. This involved displaying two images of faces not included in the main experiment on the screen in uncaricatured veridical form. These images were shown at decreasing display durations. Subjects were first shown the single images at 8 frames, then 6 frames, 4 frames and finally 2 frames, where the duration of a frame was 16.7 milliseconds. In each instance, the mask stimulus was first presented for 500 milliseconds followed by the test face image, followed again by the mask for a further 500 milliseconds. Here the mask initially acted as a signal to indicate the onset of a target image. It also functioned as a forward and backward masking device. This phase was used in order to alert subjects to the nature of the task and the mask and, particularly, that the image that they were to name was the individual target image rather than the mask stimulus.

Test phase. Subjects were given a list of the names of the ten target faces to be presented in the task and asked to rate familiarity with celebrities' faces. They were asked to rate each of the ten celebrities on a scale of 1 to 5 to indicate how familiar they believed themselves to be with a given individual's face, where 1 indicated completely unfamiliar and 5 indicated a high degree of familiarity.

Procedures were as for Experiment 2 (*Chapter 3*). Trials were given with veridical images, colour caricatured images and contrast control images presented in random order. Trials were given in blocks of 30, with 10 images representing each face from each of the three image conditions. During the first block of 30 trials, images were presented for a duration of two frames. In each subsequent block of trials the images were presented for a further two frames at a time. Four blocks of trials were presented in total, the last in which images were displayed for eight frames. Fourteen subjects were presented with the first mask stimulus during the course of the entire experiment, while the other fourteen subjects were presented with the second mask.

After the task was complete, each of the veridical images was presented in turn and subjects were asked to rate each on a scale of 1 to 5 (where 1 indicated highly unrepresentative and 5 indicated highly representative) as to how close to their personal representation of the relevant celebrity they believed the images to be.

Results and analysis

Medians were found for the post-task questionnaire to indicate how good the likenesses were seen to be by the subject population used. Three celebrities were eliminated from further analysis here because they had medians of 1 or 2 from the 5 point scale of typicality. The remaining seven celebrities (3 male, 4 females) with median typicality ratings of 4 were used collectively in further analysis (see *appendix*).

Subjects were scored for correctly identifying an image representing an individual celebrity on any given trial. The results were tallied separately for veridical, colour caricatured and contrast control images. These scores were marked separately for each of the four blocks of trials.

Recognition accuracy was subjected to a between- and within-subject analysis of variance, using the transform described for the experiments in the previous two chapters, with mask (2 levels), duration (4 levels) and type of image (3 levels) as main factors. This analysis found a significant main effect for type of mask ($F_{1,26} = 12.38$, $p < 0.005$, mse = 0.26), duration ($F_{3,78} = 132.54$, $p < 0.0001$, mse = 0.05) and type of image (*Figure 4.1*, $F_{2,52} = 4.58$, $p < 0.05$, mse = 0.02). The different masks were tested using different subject pools so these differences could be attributed to general differences between these two groups. No interaction was found between type of mask and either type of image or duration so it was inferred that the results from both masks could be considered collectively. An interaction was found between type of image and duration ($F_{6,156} = 2.28$, $p < 0.05$, mse = 0.02). A posteriori PLSD analysis of this interaction found greater accuracy for colour caricatured stimuli compared to either veridical stimuli or the contrast control stimuli (*Figure 4.2*, $p < 0.05$, each comparison) at durations of 67 ms and 100 ms, while there was no difference found between the veridical or contrast control stimuli at any duration ($p>0.05$, each comparison, *Figure 4.3*). *Figures 4.2 and 4.3* show performance for presentation durations of 33 ms was close to chance level suggesting a floor effect at the shortest stimulus duration.

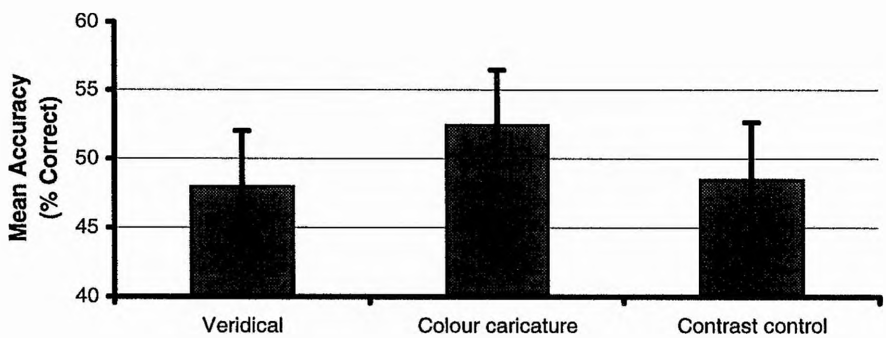


Figure 4.1. Recognition accuracy with manipulation of colour. A main effect for type of image was found across all display durations. It is evident that the only difference in accuracy is an improved discrimination for the colour caricature over the veridical and contrast control stimuli.

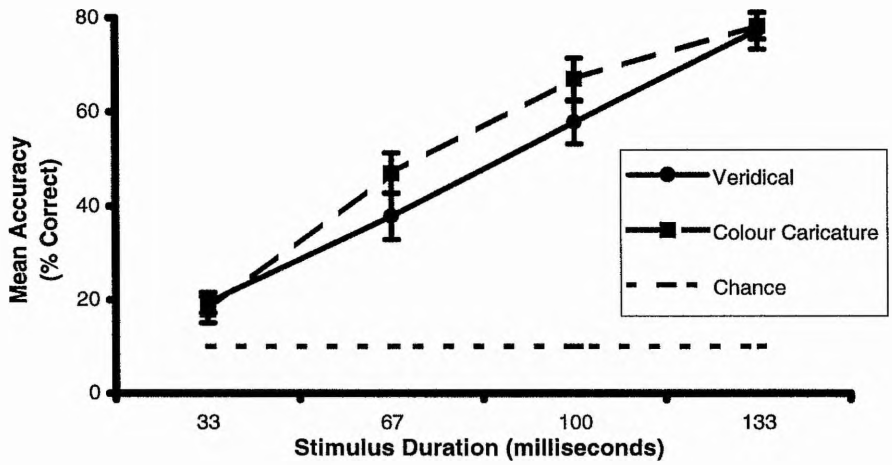


Figure 4.2. Variation in impact of caricaturing with stimulus duration. Accuracy for naming faces was significantly higher for the colour caricatures than the veridical images at 67 ms and 100 ms presentation time but not at longer or shorter durations.

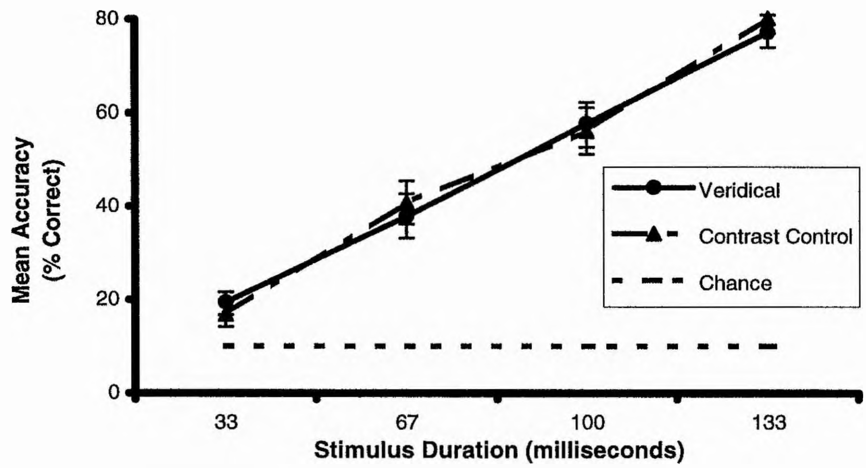


Figure 4.3. Impact of stimulus duration on recognition of contrast control. Enhancing contrast and colour saturation in the contrast control stimuli made no significant difference to accuracy for naming faces compared to the veridical images at any durations.

Discussion

Enhancing the colour information relevant to a given person's face increases subjects' ability to recognise the individual when an image is presented briefly (67 and 100 ms). Conversely, simply increasing the lightness and colour saturation contrast of an image does not increase ability to recognise an image at a brief display time. Caricaturing in colour was not found to aid recognition for familiar faces at the shortest display period employed (33 ms). The benefits of enhanced colour information on recognition performance were also attenuated at 133 ms. The diminishing of benefit at the shortest presentation can be attributed to a floor effect. For long presentation times other cues such as shape information may have more importance for recognition. As discussed with the shape caricaturing experiment in Chapter 2, it is feasible that differences between a face-like mask may have given rise to the recognition advantage found with colour-caricatures. This is suggested by the difference in general recognition performance between the two masks (although this is a between-subject factor). Further, this experiment also failed to employ proper counter-balancing in order to determine whether order of presentation may have been partially responsible for poorer recognition performance at brief presentation times. The experiment was therefore replicated in order to exclude these potential confounds.

Experiment 4

Introduction. Experiment 4 sought to replicate the findings of Experiment 3 using an improved mask and control procedure.

Methods

Subjects. These were 44 undergraduates at the University of St Andrews; 32 females and 12 males.

Stimuli. Ten veridical images, none of which had been included in Experiment 3, were caricatured in the colour domain. Each of these images was of a famous male actor and each was caricatured against a relevant prototype created from a blend of 41 male actors (*Plate 4.4*). The use of male actors alone here provides a further degree of rigour as age, sex and occupation may all play a role in cues to recognition. As in Experiment 3, the colour differences in the individual image were exaggerated from the prototypical blend by 40%. A modified form of contrast control compared to that used in Experiment 3 was also created. The mean intensity for each component in the RGB triple was calculated for both the veridical image and the colour caricature. A contrast control was then created by enhancing the RGB differences between the individual image and a neutral grey box. The amount of enhancement in the contrast control was such that the mean intensity difference for each of the co-ordinates (R,G,B) between the veridical image and the colour caricature was the same as that between the veridical image and the contrast control. This procedure prevents excessive contrast in the control with values numerically greater than 255 or less than 0. Again, the shape of each image was left unaltered. Each of the 30 images - 10 veridical images, 10 colour caricatured images and 10 contrast control images - was normalised to the same eye-centre positions.

The images were cropped so that the background, hairline and ears were removed and a black background was substituted in place. A new mask stimulus (*Plate 4.5*) was created by inverting a male face and placing random features across the image. This mask was not cropped so that faces could not be identified as a result of the position of the border relative to the mask. In this case, since the mask was not face-like it would also not be possible for each target face to be recognised in terms of facial differences from the mask. As with Experiment 3, the mask and target face stimuli had dimensions of 517 x 720 pixels.

Design. There were two independent variables in this experiment. The three types of image - veridical images, colour-caricatures, and contrast-enhanced images - and the two durations over which the images were presented (67 milliseconds and 100 milliseconds). Images were presented in blocks with the order of display duration counterbalanced between subjects. This was done in order to exclude any potential order effects and effects derived from repeated exposure to the stimuli. The dependent variable investigated was the accuracy of recognition for individual faces when presented (in either veridical, colour-caricatured or contrast-enhanced form and at either of the durations of presentation). A within-subject design was employed such that the same subjects saw all the stimuli in all three forms and displayed at both durations.

Procedure

Introductory phase. The introduction phase was similar to that of Experiment 3 except for use of the new mask stimulus.

Test phase. Subjects were given a list of the names of the ten target faces to be presented in the task and asked to rate familiarity with the celebrity faces. Instructions were the same as for Experiment 3.

Trials were given with veridical images, colour caricatured images and contrast control images presented in random order. Trials were given in blocks of 30, with 10 images from each of the three image conditions presented. Two exposure durations were employed (four and six frames). Order of presentation for the images was counterbalanced across two groups of subjects; for one group presentation time increased across blocks of trials while for the other group presentation time was decreased across block of trials. Thus for one group, during the first block of 30 trials, images were presented for a duration of four frames and in the subsequent block of trials the images were presented for six frames. This order was reversed for the second group.

A post-task questionnaire was given as in Experiment 3.

Results and analysis

One celebrity was eliminated from the analysis because the likeness in the post-task questionnaire was judged with a median of only 2 out of 5. The remaining nine celebrities were used collectively in further analysis (see *appendix*). Nine subjects (eight females, one male) were also removed from the analysis because they proved unfamiliar with the celebrities. They rated overall familiarity with the celebrities at a median and mode less than three. In two of these cases, subjects reported never having seen over half the celebrities.

A within-subject analysis of variance, using the same transform as employed for Experiment 3, was carried out on recognition accuracy across duration (2 levels) and type of image (3 levels). This analysis revealed a significant main effect for both duration ($F_{1,34} = 11.98$, $p < 0.005$, mse = 0.06) and type of image (*Figure 4.4*, $F_{2,68} = 6.56$, $p < 0.005$, mse = 0.01). A posteriori PLSD analysis of the main effect of image type revealed significantly greater accuracy for colour caricatured stimuli compared to either veridical stimuli or contrast control stimuli ($p < 0.05$, each comparison), while there was no difference found between the veridical and contrast control stimuli ($p > 0.05$, each comparison). As in Experiment 3 the advantage of colour caricatures over veridical images and contrast enhanced images was evident at both exposure durations (67 and 100 ms).

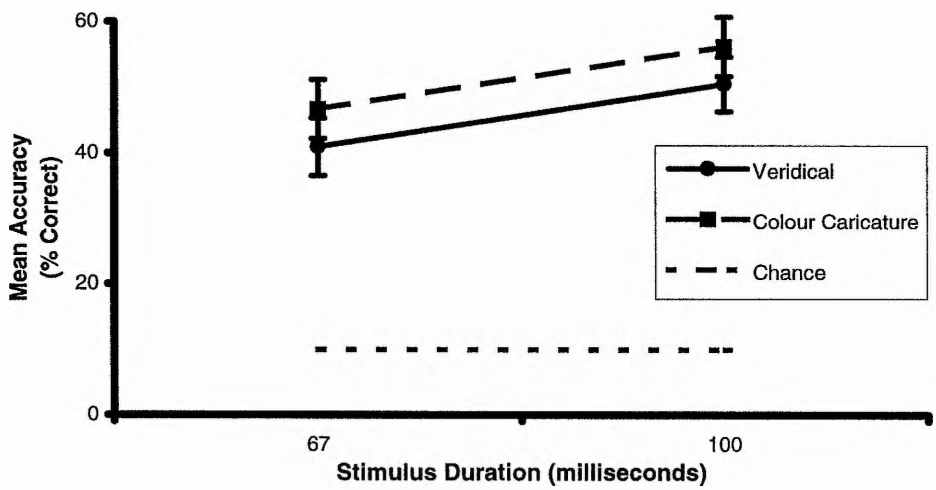


Figure 4.4. Accuracy in naming face images. The only difference in accuracy for discrimination between conditions is with the colour caricatures over the veridical and contrast control. This was consistent across stimulus duration.

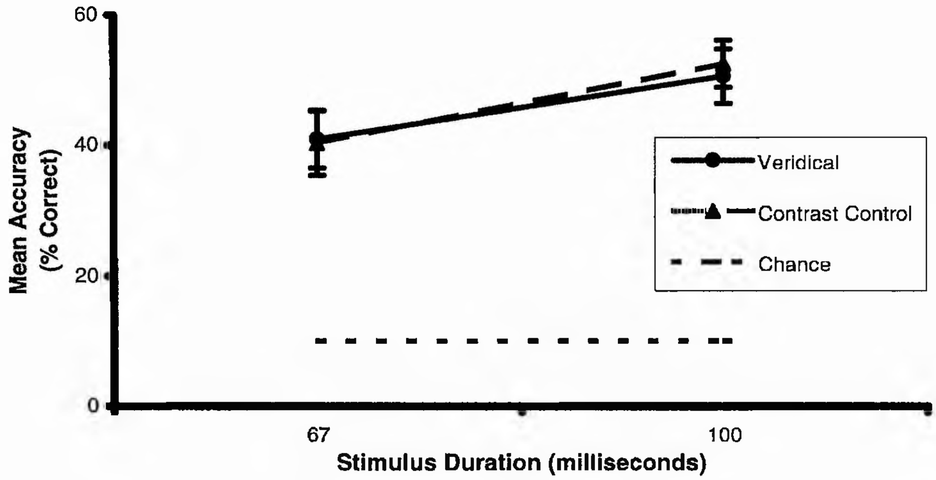


Figure 4.5. Impact of duration on recognition of contrast control. As in Experiment 3, enhancing contrast and colour saturation in the contrast control stimuli made no significant difference to accuracy for naming faces compared to the veridical image at either duration.

Discussion

Here, as in Experiment 3, enhancing the colour information relevant to a given person's face was found to increase subjects' ability to recognise the individual. Thus the benefit of colour caricaturing was replicated with new subjects, faces and masks. Conversely, simply increasing the colour saturation and lightness contrast of an image does not increase ability to recognise an image at brief presentation times. The effect was observed at both the durations to which the subjects were exposed so there is no indication that this benefit for colour requires previous exposure to the target images. The new mask also did not appear to affect the performance and in this situation it could not be argued that faces might be identified by the position of the cropping boundary relative to the mask. A further development to ensure the propriety of the mask stimuli would be to use a variety of different masks, such that some of the masks had boost contrast to investigate if this produced a difference in recognition performance with the different types of stimuli. In the findings here, there was also no difference in overall lightness contrast between the colour caricatured and contrast control stimuli so this could not be a factor which would account for preferential recognition for the colour caricatures.

General discussion

These experiments found that enhancing differences in RGB colour between veridical images and a relevant face prototype aided recognition of identity. These results, in conjunction with those found in the previous chapter, support the suggestion that colour information is helpful when differentiating between stimuli which have highly similar configurations and are natural objects (Price and Humphreys, 1989; Wurm *et al*, 1993; Humphrey *et al*, 1994).

It should be noted that the impact of colour in these experiments is relatively modest. In the experiment described in the last chapter, colour presence accounted for an improvement in accuracy of 4% in absolute terms or 10% relative to baseline performance for achromatic faces (52% with and chance level of 10%). In the experiments described here enhancing colour information improved performance by 17-18% relative to baseline performance for veridical images defined in the same way. This indicates that face recognition can rely on sources of information other than colour as recognition accuracy is substantially greater than chance in all conditions.

The results here indicate that further work investigating cues to face recognition should concentrate on relevant colour information in addition to shape information. While prior experiments involving photorealistic images have used black-and-white images, it is clear here that information can also be gleaned from the colour domain.

Short duration presentations

The presentation time paradigm introduced for Experiment 1 (*Chapter 2*) has proved to be an effective tool for exploring the effects of manipulations in the colour domain on recognisability of faces. The effect of caricaturing in RGB colour-space appears at short exposures of 67 and 100 ms in the experiments described here. This effect is robust since colour caricatures were still found to be beneficial for recognition in Experiment 3 at 100 ms even after representations of each face had been presented in three different conditions (veridical, colour caricature and contrast control) on two previous blocks of trials (33 and 67 ms). The results found here indicate the usefulness of varying presentation time as a measure of the effect of caricaturing. Caricature effects have not been clearly demonstrated for photographic images using reaction-time paradigms where images are displayed for longer periods of time (Benson and Perrett, 1991).

Image degradation and caricature advantage

In Experiment 3 no differences between type of image were apparent at the longest exposures used. This highlights the usefulness of the presentation time paradigm. Short duration exposures can be considered a form of image degradation. Indeed Sergent (1984) has argued that brief presentations favour information contained in low spatial frequencies and strategies of visual analysis that rely on judgements of spatial configuration. These conditions may emphasise the configural information thought to be utilised in familiar face processing (Yin, 1969; Sergent, 1984).

Line-drawn stimuli degrade the images of faces and impair recognition through the removal of all textural information (Davies *et al*, 1978; Leder, 1996). Caricature advantages are consistently found with line-drawn faces. Using an interactive caricaturing technique, perceived best likenesses for line-drawings of faces have been found with caricatures at a mean of 42% over veridicality (ranging from 7% to 85%; Benson and Perrett, 1994) while the same degree of caricaturing with photographs is seen to be very unreal (Benson and Perrett, 1993). Thus it seems that at least some degree of degradation is essential in order to detect a caricature advantage, that is caricatures benefit recognition when processing is compromised in some way.

Enhancing any identifiable attribute such as texture, shape or colour may be beneficial at brief inspection periods where there is very little information available to the viewer. At the shortest exposure durations the recognition system is challenged so any extra information that is provided by amplifying relevant image attributes should be beneficial while at longer duration times this supplementary information may become less vital for discrimination. In best-likeness or free-response tasks the unreal nature of caricatured stimuli may actually incline subjects to reject the image as a representation of the individual concerned (for discussion see Rhodes *et al*, 1987; Rhodes and McLean, 1990; Benson and Perrett, 1991). In the paradigm used here, the subjects were presented with a restricted list of the individuals who were represented by the faces to be displayed. Subjects were required only to denote the individual best represented by the image displayed. As a result the 'reality principle', whereby exaggerated images are seen as distortions, would be unlikely to come into play.

Using the same presentation time paradigm, it was found that shape caricaturing can benefit recognition at brief stimulus exposures (see Experiment 1, *Chapter 2*); again the caricature advantage disappeared at longer duration exposures. Thus restriction of a caricature advantage to short exposure durations cannot be attributed to specific idiosyncrasies of colour (e.g. a sluggish parvo-cellular pathway, Livingstone 1988) or texture processing. The shape caricature effect disappears at briefer durations than the effect with colour. This may be because brief durations, masking or both might favour the influence of edge-based rather than surface-based information because edge-based information is instrumental for primal semantic access (Price and Humphreys, 1989).

Caricaturing and contrast

The benefits of colour caricaturing cannot be explained by contrast and saturation enhancement. Indeed the contrast control stimuli were recognised with equal accuracy to veridical stimuli. This lack of effect of contrast enhancement may be due to the caricature process acting at a late stage in perceptual processing by providing more relevant information to representations of familiar faces. Contrast may only enhance early visual processing and may be less important for higher stages of visual processing such as those underlying face recognition. In this context, it is relevant that cells which are responsive to faces show considerable generalisation over contrast and lighting conditions in studies with monkeys (Hietanen *et al*, 1992; Rolls and Baylis, 1986).

The experiments described here could be replicated using a larger number of target faces. This would allow us to see if there is a change in sensitivity to caricaturing in shape and/or colour when more faces are to be discriminated. Shape-caricaturing could also be investigated using grey-scale images in addition to colour images to see if performance is differentially affected when colour information is no longer present. It is likely that shape-caricaturing would be of greater benefit under these circumstances. Caricaturing in shape and colour simultaneously would also indicate whether both components produced a cumulative caricature advantage to recognition. The use of monitors with high refresh rates could allow tachistoscopic presentation of images for even briefer periods of time, enabling the investigation of the caricature advantage without the need for the use of masking. The presentation time paradigm could be further extended to study the effect of caricaturing on discrimination of other perceptual facial attributes such as gender and race.

In conclusion, these experiments have shown that colour information is relevant to the recognition of face identity. It was found that Brennan's (1985) algorithm for caricaturing faces in the shape domain can be extended to the colour domain. As with shape caricaturing, enhancement of distinctive colour and intensity information can increase the accuracy of recognition for famous faces.

Plate 4.1. An example of a colour caricature. Here, the colour differences between a picture of Arnold Schwarzenegger (left) and an average actor have been exaggerated in order to produce a colour caricature (right). This is done by warping the actor composite to the shape of Arnold Schwarzenegger and exaggerating the difference between the corresponding pixels of the images, which are now the same shape, by 40%.

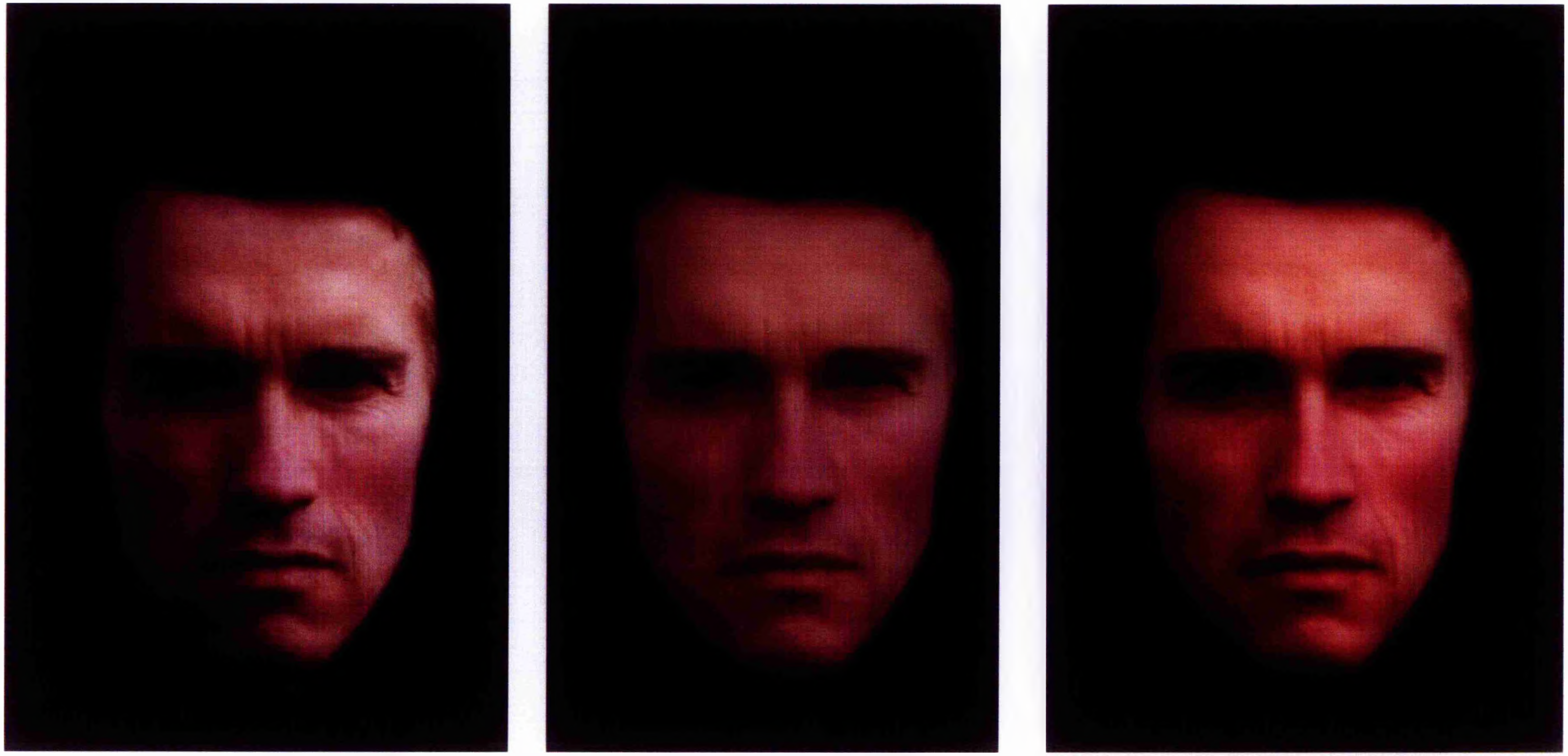


Plate 4.2. A colour caricature (left) was produced by enhancing the RGB differences between a starting image (centre) and a prototype (Plate 1.4, left). A contrast control was also produced (right) which increased the contrast of all the images in a systematic manner. This was digitally equivalent to turning up the colour saturation and contrast buttons on a television set and was performed by colour caricaturing each veridical image against a mid-grey rectangle. For demonstration purposes the starting point here is a blend such that all RGB values are equidistant between those for the prototype and those for the individual.

Plate 4.3. A mask generated for use in Experiment 3. This mask is derived from the average shape and colour information for 60 models, 30 male and 30 female. This mask was the same shape as the mask used in Experiment 1 (and also used in Experiment 3). Like the target stimuli in the experiment and the other mask, the background information and external features were removed and replaced with a black background.

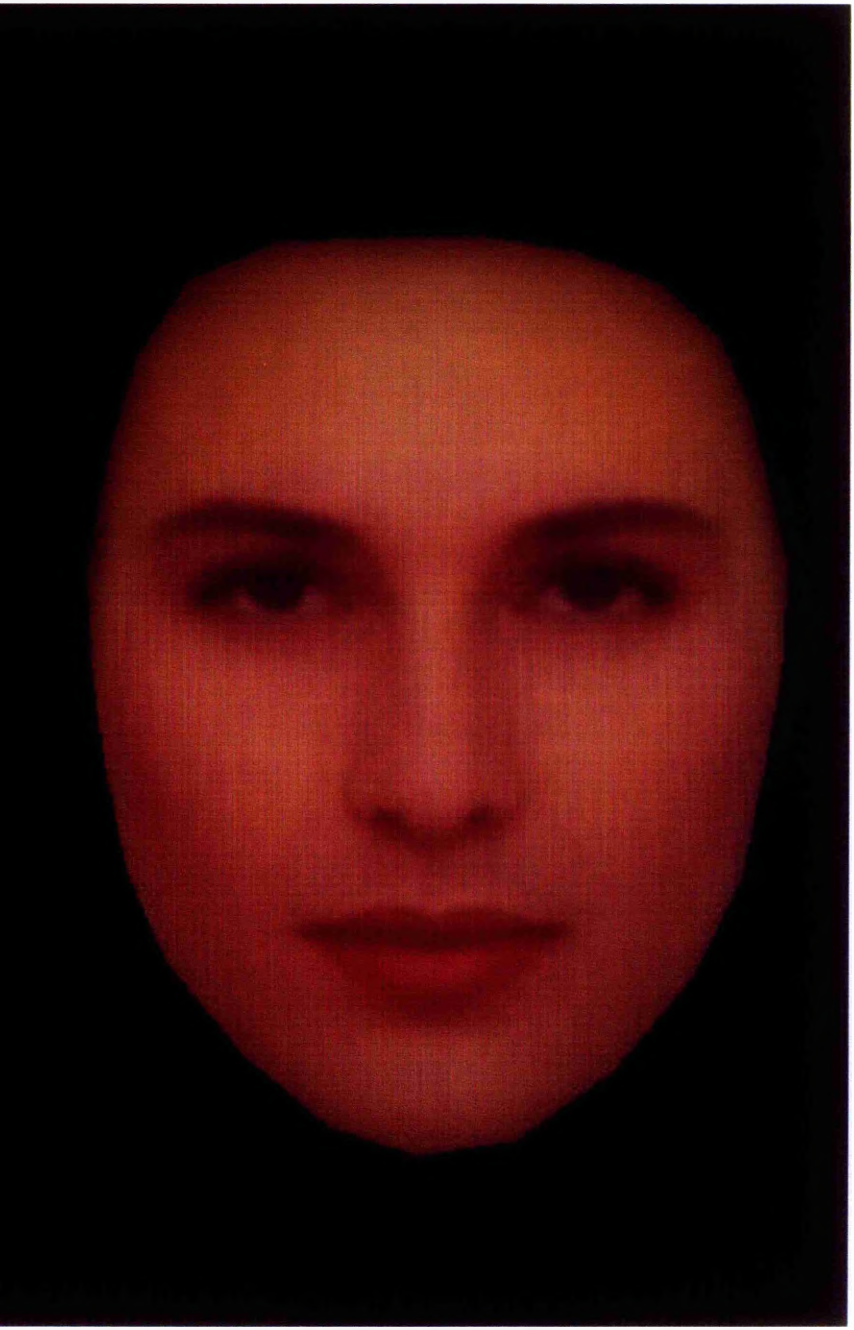


Plate 4.4. For the prototype used in Experiment 4, the shape of each component actor image was warped to the average face shape derived from a set of 41 actors and then the RGB colour value for each pixel was averaged. A greater number of individuals was incorporated to produce the average shape and colour information for this class of face in this blend than in the previous studies.

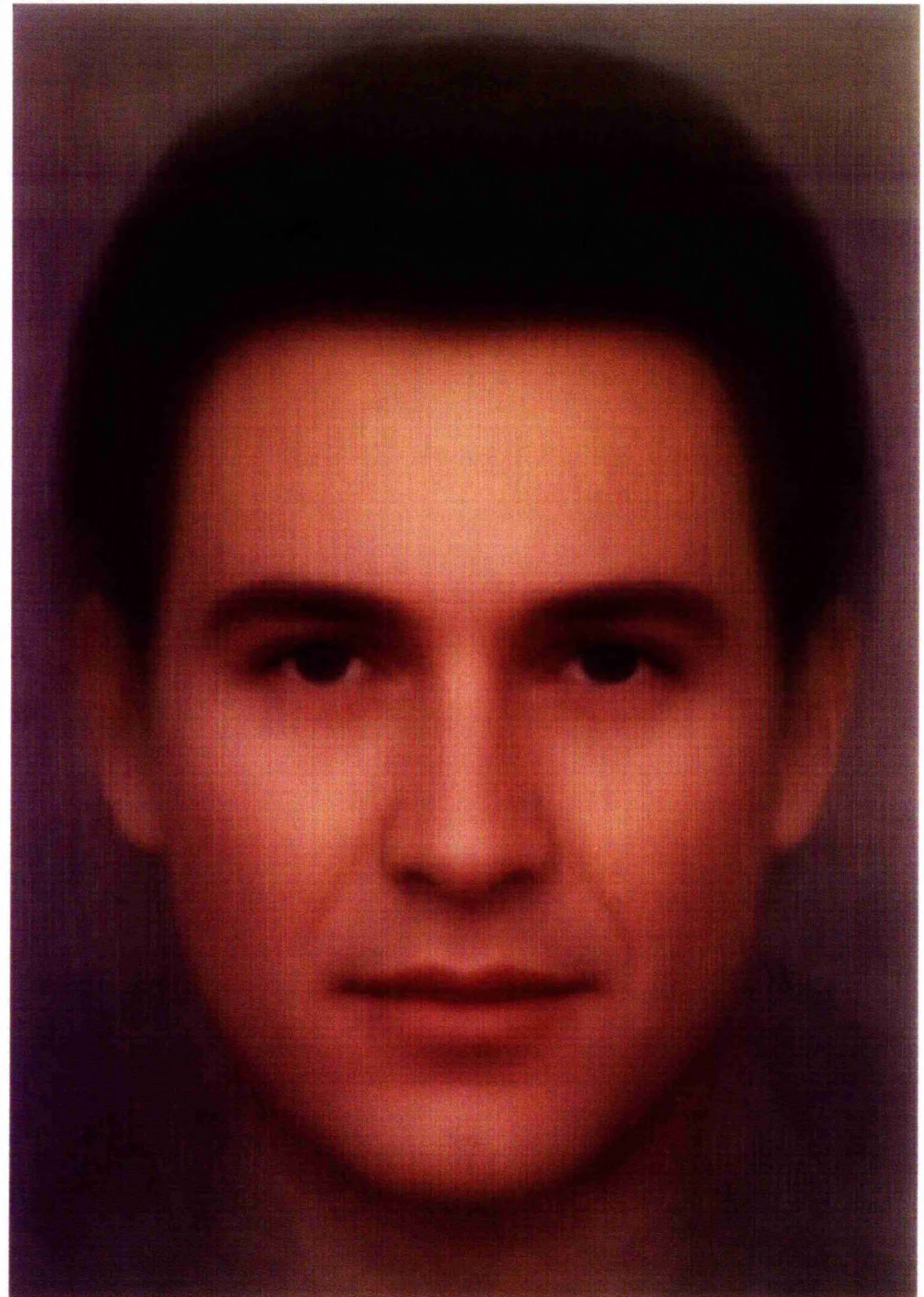


Plate 4.5. A mask generated for use in Experiment 4 consisted of an inverted face covered with several facial elements at random positions and orientations.

Chapter 5 - Experiment Five: Perception Time and Intensity Caricature

Introduction

In the previous chapter it was shown that colour caricaturing (in RGB colour space) increases the distinctiveness of an individual's face relative to a norm. The experiments described there indicate that this enables the image to be recognised more efficiently at brief presentation times. Contrast control procedures manipulating contrast and colour saturation did not increase the distinctiveness of faces. In the same way, colour caricaturing of a middle-aged prototype face against a youthful prototype face increased apparent age, while increasing luminance contrast and colour saturation did not affect perceived age (Burt and Perrett, 1995).

The recognition advantage found by exaggerating the colour information of a given face from a gender-relevant prototype could be attributable to different aspects of RGB colour space - lightness (or brightness), hue and saturation. The lightness component provides textural information relevant to a face and is visible in both colour and grey-scale images while chroma (hue and saturation) defines the actual variation in colour pigmentation. For instance, the pigmentation (dark red mark) on Gorbachev's head may be characteristic. This mark can be made more obvious by enhancing either the chroma information in order to make it more red or by enhancing the difference in lightness in order to make it darker. The separation of these components could be investigated if image manipulations were performed in HLS colour space.

Lightness may be an important cue to identity as varying lighting on a face can provide some three-dimensional face information and photographic negatives of faces have been shown to be more difficult to recognise (Bruce and Langton, 1994; Johnston *et al*, 1992; Kemp *et al*, 1990) than normal photographs. Hue and saturation information may provide cues to the texture of an individual face, in particular, variations in skin pigmentation such as freckles. The role of texture defined by lightness or chroma information could be further investigated by placing individual facial textures on a constant face-shape. Indeed information remaining in so called shape-free faces has been found to be a major contributor to facial recognition in studies of automated face classification using principal component analysis (Costen *et al*, 1996; Hancock *et al*, 1996). Even when faces are warped to an average 2-D shape some luminance cues to three-dimensional face shape for the individual remain from shadows.

The experiment described here attempted to look at the effect of enhancing relevant intensity information for a face without the manipulation of pertinent colour information.

Methods

Apparatus. A Silicon Graphics 4D-25 Personal Iris computer was used with a non-interlaced monitor synchronised at 60 Hz; an Epson GT-6500 scanner was connected to an Escom 486 DX2/66 IBM-compatible personal computer.

Subjects. 15 undergraduates at the University of St Andrews participated in the survey while nine undergraduates took part in the experiment proper; 5 males and 4 females.

Stimuli. Stimuli were created by scanning photographic images of celebrity faces into the computer at a resolution of 150 dpi. None of the celebrities in these images wore glasses or had facial hair. These images were all scanned in full 24-bit colour at the dimensions of 517 pixels in width by 720 pixels in height. A questionnaire was given to 15 undergraduate students at the University of St Andrews. From a list of 158 established male and female celebrities, these students were asked to rate on a scale of one to five how familiar they believed themselves to be with the actual face of each of these celebrities. Each student was then shown several images of each of the five male and five female celebrities that they had felt to be the most recognisable. For each celebrity, the student was asked to put into rank order how well they thought each image of the individual celebrity represented their personal impression.

The grey-scale version of the mask created in *Chapter 3*, developed by taking features from the internal facial area of several faces and placing them at random on a face background, was used (see *Plate 3.4c*).

From the survey, 15 images - each of an individual celebrity - selected from the faces rated as being most recognisable were chosen for use in the study. In this collection of 15 images were eight males and seven females. Each of these images was converted from a full-colour image into a 256 grey-scale image on the Personal Iris using the formula

$$grey = 0.30 * red + 0.59*green + 0.11*blue$$

Each of these images was caricatured in intensity against a gender relevant prototype created from a blend of 22 actors or 22 actresses (*Plate 1.4*). This caricaturing process enhanced the intensity differences in the individual grey-scale image from the prototypical blend by 40 percent. A contrast control was also created by enhancing the differences in intensity between the individual image and a grey box by 40%. This caricaturing process is carried out by adding the difference in RGB values between the target image and the prototype for each pixel to the image by the chosen percentage. This was equivalent to the 'colour-caricaturing' and contrast procedure for the colour stimuli described for Experiment 3, *Chapter 4*.

In each image, the internal features from the face were cropped from the hair, ears and background so that these could not be a confounding factor. As noted, facial hair and glasses were not present in any of the photographs so that this could not be a confounding factor. A black background was then substituted for the former background and external features of each image. Each image was normalised so that that inter-ocular distance was the same as that for the face-basis of the mask. The mask and face stimuli had the dimensions 517 x 720 pixels.

Design. There were two independent variables in this experiment. The three types of image - veridical images, intensity-caricatures, and contrast-enhanced images - and the three durations over which the images were presented (33 milliseconds, 67 milliseconds and 100 milliseconds). Images were presented in blocks with the order of display duration partially counterbalanced between subjects. This was done in an attempt to exclude any potential order effects and effects derived from repeated exposure to the stimuli. Three permutations of order of presentation were given across groups of subjects. The dependent variable investigated was the accuracy of recognition for individual faces when presented (in either veridical, colour-caricatured or contrast-enhanced form and at either of the durations of presentation). A within-subject design was employed such that the same subjects saw all the stimuli in all three forms and displayed at both durations.

Procedure

Introductory phase. Subjects were initially instructed as to the nature of the task in which they were to participate. This involved displaying two target images that were not to appear in the test phase on the screen in uncaricatured form. Subjects were first shown the images at six frames, then at four frames and finally two frames, where a frame was 16.7 milliseconds. The maximum presentation time was reduced from eight frames to six frames as a ceiling effect had been found for this time during Experiments 1 and 3 (*Chapters 2 and 4*). In each instance, a mask was first presented for 500 milliseconds followed by the test face image, followed again by the mask for a further 500 milliseconds. Here the mask acted as a signal to indicate the onset of a target image and as a forward and backward masking device. This phase was used in order to alert subjects to the nature of the task, particularly that the image that they were to name was the individual image and not the mask.

Test phase. Subjects were given a list of the names of the 15 target faces to be presented in the task. They were asked to rate each of the 15 celebrities on a scale of 1 to 5 to indicate how familiar they believed themselves to be with a given individual's face, where 1 indicated completely unfamiliar and 5 indicated a high degree of familiarity.

Subjects were informed that the experimenter would press the computer space-bar. Following this there was a 500 millisecond pause then a mask would be presented for 500 milliseconds followed briefly by a target face followed again by the mask. Subjects were told that the nature of the task was to determine the nature of the relationship between recognition and presentation time. They were asked to write down the name of the celebrity which the target face represented following each trial from the list provided. The experimenter would then press the space-bar again for the next trial. Trials were given with veridical images, intensity caricatured images, and contrast control images presented in random order. Trials were given in blocks of 30. These trials were counterbalanced over three periods of duration - two, four and six frames - such that in each of these blocks of trials each of the 15 images was presented at random in two of the three conditions. After a subject had viewed each image at all three presentation times the subject had seen each image a total of six times, twice under each condition with 15 images from each of two image conditions presented. Order of presentation was also counter-balanced such that for three subjects during the first block of 30 trials, images were presented for a duration of two frames; in the subsequent block of trials the images were presented for a further two frames; finally, in the third blocks of trials images were displayed for six frames. For another three subjects, the order of presentation was altered so images were presented in the reverse order - six frames, four frames and two frames. The last group of three subjects were presented images in the following order: four frames, six frames and finally two frames. In this manner any priming effects could be investigated.

After the task was complete, each of the veridical images was presented in turn and subjects were asked to rate each on a scale of 1 to 5 (where 1 indicated highly unrepresentative and 5 indicated highly representative) how close to their personal representation of the relevant celebrity the believed the images to be.

Results and analysis

The mode was found for the post-task questionnaire to indicate how good the likenesses were seen by the subject population used. Two of the celebrities were then eliminated from further analysis because they had modes of 2 from the 5 point scale of recognizability. The remaining thirteen celebrities all had modes of 4 or 5 and were used collectively in further analysis.

Subjects were scored for correctly identifying an image representing an individual celebrity on any given trial.

A within-subject analysis of variance (which used the transform described for Experiment 1) was carried out on the data with duration (3 levels) and type of image (3 levels) as main factors. This analysis found a significant main effect for duration ($F_{2,16} = 22.48$, $p < 0.0001$, mse = 0.11), but not type of image (*Figures 5.1 and 5.2*, $F_{2,16} = 0.57$, p=0.58, mse = 0.02). No interaction was found between type of image and duration ($F_{4,32}$=1.98, p=0.12, mse = 0.02).

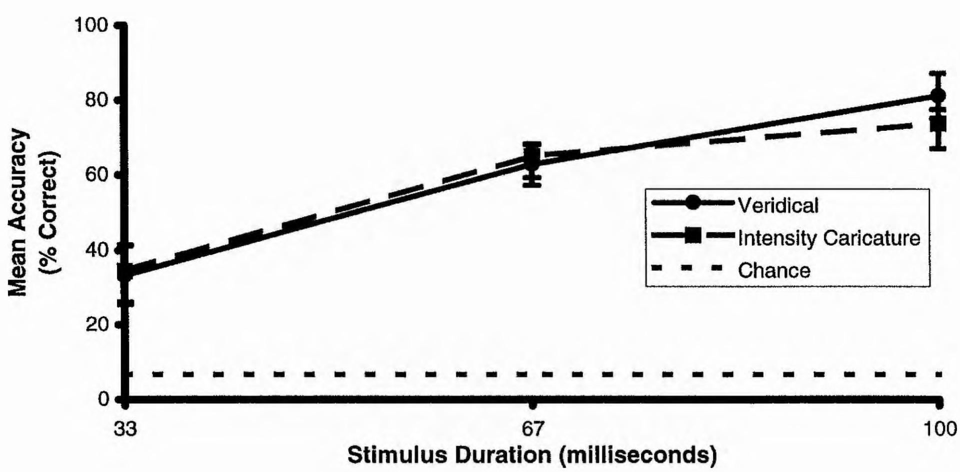


Figure 5.1. Recognition of intensity caricatures at various durations. Enhancing intensity information pertinent to a particular face made no significant difference to accuracy for naming faces compared to the veridical image at any duration.

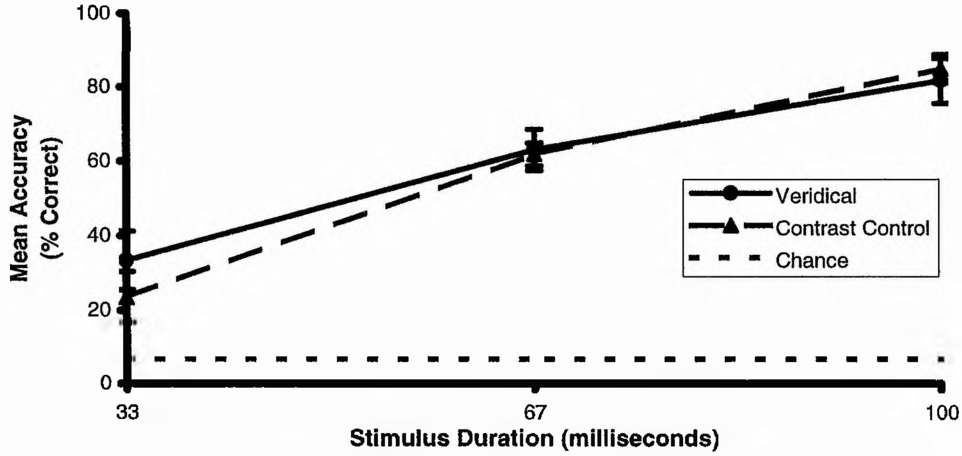


Figure 5.2. Impact of duration on recognition of contrast control. As in the previous experiments (Chapter 4), enhancing stimulus contrast made no significant difference to accuracy for naming faces compared to the veridical image.

Discussion

Neither enhancing the intensity information relevant to a given person's face nor increasing the contrast for a face increases the ability to recognise the individual when an image is presented for a very brief period of time. The latter replicates the result found for Experiments 3 and 4 (previous chapter). The failure to find a benefit for exaggerated relevant information in the luminance domain again shows the importance of colour to recognition. Whereas exaggerating relevant colour in an image gave rise to better recognition accuracy, increase in relevant intensity in a grey-scale image was not found to aid recognition. This suggests that the configural information provided by the enhancement of relevant shading and texture information affected by variation in luminance is not a sufficient aid to recognition with brief presentation times. It is the colour hue and saturation information which is more important under these circumstances.

Perhaps a more sensitive measure needs to be employed to investigate the effects of intensity on recognition. Another, more likely, possibility given the strength of the findings for the studies described in *Chapters 3 and 4* is that the information contained within the hue and saturation information makes a much greater contribution to the discrimination of the faces than luminance information alone. As discussed in *Chapter 4*, the use of several masks of varying intensity across the image might allow to control for any bias the mask may conceivably give to recognition accuracy for a particular type of image. Further, although several orders of presentations were employed here this was not exhaustive and complete counter-balancing of the experiment may potentially have produced a different result. One further possibility is that the intensity was, in fact, caricatured at too strong a degree. Beyond a certain point in the caricaturing process, intensity may actually be topping out across certain regions of the face.

Chapter 6 - Experiments Six and Seven: Interactive Caricature and Best-likeness

Introduction

It is important to note the difference between the requirements for the best-likeness judgements in the studies described here and the caricature advantage for recognition found in the studies described by Rhodes *et al* (1987) and in Chapter 2 to Chapter 5 of this thesis. In those studies caricatures were found to produce a reaction time advantage to recognition (Rhodes *et al*, 1987) or to produce an advantage with caricatures in a presentation time paradigm (Chapter 2 to Chapter 5). Under these circumstances subjects are not required to make any judgement as to the veracity of these caricatures and the caricatured stimuli are merely shown to be an aid to the viewer in discriminating one individual from others. Thus even though a caricature can be recognised quicker than a veridical it is not necessarily seen as a best-likeness. For instance, in the Rhodes *et al* (1987) study the quickest reaction times to recognition were found for 50% caricatures while 25% caricatures and veridical images were found to be the more popular selections for best-likeness.

Rather than use a reaction time to recognition paradigm or a presentation time paradigm, other investigations of the caricature effect have employed techniques which allow subjects to select which image they feel to be the best likeness for an individual (Benson and Perrett, 1991; Benson and Perrett, 1994; Carey, 1992; Rhodes *et al*, 1987; Rhodes *et al*, 1996). Using an interactive caricaturing paradigm where subjects moved a computer mouse from left to right in order to manipulate the degree of caricature or anticaricature of line-drawing representations of famous faces, it was found that on average a 42% caricature was chosen as a best likeness for famous faces (*Figures 6.1 and 6.2*, Benson and Perrett, 1994).

Figure 6.1. In Benson and Perrett's (1994) study using an interactive technique to find best likeness for line-drawn caricatures. The caricatures such as that for Jack Nicholson (right) were found to be better likeness than the veridical line-drawn representation.

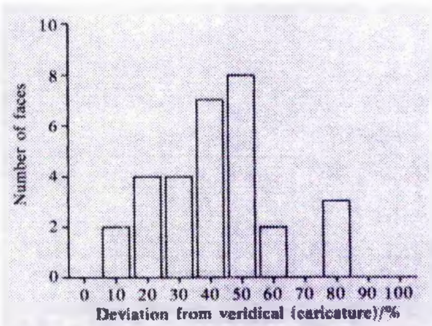


Figure 6.2. From Benson and Perrett (1994). In this study, which used an interactive technique to find best likeness for line-drawn images, an interpolated mean of 42% was found for best-likeness

Faces considered to be more distinctive by experimental subjects were caricatured to a lesser extent whereas with more typical faces, where deviations from the norm where smaller, marked exaggeration was required before a drawing was deemed to be an accurate representation of the selected individual. Thus distinctiveness correlated negatively with the percentage of caricature for best likeness (Benson and Perrett, 1994). In Rhodes *et al*'s (1987) study the best likeness derived by subjective ratings of line-drawing representations of faces was interpolated to be a 16% caricature. In this task, subjects were asked to choose a best-likeness between a 50% anti-caricature, 25% anti-caricature, the veridical, a 25% caricature and a 50% caricature. Line-drawings caricatured by 25% were considered to be equally good likenesses to veridical line-drawings of familiar people. The 50% caricatures and 25% anticaricatures were judged to be good likenesses but poorer than the 25% caricature and veridical (*Figure 6.3*) so although 50% caricatures produced a strong reaction-time advantage, they were judged to be poorer likenesses than veridical representations.

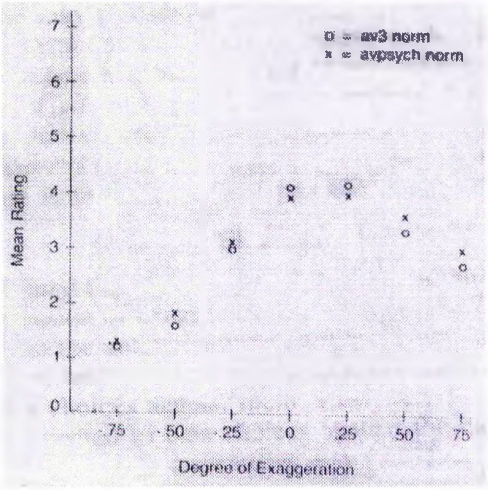


Figure 6.3. From Rhodes, Brennan and Carey (1987). Mean rating versus exaggeration level for each type of norm face. 25% and 50% caricatures were judged to be equally good likenesses of individuals while other caricatures and anticaricatures were seen as poorer likenesses

Benson and Perrett (1991) also found an interpolated best-likeness which was a caricature using black and white photographic stimuli. The degree of caricature interpolated as best likeness in their study was milder at 4.4% with 0% chosen most often (*Figure 6.4*). Benson and Perrett also found the best likeness to show a significant correlation with expert ratings of the quality of the caricaturing process. Rhodes *et al* (1996) found the most popular choice for best likeness was a 10% caricature with an interpolated mean of 11% using a larger range of static line-drawn face images (13) for which the hair, brows and irises and been filled and facial lines such as double-chins and dimples added. Fifteen famous faces were employed in this study, each with caricature levels ranging from 60% anticaricature to 60% caricature (*Figure 6.5*).

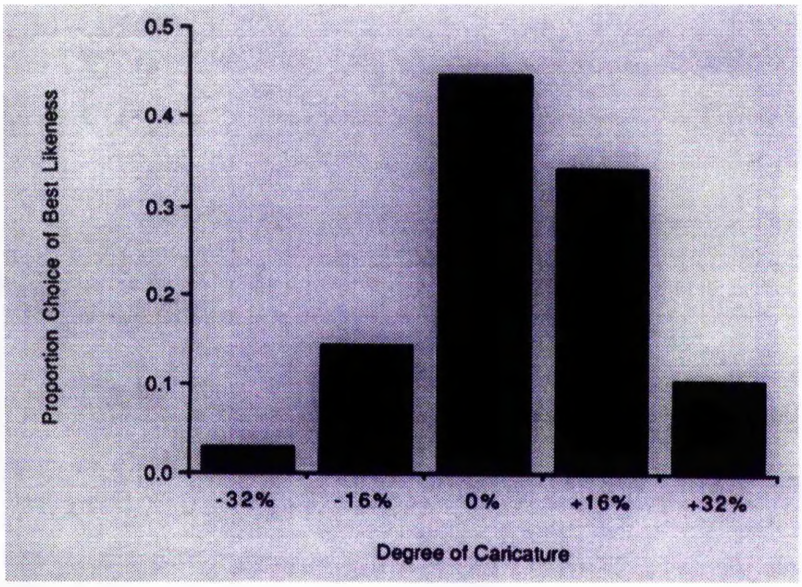


Figure 6.4. From Benson and Perrett (1991). Data in their study showed a preference for caricatures over caricatures with an interpolated mean best-likeness of 4.4%. However, the image chosen most often as best likeness was the veridical or 0% caricature.

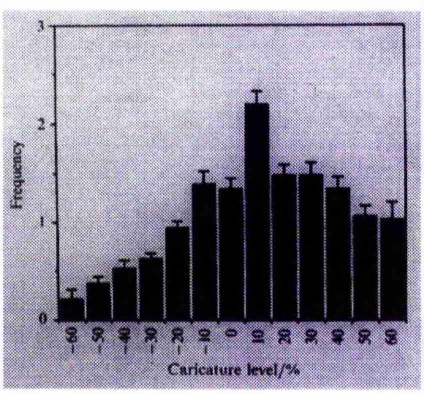


Figure 6.5. From Rhodes *et al* (1996). Mean frequency of best-likeness choices as a function of caricature level with line-drawn stimuli.

Rhodes and McLean (1990) found that a caricature advantage for birds was restricted to recognition at a shorter exposure time using a reaction time paradigm. In their studies subjects were presented with a bird name (e.g. blackbird) followed by a drawing of a bird which remained until the subject responded with a key denoting whether or not this drawing represented the bird named. 25% and 50% caricatures of passerines were recognised with poorer accuracy than the veridical drawings (*Figure 6.6*). The highest levels of caricaturing for best-likeness choices with line-drawn stimuli (interpolated means 42%, Benson and Perrett, 1994; 16%, Rhodes *et al*, 1987) followed by filled-in line drawings (11%, Rhodes *et al*, 1996) and the least exaggeration with photographic images (4%, Benson and Perrett, 1991) in previous studies suggests that a greater amount of information is available in photographic images. A line-drawn figure is unlikely to be representative of a face which inherently has a given texture as exhibited by the skin, eyes and so on. Less exaggeration is therefore required to introduce further information to make the face image more identifiable as the individual it represents. This suggests the involvement of a 'reality principle' such that when viewing a face or bird for a prolonged period of time, judgement as to whether or not the viewed object is an instance of a particular type (e.g. whether a caricatured or uncaricatured drawing of a blackbird is indeed a blackbird) depends on how closely the depiction of the object matches that perceived by the viewer as subjects can distinguish valid instances from those which are erroneous. Nonetheless, in speeded reaction time and presentation time paradigms an advantage is found for the caricatured, "unreal" representations.

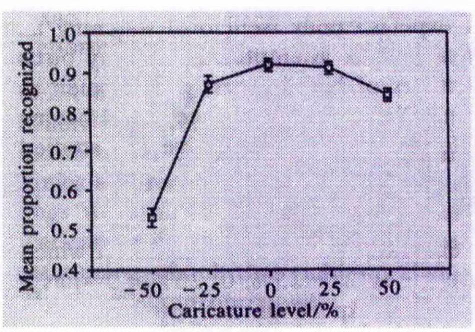


Figure 6.6. From Rhodes and McLean (1990). Mean proportion of drawings recognised as a function of caricature level. 25% and 50% caricatures were recognised with poorer accuracy than the veridical line drawings.

In the experiments described in this chapter the caricaturing process was used to investigate perceived best-likeness for photorealistic images using stimuli similar to that recently used in recognition studies (*Chapter 2* through *Chapter 5*). The paradigm used here was designed to not just to replicate that used by Benson and Perrett (1991) but sought to measure best-likeness judgements with a greater deal of sensitivity. The methodology and choice of stimuli used here differed in several ways. First, all the face stimuli used in the experiments which follow were photographed and rendered in full-colour rather than grey-scale. Second, external features and background were removed as discussed in the previous experiments. Third, an interactive technique was used such that rather than selecting a best-likeness from a sequence of static images displayed simultaneously, subjects were allowed to manipulate a face in real-time so as to alter the degree of anticaricature or caricature. Fourth, a further modification was made in Experiment 6 so as to investigate best-likeness purely in the colour domain.

The images in the experiments were shown to subjects in 24-bit colour in a large graphics window on a well-illuminated computer screen The stimuli in the Benson and Perrett (1991) study were presented as 5" x 7" black and white photographs with background information kept fully intact. Other studies (Benson and Perrett, 1994; Rhodes *et al*, 1987; Rhodes *et al*, 1996) have differed further in that line drawn stimuli have been presented. The use of high-quality colour photorealistic stimuli in the current experiment may serve to limit the amount of caricaturing desired to produce a best likeness as much more information pertinent to recognition of the face, texture and colour is present (as discussed in *Chapter 3)*. Stimuli in the experiments described in this chapter were cropped digitally in order to remove background information. As discussed in *Chapter 1*, internal face information is thought to be more than sufficient for recognition of famous faces (Ellis *et al*, 1979; Young *et al*, 1985). Further, caricaturing of images is likely to invoke distortions with respect to the background which may provide a cue to the veracity of the images.

In the Benson and Perrett (1991) study subjects were shown faces in sets of four for any given faces. For a particular face they were shown images either with the series 32% anticaricature, 16% anticaricature, veridical and 16% caricature or with the series 16% anticaricature, veridical, 16% caricature and 32% caricature. All four stimuli were displayed to the subjects simultaneously on strips and the participant was asked to judge which he/she felt to be the best likeness and to rate each of the four images for quality of representation for the individual it depicted. Rhodes *et al* (1996) presented line-drawn stimuli as a13-image continuum from a 60% anticaricature to a 60% caricature and individuals were asked to select a best-likeness from this set. In the present study subjects were given the chance to manipulate the level of caricature for the faces in real time. In this manner they were able to judge how close a likeness the image was to their perception of the individual at any point within the caricature range of images provided. The subject was not compelled to choose a particular image from a series of four or 13 discrete images displayed together. They were granted the possibility to opt for a best likeness at some point between two otherwise discrete, static images. It is likely that given free reign to choose one's ideal precise representation of a person's face from a full range of colour photographic images the selected best-likeness would be quite different to that found in the Benson and Perrett (1991, 1994) or Rhodes *et al* (1996) studies. From the distribution of choices for best-likeness found in these prior studies one could speculate, however, that the mean choice for best-likeness would be a mild caricature.

In order to perform the interactive task described using photorealistic images an interactive technique has been developed by which an individual, using a mouse, can change the appearance of a colour photorealistic face from an anticaricature, in colour or shape, to the associated caricature smoothly and in real time. This allows viewers to choose the exact point at which they feel the image presented to be the best likeness of an individual.

In Experiment 6, the perceived best likeness when faces are caricatured in shape was investigated. In Experiment 7 the interactive technique was extended to identify choice for best likeness for the manipulation of intensity, hue and saturation information by caricaturing in RGB. The validity of the latter as an aid to recognition has already been found in caricature studies (Experiments 3 and 4, *Chapter 4*).

General Experimental Methods

Apparatus. A Silicon Graphics R4400 Indigo high impact computer with 4MB TRAM texture memory was used for stimulus creation and presentation in 24-bit and real-time with monitor set at medium contrast (factory setting); an Epson GT-6500 scanner was connected to an Escom 486 DX2/66 IBM-compatible personal computer.

Stimuli. The facial areas of photographs of 60 male actors were scanned at a resolution of 517 pixels in width by 720 pixels in height at 150 dpi in 24-bit colour. Individual stimuli used in the experiments consisted of a subset of these images. Each of the individuals was clean-shaven and wore no spectacles. The features on the faces were defined using 179 feature points placed manually in standardised positions. A composite was created for the set of actors (*Plate 4.4*). This composite was formed by using the position data from all the faces in the appropriate set to compute an average face shape. Average feature colour information was then calculated by digitally averaging corresponding pixels of individual faces from each face in the set after they were warped into the group average shape. These methods are described in greater detail elsewhere (Rowland and Perrett, 1995). In each image, the internal features from the face were cropped from the hair, ears and background so that distortion of these latter items would not affect judgements to best-likeness. The internal features were then placed on a black background (see *Plate 2.2).*

Experiment 6

Introduction. Experiment 6 sought to investigate the perceived best-likeness of individuals when caricatured in shape.

Methods

Subjects. 30 undergraduates at the University of St Andrews: 15 females and 15 males.

Stimuli. From the pool of cropped images of 60 male actors, six were selected which had been rated as highly familiar in an earlier familiarity questionnaire. Each of these veridical images was normalised for inter-ocular position and caricatured in shape against the actor prototype. A shape caricature is defined as an exaggeration of the differences between an individual face shape and the average face shape. Two such caricatures were made for each individual. For each, one caricature was produced by exaggerating the difference in x, y position of the corresponding points between the individual and the prototype by 25%. The colour information of the veridical was warped so that the image is stretched to accommodate these changes in position. A second caricature was produced the same way, but with an exaggeration of 50% of the difference between the individual and the prototype applied. Two anticaricatures were produced similarly by reducing the differences for each corresponding feature point between the individual and the prototype such that the greater the degree of anticaricature the more average the shape of the individual. All the resulting images retained all the colour information pertinent to the given individual and as such the colour information was not altered across the images. The image progression is illustrated in *Plate 6.1*.

A morph was created utilising the Silicon Graphics hardware. This used the resultant five images (two caricatures, two anticaricatures, and the veridical) to enable movement between a 50% anticaricature and a 50% caricature that could be made in real-time. In order to do this the original images were first stretched to be 512x512 pixels in size. Feature point data were also stretched by the same factor to maintain registration. The Delaunay refinement algorithm (Ruppert, 1995) created a triangular tessellation from the feature point data. This tessellation structure projects the texture into the desired shape. To do this, the original (unstretched) version of the feature point data was used to map the texture map triangles on screen in the correct dimensions. Texture, in this sense, can be defined as the colour values given to represent the various pixels across an image and through mapping with the associated feature points can be linked to the shape of the face (Rowland, 1997). The amount of texture which is present in the displayed image is controlled using an alpha component. This specifies the transparency, how the values of two overlaid pixels are combined, between two images in the continuum. Thus if the images are manipulated in real time to be halfway between a 25% and 50% caricature, the shape will be represented by an average of the feature point data for 25% and 50% caricature and the textures of both the 25% and 50% caricatures will be projected onto this shape with equal intensity through the alpha channel. All five images were used in order to ensure smooth continuity across the sequence using four continuous, but discrete, morphs.

Design. There was one independent variable in the experiment: target face. The dependent variable was the precise level of caricature (across an infinite continuum between a 50% anti-caricature and a 50% caricature) perceived by an individual subject to be the best-likeness of the individual which the target face represented.

Procedure

Subjects were given a list of the names of the six target faces to be presented in the task and asked to rate familiarity with celebrities' faces. They were asked to rate each of the six celebrities on a scale of 1 to 5 to indicate how familiar they believed themselves to be with a given individual's face, where 1 indicated completely unfamiliar and 5 indicated a high degree of familiarity.

Trials were given with each individual celebrity presented in a pseudo-random order. Each of the six individuals was displayed in 24-bit colour in a square window with a size of 1000x1000 pixels. Subjects were told that each individual would be from the rating list given and subjects were asked to enquire as to the names of any individuals whom they were unable to identify. Each subject was asked to hold the left mouse button while moving the mouse from left to right within the window. This latter movement resulted in the increasing or decreasing of the level of caricature in real-time. In 50% of the trials leftward movement of the mouse resulted in decreasing the level of caricature, while in the remaining trials leftward motion of the mouse resulted in an increase in caricature level. In all cases the full range of caricature was from -50% to +50%. Subjects were asked to press the space-bar when they felt the image they had selected through movement of the mouse represented the best-likeness within the range available for the individual concerned.

After the task was complete, each of the veridical images was presented in turn, in a form which was not cropped, and subjects were asked to rate each on a scale of 1 to 5 (where 1 indicated highly unrepresentative and 5 indicated highly representative) as to how close to their personal representation of the relevant celebrity they believed the images to be. The images shown at this stage were not cropped as absence of background information is likely to affect subjects ratings due to lack of contextual information. Displaying images in cropped format here would have skewed ratings toward the lower end of the spectrum as any face with background and external features such as ears removed is likely to be described as somewhat unrepresentative of the individual.

Results and analysis

Medians and modes were found for the post-task questionnaire to indicate how good the likenesses were seen to be by the subject population used and how familiar the subjects believed themselves to be with each individual represented. All the veridical images were determined as good likenesses in this manner with medians and modes of 3 and greater on the 5 point scale of typicality.

The percentage of caricature selected for each set of images representing an individual celebrity was calculated in a by-subject analysis.

The best-likeness was found to be anticaricature (mean -8.2%) such that the subjects were making the shapes of the faces slightly more average when choosing the image they felt best represented the target. The mean caricature level was significantly lower than the level of 0%, or veridical, which would be predicted by the null hypothesis (*Figure 6.7* $t=-3.39$, $p<0.005$, $df=29$).

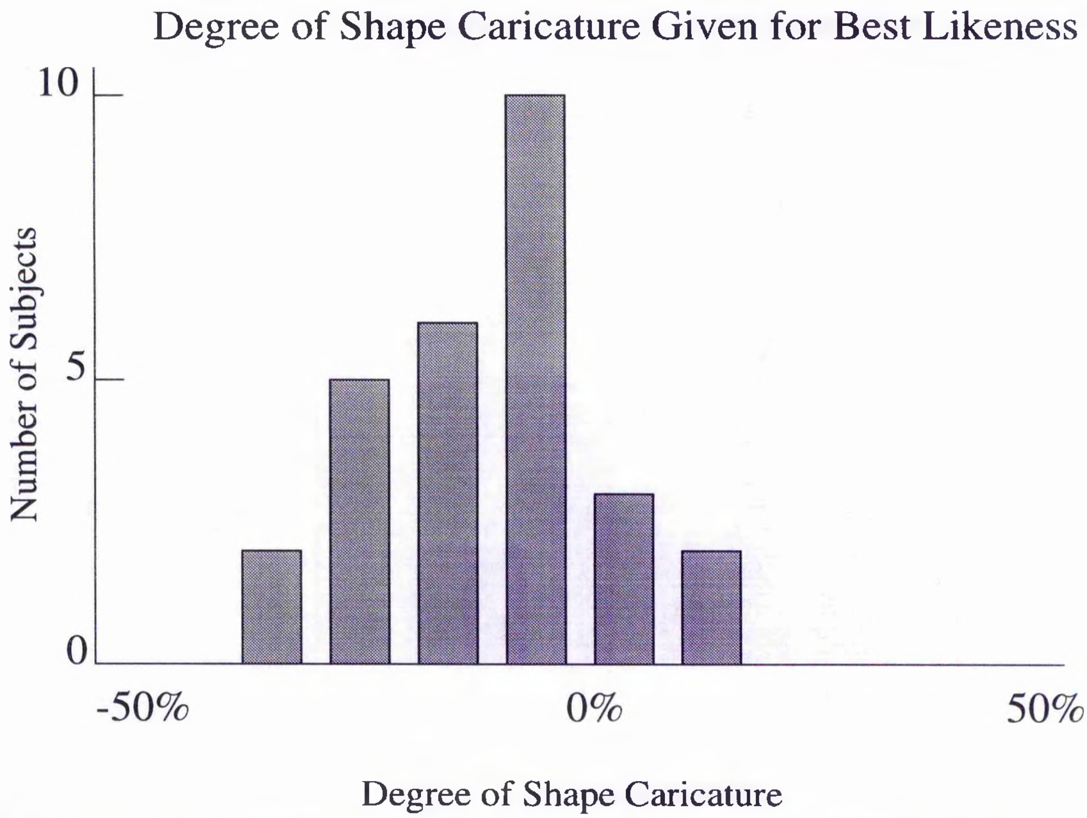


Figure 6.7. Degree of shape caricature given for best likeness. An anticaricature was found for best likeness when interactively manipulating the shapes of faces. This histogram describes the range of best-likeness selections in bins of 10%.

Using the set of faces a random factor, the mean caricature level was not found to be different from 0% ($t=-1.999$, $p=0.102$, $df=5$, 2-tailed). The failure of this analysis to reach statistical significance could be attributed to the limited number of target faces used. In fact, the faces were subjected to a one-way ANOVA with the target face as the only factor (six levels). A significant main effect was found for this factor ($F_{1,29}=11.52$, $p<0.005$). To investigate further each of these target faces was further tested against the veridical value of 0. As such the best-likeness chosen for three of the faces did not differ significantly from veridical (Kevin Costner, mean=0.004, $t=0.86$, $p=0.40$, $df=29$; Bruce Willis, mean=0.0009, $t=0.19$, $p=0.85$, $df=29$; Al Pacino, mean=0.008. $t=-1.83$, $p=0.077$, $df=29$), while for three others the best-likeness chosen was a significant anticaricature (Christian Slater, mean=-0.10, $t=-2.11$, $p<0.05$, $df=29$; Robert Deniro, mean=-0.12, $t=-2.61$, $p<0.05$, $df=29$; Harrison Ford, mean=-0.24, $t=-5.44$, $p<0.001$, $df=29$).

Discussion

The best-likeness was found to be an anticaricature (mean 8.2%) with best-likeness judgements as a rule skewed to anticaricatures being chosen as most representative. This finding contrasts with that found for the presentation time experiment described in *Chapter* 2 and findings found for reaction time studies of recognition for caricatured line-drawings and photographic images in previous studies. In these studies, caricatures were found to improve recognition accuracy or to reduce reaction time. Under these paradigms stimuli were only displayed for a brief period of time. Exposure of the images to the subject for an extended period of time during an interactive caricaturing procedure allows the subject greater time to analyse the image. In a reaction time or presentation time paradigm subjects are only shown the faces for a very brief period of time and so it is likely they do not get the opportunity to ascertain whether the image they have been presented with has been distorted in any way. Whereas in the presentation time and reaction time paradigms involve only brief exposure to the stimulus in which any exaggeration of relevant information from a norm is a benefit to recognition in the best likeness paradigm the subject is free to survey the image and move the image through the boundary from 50% anticaricature to a 50% caricature.

The results here differ from those found by Benson and Perrett (1991) and Rhodes *et al* (1996). The results from their respective studies indicated that positive caricatures were seen to be better likenesses than veridicals or anticaricatures with an interpolated mean best-likeness at a caricature level of 4.4% and 11%. However, these findings were not pervasive. Of the seven faces used in the Benson and Perrett (1991) experiment, caricatures were selected as best-likenesses for only three of them while best likeness judgements for the other four were not seen to differ significantly from veridical. Further, of the three faces that were caricatured that which was caricatured most was Margaret Thatcher. It is likely that viewers had seen the then Prime Minister's face caricatured many times previously. It should be noted that although the distribution of face image chosen as best-likeness was skewed towards caricature the most popular choice for best likeness was the veridical. Similarly, in Rhodes *et al*'s (1987) study using line-drawn stimuli veridical drawings were selected as best likenesses with about the same frequency as 25% caricatures. Again, the best likeness was interpolated to be a caricature with a mean level of 16%. There are several other possibilities as to why the findings here differ from those found by Benson and Perrett using photorealistic stimuli. These include differences in quality of the stimuli and in the paradigms used.

Quality of stimuli

In the experiment described here the stimuli which were presented had been cropped digitally in order to remove background information. These images were also displayed in 24-bit colour on a well-illuminated computer screen with background to the images replaced with a uniform black background. The stimuli in the Benson and Perrett (1991) study were presented as 5" x 7" black and white photographs and included the background. Other studies (Benson and Perrett, 1994; Rhodes *et al*, 1987; Rhodes *et al*, 1996) have differed further in that line drawn stimuli have been presented. The use of high-quality photorealistic stimuli in the current experiment may have reduced the amount of caricaturing introduced to produce a best likeness as much more information pertinent to recognition of the face, texture and colour, is present as discussed in *Chapter 3*. Certainly as the amount of information available in the stimuli has increased the level of caricature for best-likeness has decreased in this and previous studies (line-drawn stimuli: 42%, Benson and Perrett, 1994; 16%, Rhodes *et al*, 1987; enhanced line-drawings: 11%, Rhodes *et al*, 1996; grey-scale photorealistic images: 4%, Benson and Perrett, 1991; full-colour photorealistic images:-8%, this study).

Paradigm differences

In the Benson and Perrett (1991) study subjects were shown faces in sets of four for any given faces such that for a particular face they were shown images either with the series 32% anticaricature, 16% anticaricature, veridical and 16% caricature or with the series 16% anticaricature, veridical, 16% caricature and

32% caricature. All four stimuli were displayed to the subjects simultaneously on strips and the participant was asked to judge which he/she felt to be the best likeness and to rate each of the four images for quality of representation for the individual it depicted. In the present study subjects were able to manipulate the level of caricature for the faces in real time. In this manner they were able to judge how close a likeness the image was to their perception of the individual at any point within the -50% to +50% caricature range. The subject was not compelled to choose between one of four static images being shown. It is likely that it was this free range of choice that resulted in subjects choosing an image that was on average a slight anticaricature. It can also be noted that in the Benson and Perrett (1991) study the levels of image caricature appeared at 16% intervals and thus images near the 8.2% mean level of anticaricature selected here would not have been an option. It is likely that the paradigm used by Benson and Perrett (1991) was not sensitive enough to produce this finding.

The findings here are not based on an interpolated best-likeness but show the level of caricature which was actually selected by participants in the experiment. In a prior interactive caricaturing study line drawings have been used rather than photorealistic stimuli (Benson and Perrett, 1994). In that study, a 45% caricature was typically chosen as the best-likeness. This study finds a slight anticaricature as best-likeness, but while the use of line drawings again limits the information available to the viewer, photorealistic colour images contain a great deal of the information attached to a real face. Thus, in a situation were an individual is free to choose the best-likeness of individuals from within a restricted contiguous range using images of the quality of colour photographs, this likeness is found to be an anticaricature but this was not pervasive across the faces used and a by face analysis did not find a mean significantly different from 0% (veridical). It must be noted here, as discussed in *Chapter 2*, that there were only six faces shown in this experiment and as such by face analysis lacks the sensitivity of the by subject analysis provided. Inclusion of more faces in the experimental procedure would circumvent this problem and allow one to determine if there is a relationship between aspects of the face (expression, typicality/distinctiveness, recognizability, familiarity, representativeness of the image) and degree of caricature and anticaricature selected. It may be that the more typical the face the more it is likely to be caricatured, or less-likely to be anticaricatured, to produce a best-likeness as a greater degree of caricature is possible before the face will leave the bounds which constrain feature variability of the human face[3] .

Contrary to the presentation time paradigm subjects are exposed to the faces for prolonged periods of time and facial distortions are more easily identified. It is possible that anticaricatures are being chosen in order to compensate for the increased levels of distortion found on the caricatured side of the transformation. Subjects may choose a level of caricature closer to the mean shape in order to preserve a more human looking face. One can speculate here as to why some faces were anticaricatured to produce a best-likeness while in other cases the veridical was chosen as a best-likeness. It may be that those faces which are given the greatest degree of anticaricature are those which are originally more distant from the norm and are therefore reduced more to make them closer to the norm, in order to match a configuration which appears more face-like to the viewer. As texture is constant for these faces, much of the information required to identify the face is present even when the shape is that of an average face. The presence of colour, incorporating textural information, has already been shown to be an important aid to face recognition in *Chapters 3 and 4*. It is likely to be an important factor for facial recognition, perhaps more so than shape and in addition to benefiting recognition in a presentation time paradigm it may be found to be more pertinent when subjects are required to identify a best likeness. This was investigated in Experiment 7.

---

[3]**Footnote.** It is worth noting here that when subjects were shown line-drawings of the five faces used in the analysis for Experiment 6, Harrison Ford's face was selected as the most distinctive. His was also the only face that subjects were able to identify purely from line-drawn information (66% of subjects responded accurately with 0% responding accurately for the other faces) further attesting to the lack of information available in line-drawings of faces. The faces of Harrison Ford and Robert Deniro were also the only two faces given a mode of 5 for representativeness of the veridical image in the post-task questionnaire indicating that highly distinct and/or representative faces may be anti-caricatured in best-likeness tasks.

Experiment 7

Introduction. Experiment 7 sought to investigate the perceived best-likeness of individuals' faces when caricatured in colour or contrast was enhanced. This experiment used the interactive technique employed for Experiment 6 to investigate the best-likeness chosen for colour-caricatured and contrast-enhanced images introduced in Experiments 3 and 4.

Methods

Subjects. These were 40 undergraduates at the University of St Andrews: 15 females and 15 males. 30 of the subjects were the same as those who participated in Experiment 6.

Stimuli. Six veridical images, none of which had been included in Experiment 6, were caricatured in the colour domain. Each of these images was of a famous male actor and each was caricatured against a composite of 60 actors. The colour caricaturing process was carried out by warping the shape of the relevant prototype face image to that of the target face image. In a manner similar to Experiment 6, four different levels of colour caricatures were produced - a 40% and 80% caricature and a 40% and 80% anticaricature. For each corresponding pixel, the difference in each component (r,g,b) of the RGB values between the target image and the prototype was multiplied by the appropriate percentage and then added to the target image. The result was a colour-caricatured image with the colour and intensity difference between the target image and the prototype exaggerated or produced by that percentage factor (see Burt and Perrett, 1995; Rowland and Perrett, 1995 and *Chapter 4* for a more detailed description). Contrast-enhanced and contrast-reduced images of the same faces were also created by enhancing the differences in RGB values between the individual face image and a uniform 'grey box' - a mid-grey rectangle with RGB values (127,127,127). Here, the co-ordinates (R,G,B) represent the intensities of the red, green and blue intensity values where 0 indicates zero intensity and 255 indicates maximum intensity. In this manner contrast enhancement also alters colour saturation. The amount of contrast enhancement here was such that the mean difference in pixel intensity and thereby colour saturation between the veridical and the contrast enhanced image was the same as the mean difference between the veridical image and the associated colour caricature, e.g. the mean difference in the RGB triple of a 40% colour caricature of an individual and the veridical was the same as that between the '40%' contrast enhanced image and the veridical.

To explain further each pixel position, k, in the 517 x 720 pixel face array can be described as a triple by representing the RGB colour values for the particular pixel as F(s, k) for the source face, for instance. If the corresponding pixel for a prototype face is F(p, k) then the equivalent pixel in a colour caricature against the prototype face, $F(c_p, k)$ can be described as follows:

$$F(c_p, k) = F(s, k) + l[F(s, k) - F(p, k)]$$

where l is the caricature factor, say 0.4 for a 40% caricature (Rowland and Perrett, 1995).

The pixel values for a contrast-enhanced representation of the same face, $F(c_g, k)$ produced by caricaturing against a grey-box, F(g, k), can be calculated with a corresponding formula:

$$F(c_g, k) = F(s, k) + m[F(s, k) - F(g, k)]$$

here the value of each component of each pixel, k, for F(g, k) is the value 127.

The value of m for the contrast controls generated was calculated such that the mean pixel intensity

$$\Sigma(F(c_p, k) - F(s, k))/517\text{x}720$$

calculated for all pixel values, k, across the 517 x 720 array is equivalent to the mean pixel intensity

$$\Sigma(F(c_g, k) - F(s, k))/517x720$$

calculated for all k across the array.

This procedure prevents excessive contrast in the control with values numerically greater than 255 or less than 0. This contrast-enhanced image is digitally equivalent to turning up the colour saturation and contrast on a television set (for further explanation see *Chapter 4*). The shape of each face image was left unaltered.

A separate morph for each of the colour caricature and contrast enhanced conditions was produced purely in the colour domain. This used the resultant five images for each condition to enable movement between an 80% anticaricature or contrast-reduced image and an 80% caricature or contrast-enhanced image in real-time. The images used for the progression are illustrated (colour caricature, *Plate 6.2*; contrast control, *Plate 6.3*).

The amount of texture which is present in the displayed image is controlled using the alpha component of RGB colour space which specifies the transparency between two images in the continuum. Thus if the images are manipulated in real time to be halfway between a 40% and 80% caricature the textures of both the 40% and 80% caricatures will be projected onto the veridical shape (as this remains constant) with equal intensity through the alpha channel. All five images were used in order to ensure smooth continuity across the sequence using four continuous, but discrete, morphs (where morph here consists of this texture component alone).

Design. Thirty subjects participated in the main body of the experiment. There were two independent variables in the experiment: target face and type of image manipulation (colour caricature or contrast-enhancement). The dependent variable was the precise level of caricature or contrast enhancement perceived by an individual subject to be the best-likeness of the individual which it represented (in both instances this varied across a continuum and the levels were infinite between a 80% anti-caricature/reduced contrast image and an 80% caricature/enhanced contrast image). In a further between-subject paradigm involving 10 subjects, each group of five subjects was tested in conditions whereby the monitor was set at either a high contrast or low contrast level. Here, the target face and image types being manipulated remained as within-subject variables.

Procedure

Subjects were given a list of the names of the six target faces to be presented in the task and asked to rate familiarity with the celebrity faces. Instructions were the same as for Experiment 6.

Trials were given with the series of images - colour caricatured images and contrast control images presented in pseudo-random order. Each of the six individuals was displayed in 24-bit colour in a square window with a size of 1000x1000 pixels. Subjects were told that each individual would be from the rating list given and subjects were asked to enquire as to the names of any individuals which they were unable to identify. Each subject was asked to hold the left mouse button while moving the mouse from left to right within that window. This latter movement resulted in the increasing or decreasing of the level of caricature in real-time. In 50% of the trials leftward movement of the mouse resulted in decreasing the level of caricature, while in the remaining trials leftward motion of the mouse resulted in an increase in caricature level. In all cases the full range of caricature was from -80% to +80%. Subjects were asked to press the space-bar when they felt the image they had selected through movement of the mouse represented the best-likeness within the range available for the individual concerned. The contrast-enhanced series and colour-caricatured series comprised the same target individuals but were presented to subjects separately. The series were presented so that contrast-enhanced series and colour-caricatured series for different target faces were presented in pseudo-random order and the position of target faces in the presentation sequence was counterbalanced across subjects.

A post-task questionnaire was given as in Experiment 6.

Results and analysis

Medians and modes were found for the post-task questionnaire to indicate both how representative the images were seen to be by the subject population and how familiar these subjects were with the faces of the individuals represented. One face was removed as the subjects used reported a low familiarity with it.

The percentage of caricature selected by each subject for each set of images representing an individual celebrity was calculated.

In the colour-caricaturing instance, the best-likeness was found to be a caricature (mean 14.5%) such that the subjects were selecting faces through which relevant colour had been exaggerated. Subjects were also found to select contrast-enhanced images as ideal likenesses (mean 28.7%) such that subjects expressed that faces with higher-contrast to be much better likenesses than veridical. Both effects were found to be statistically significant across both the subjects ($t=4.56$, $p<0.001$, $df=29$; $t=9.68$, $p<0.001$, $df=29$, *Figure 6.8 and 6.9*; respectively) and the set of faces used ($t=5.05$, $p<0.01$, $df=4$ and $t=8.69$, $p=0.001$, $df=4$; respectively) such that degree of caricature was significantly greater than that suggested by the null hypothesis (0%). The best-likeness chosen with enhanced contrast was found to be significantly greater than the best-likeness chosen for colour caricature in both within-subject ($t=6.67$, $p<0.001$, $df=29$) and within-item ($t=7.26$, $p<0.005$, $df=4$) analyses. In a separate analysis, a 2-way between-subject ANOVA with type of manipulation and target face as within-subject factors and monitor contrast setting as a between-subject factor was also employed. Subjects tested in conditions whereby the monitor controls were used to set the screen at either high contrast or low contrast levels did not produce a significant effect of degree of manipulation for level of colour caricature or contrast enhancement selected as best-likeness ($F_{1,8}=1.86$, $p=0.21$).

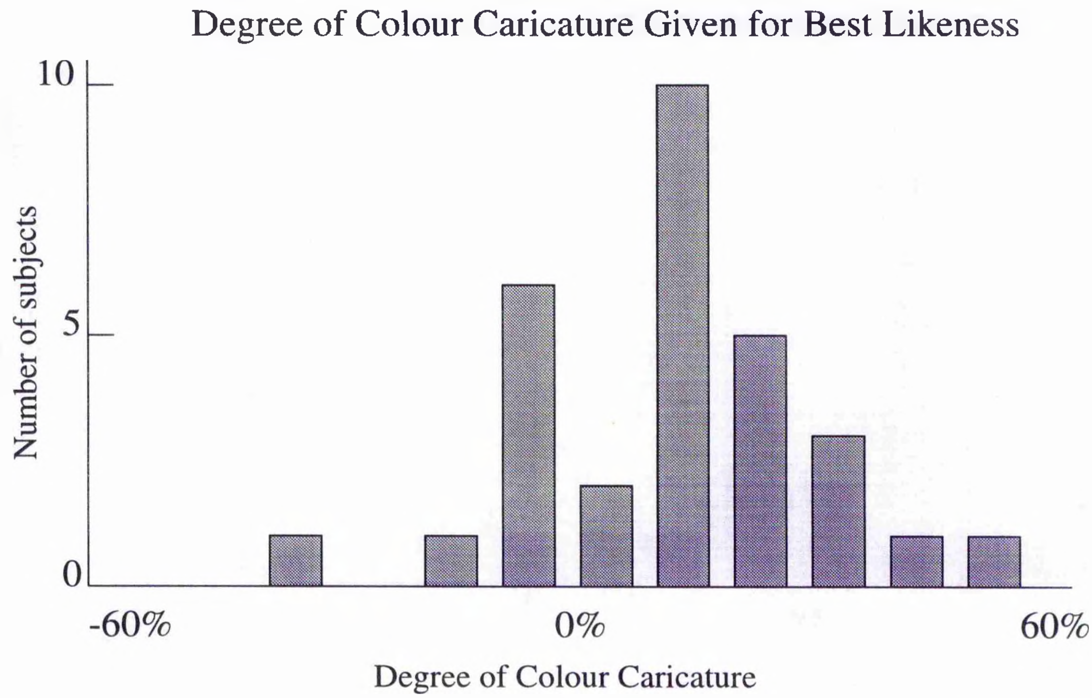


Figure 6.8. Degree of colour caricature given for best likeness. A caricature was found for best likeness when interactively manipulating the colour of faces. This histogram describes the range of best-likeness selections in bins of 10%.

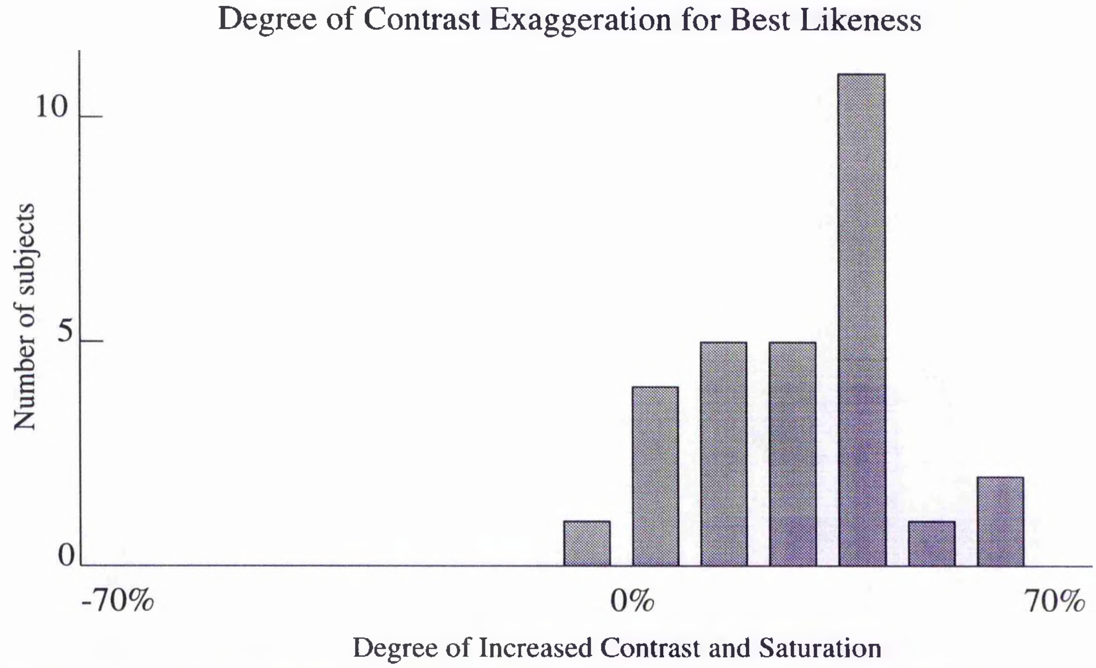


Figure 6.9. Degree of contrast exaggeration for best likeness. Contrast-enhanced images were found to be the best likenesses when interactively manipulating the colour found in images of celebrities. This histogram describes the range of best-likeness selections in bins of 10%.

Discussion

The best-likeness in the first section of the experiment, involving the manipulation of colour relative to an appropriate facial norm, gave rise to a best-likeness with a caricature (mean 14.5%). The best-likeness with contrast and saturation enhancement was found to be at double the level of exaggeration (28.7%). This contrasts with the situation whereby images are presented for brief periods of time using the presentation time paradigm of *Chapter 4*; there colour information specifically relevant to that face and available in a 40% colour caricature was useful for recognition compared to a veridical image and contrast enhancement itself was not beneficial for recognition. Thus the optimal stimulus for recognition may not be seen as the best likeness.

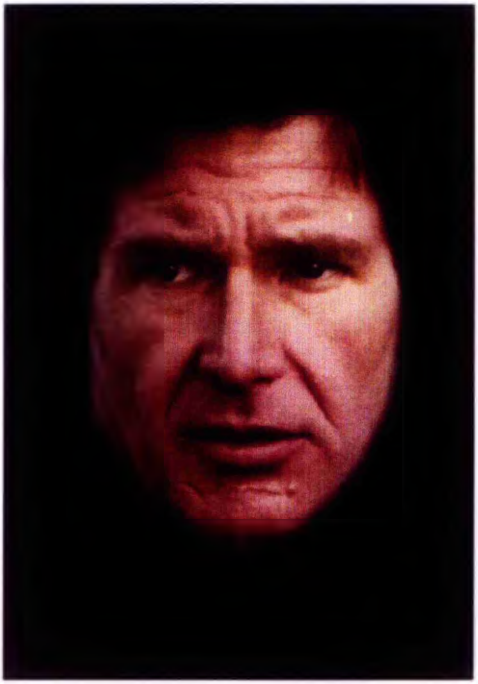
The findings here also suggest that when subjects are looking at images for an extended period of time they have a tendency to select images in which colour information relevant to the face has been exaggerated or the contrast has been enhanced to at least some degree. Enhancing the colour information exaggerates individual texture and contour information giving cues to the 3D structure of the face which may be important to selecting a facial images as having the best-likeness. This also indicates the importance of information regarding complexion in a face. A bland average-coloured face where textures have been blurred together is also unlikely to be selected as the best-likeness for an individual person with a known distinctive complexion and hence best-likenesses are chosen with some degree of caricature. It appears that subjects are also inclined to select for best-likeness images with higher contrast. Manipulating the contrast level of the monitor screen such that contrast was either very high or very low did not affect the viewers decision to select high-contrast images as best-likenesses. Thus it is unlikely that the images displayed were low contrast and subjects were merely selecting normal levels of contrast for the face by this manipulation. In these pictures the images appear to stand out more from the background and seem more vivid to the viewer. The colour caricature may be exaggerated to a lesser degree than the contrast control in order to keep the resultant face as face-like as possible. The level of exaggeration required to bring out facial structure and wrinkles is lower when caricaturing against a face norm than a neutral-grey image. The information found in a high colour caricature is likely to exceed that found within a normal range of faces as it represents information which is exaggerated from a prototypical face. However, enhancing the contrast of a face does not affect the face in this way so this may lead to a greater degree of contrast enhancement selected for the best-likeness.

General Discussion

Rhodes *et al* (1987) proposed two models which may be responsible for the caricature advantages described here. The first maintained that caricatured representations rather than veridical representations are stored in memory. In this case caricatures would be more efficiently recognised because they are closer to the stored representations. Alternatively, the second model proposed that representations stored in memory are veridical but caricaturing aids the process of matching the input image to the veridical representation. Here, the caricaturing process would constrain the search, because exaggerating features would make it easier to realise qualitatively what kinds of features the target faces possesses.

This second model appears to be the more likely given the findings described in this and the preceding chapters. Caricaturing either shape information or texture information for a face against a relevant norm increases recognition accuracy using a presentation time paradigm. This has been shown with images caricatured up to a level of 50% and presented for a brief time. So it would seem that exaggerating the features against a relevant norm does allow viewers to identify a target face from a set of possibilities. Caricaturing in shape has also been shown to reduce the time taken to decide if the face represents a given individual in a face-name matching task (Benson and Perrett, 1991). However, individuals do not find it necessary to caricature the shape of an individual's face to identify a best likeness indicating that a face is not somehow stored in caricatured form in the memory. Increasing contrast for a face provides a better likeness than exaggerating the colour information against a relevant norm. In this way, enhancing contrast produces a better likeness but does not aid recognition (see *Chapter 4*). This indicates that individuals may have a general preference for enhanced contrast when viewing faces but such preferences do not explain a caricature advantage. It is also possible that, as discussed in *Chapter 3*, colour information is much more important in face recognition due to the homogeneity of faces as a group and as such shape plays a secondary role in recognition and judgement of identity. The finding that faces are colour-caricatured to some extent in an interactive best-likeness task while veridical or anticaricatured images are chosen when manipulating shape may provide support for this. To help pin-point this relationship, it would be interesting to produce stimuli which consisted of prototypical texture or vice-versa but were caricatured in shape to be studied using both presentation time, or reaction time, and best-likeness paradigms. Through this technique it may be shown that face images in which individual colour but not shape information is available may be recognisable as the individuals represented whereas images of faces where only the shape information is available would not be recognised. It is also possible that under such circumstances a caricatured face shape with neutral texture may be chosen as best likeness in an interactive best-likeness paradigm as the images would already be distorted in the sense that they would contain texture information not associated with the individual presented; recognisability and likeness for the individual would already be impaired and no manipulation could produce a truly ideal depiction.

There is much scope for study of the factors prompting facial recognition and/or judgement of facial likeness. The information relevant for recognition could be studied directly as described above to see if texture information alone or shape information alone lead to better recognition of a face when the rest of the information provided was adopted from an average face. Under these conditions, the effect of caricaturing in colour and/or shape could also be explored both in identifying best likeness and in recognition under inhibited viewing conditions using a reaction time or presentation time paradigm. In this manner the strength of utility of various types of information could be explored when facial processing has been handicapped in some way.



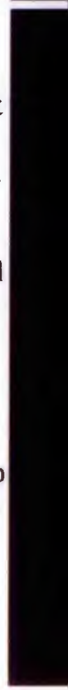
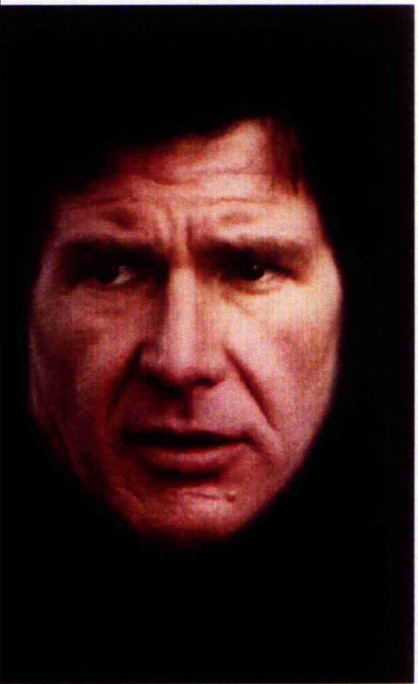
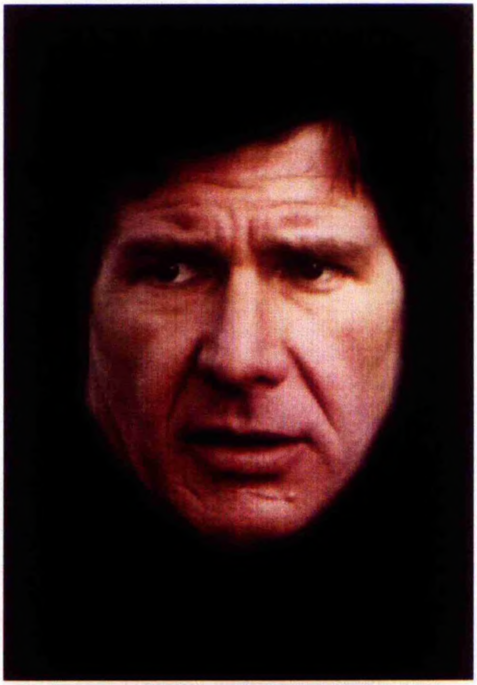
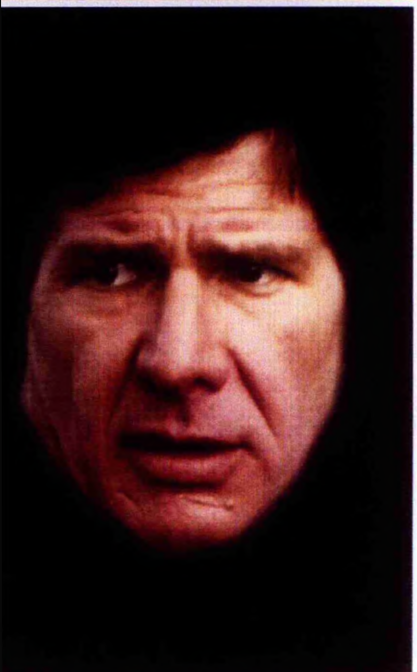
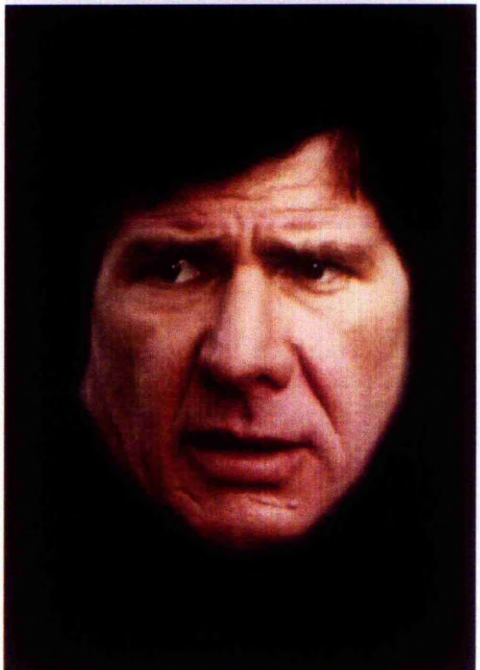


Plate 6.1. These images illustrate the range of choice the subject is given to select which he/she believes to be the best likeness for a given individual. The top left image is warped to be halfway between the shape of the average actor prototype and the individual face, i.e. a 50% anticaricature. The top middle image is a 25% anticaricature and the top right image is a veridical or accurate reproduction of a photograph of the individual actor, Harrison Ford. The bottom left and bottom right images have been caricatured in shape by 25% and 50% respectively. All images retain the texture information for the individual. When individuals manipulate the image on-screen it moves fluidly between these five images from a 50% anticaricature to a 50% caricature and all the possible configurations in-between.

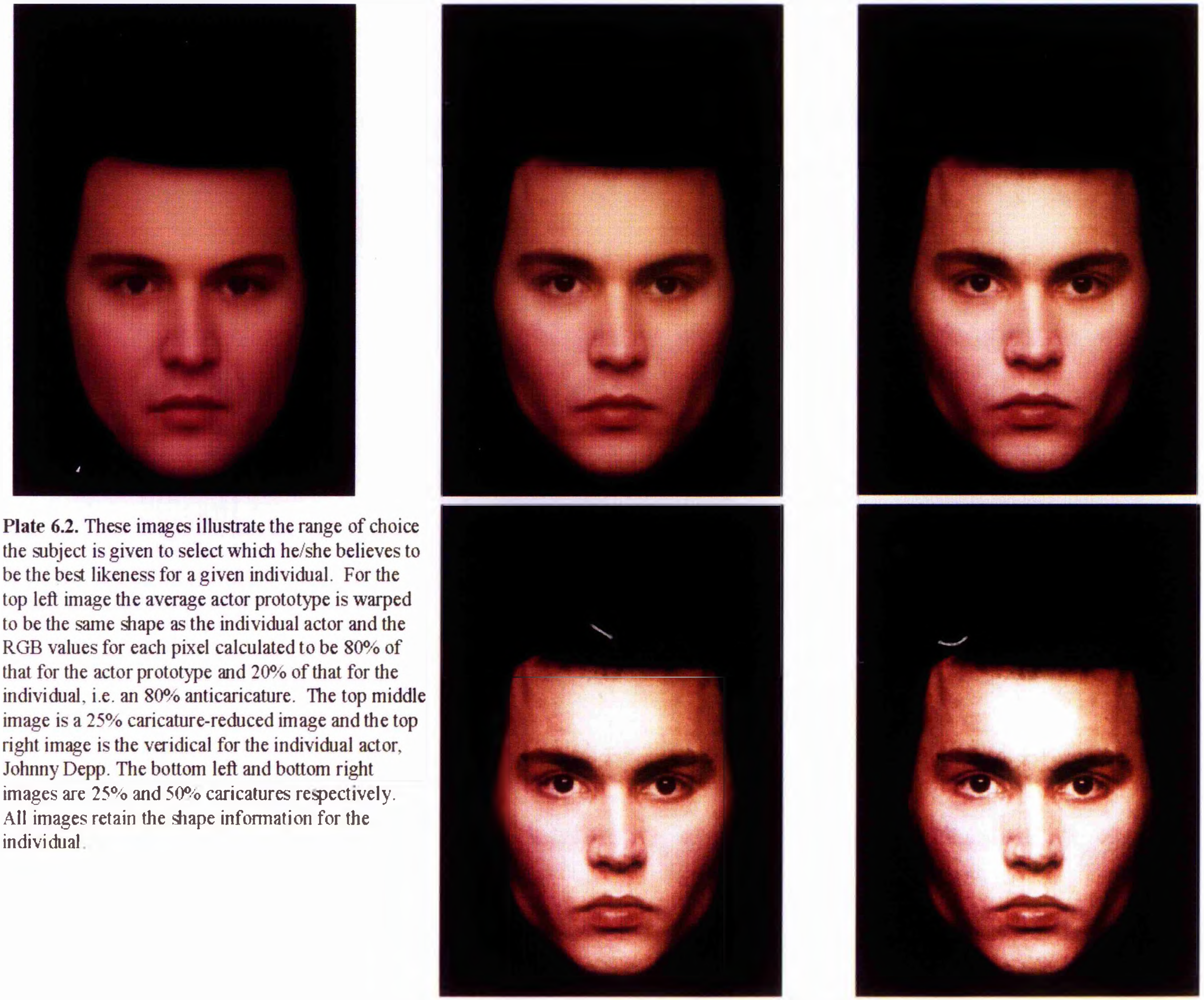


Plate 6.2. These images illustrate the range of choice the subject is given to select which he/she believes to be the best likeness for a given individual. For the top left image the average actor prototype is warped to be the same shape as the individual actor and the RGB values for each pixel calculated to be 80% of that for the actor prototype and 20% of that for the individual, i.e. an 80% anticaricature. The top middle image is a 25% caricature-reduced image and the top right image is the veridical for the individual actor, Johnny Depp. The bottom left and bottom right images are 25% and 50% caricatures respectively. All images retain the shape information for the individual.

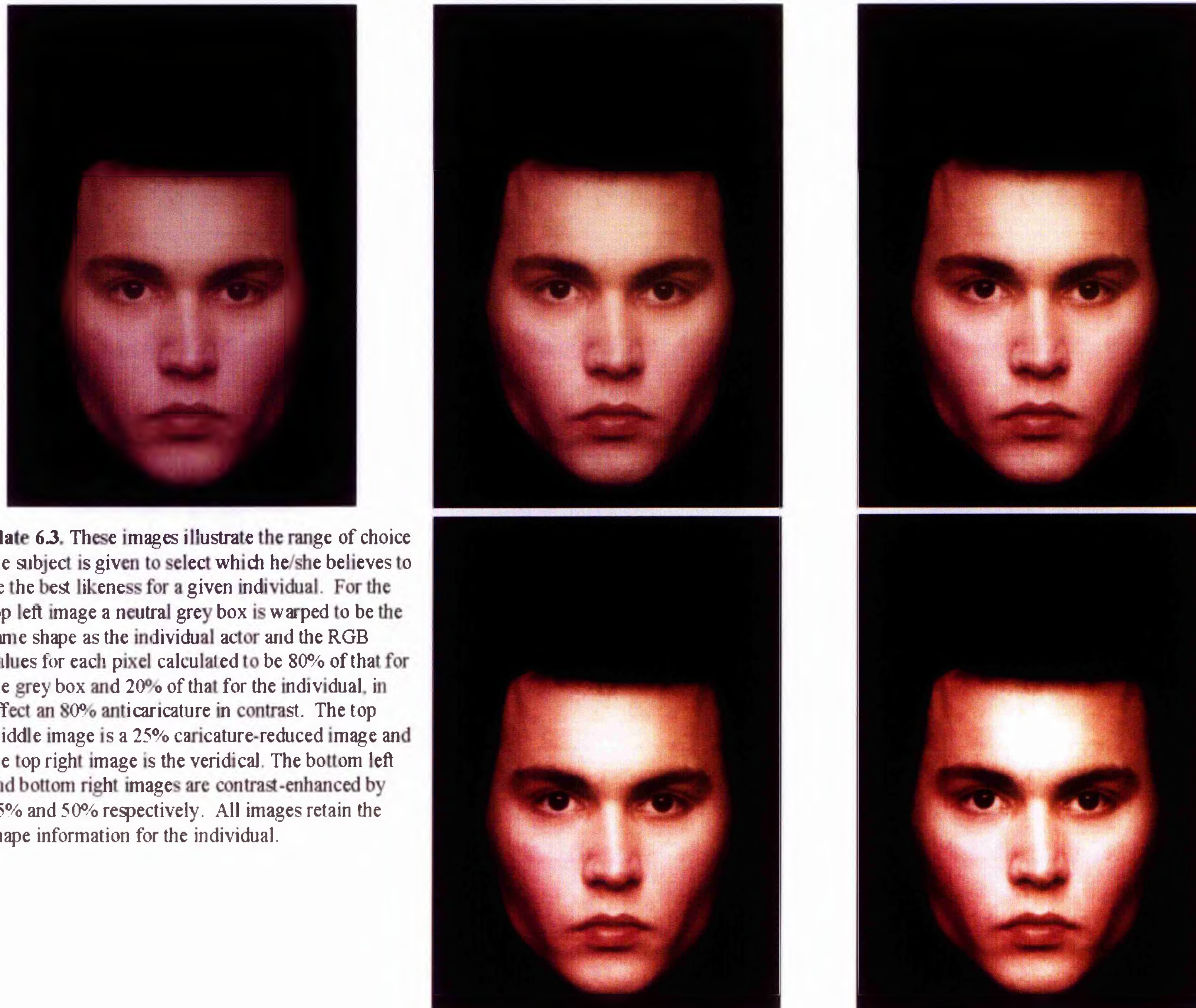


Plate 6.3. These images illustrate the range of choice the subject is given to select which he/she believes to be the best likeness for a given individual. For the top left image a neutral grey box is warped to be the same shape as the individual actor and the RGB values for each pixel calculated to be 80% of that for the grey box and 20% of that for the individual, in effect an 80% anticaricature in contrast. The top middle image is a 25% caricature-reduced image and the top right image is the veridical. The bottom left and bottom right images are contrast-enhanced by 25% and 50% respectively. All images retain the shape information for the individual.

Chapter 7 - Attractiveness (an introduction)

The utility of facial prototyping is not restricted to studies of identity. The information present in a prototype has allowed researchers to study the perception of individuals or groups of individuals in respect to many facial attributes. These include the study of factors underlying the perception of age (Burt and Perrett, 1995), gender (Brown and Perrett, 1993; Bruce and Langton, 1994; Hill *et al*, 1995), race (Hill *et al*, 1995) and facial attractiveness (Grammer and Thornhill, 1994; Langlois and Roggman, 1990; Perrett *et al*, 1994).

Cultural contingencies

It has been stated often that beauty is in the eye of the beholder. It is true that what is perceived as sexually attractive varies greatly from one individual to another. Many factors are likely to influence the development of individual interpersonal relationships. Cultural factors such as fashion, social status, and cognitive factors such as stereotypes and one's own perceived attractiveness are all likely to affect our attractiveness preferences and individual partner choice (Kowner and Ogawa, 1993; Kowner, 1995). Indeed people with similar levels of attractiveness tend to affiliate (Cash and Derlega, 1978; Feingold, 1988; Korthase and Trenholme, 1982; Stroebe, Insko, Thompson and Layton, 1971). Judgements to this extent are thereby subjective.

Western subjects have, however, shown consistency in evaluating attractiveness (Hatfield and Sprecher, 1986; Iliffe, 1960) and this agreement in attractiveness judgements between individuals is pervasive across cross-cultural judgements (Bernstein, Lin and Mclellan, 1982; Cross and Cross, 1971; Cunningham, 1986; Cunningham *et al*, 1995; Johnson, Dannenbring, Anderson and Villa, 1983; Maret; 1983; Maret and Harling, 1985; McArthur and Berry, 1987; Morse, Gruzen and Reis, 1976; Richardson, Goodman, Hastorf and Dornbusch, 1961; Thakerar and Iwawaki, 1979; Wagatsuma and Kleinke, 1979; Weisfeld, Weisfeld and Callaghan, 1984; Zebrowitz, Montepare and Lee, 1993). There has been debate as to the cultural stability of attractiveness. Cognitive theory, on the one hand, states that preferences for attractive faces are acquired early in life through cultural transmission and exposure to standards of attractiveness ascribed by society and the media (see Etcoff, 1994). The evolutionary hypothesis, on the other, suggests that perception of attractiveness should be influenced by similar factors, and thereby be stable, across cultures. Darwin (1871) was struck by cultural differences in preference for differing skin colour and amount of body hair, as well as practices of teeth drilling and filing and lip ornamentation. However, the latter cultural practices may be purely indicative of vanity and courtship measures used within cultures while unadulterated physical appearance may be more indicative of mating success. Variability in some aspects of preferred physique or ornamentation does not preclude the possibility of other universally alluring characteristics (Horvath, 1981; Lott, 1979) so that certain facial configurations could be intrinsically attractive. Darwin (1871) even cited an earlier observation akin to studies in progress with collaborators in the Amazon "Mr Winwood Reade... who has made ample opportunities for observation, not only with the Negroes of the west coast of Africa, both those of the interior who had never associated with Europeans is convinced that their ideas of beauty are, on the whole, the same as ours". Several recent studies have shown that culture or class had little effect on ratings of physical attractiveness and similarly these ratings were not affected by the age of rater (Alley, 1983; Bernstein, Lin and McLellan, 1982; Cross and Cross, 1971; Maret, 1983; Udry, 1965). Physical attractiveness for females is substantially mediated by facial appearance (Alicke *et al*, 1987) and female facial attractiveness has been the focus of the majority of studies on this topic. More recent studies have broadened this scope to the study of male facial attractiveness and also to male and female body morphology. While overall impressions of physical attractiveness are highly reliable cross-culturally (Jackson, 1992), the majority of studies of these comparisons concentrate primarily on facial appearance rather than that of other physical attributes.

There is agreement in such attractiveness judgements between Asian-American and Caucasian females for male and female faces (Wagatsuma and Kleinke,

1979) and between males and females of European, African, Korean, Taiwanese, Japanese, Thai, Chinese and Hispanic ethnic origin living in the United States regardless of ethnicity of the target face (Cunningham *et al*, 1995; Zebrowitz, Montepare and Lee, 1993). Chinese, Indian and English females gave similar judgements for Greek males (Thakerar and Iwawaki, 1979) while South African and African-American males and females (Morse, Gruzen and Reis, 1976) and black and whites American males and females judging black and white male and female faces (Cross and Cross, 1971) also agreed. Cunningham (1986) noted that hair and skin colour were not related to attractiveness in his cross-cultural studies.

Infants aged three to six months shown pictures of female faces, which adults have judged to be attractive or unattractive, show a preference for the attractive ones through displaying prolonged attention (Langlois *et al*, 1987; Langlois, Roggman and Reiser-Danner, 1990; Samuels and Ewy, 1985; Samuels *et al*, 1994; Shapiro *et al*, 1984). Langlois *et al* (1991) found that this similarity to adult preferences extended to infants looking at male faces, faces from a different ethnic background and children's faces. Further, Dion (1973) noted that pre-school children give ratings similar to those of adults when judging other children's attractiveness. Given the substantial differences in structure of male and female, black and white and adult and infant faces (Farkas, 1981) these findings indicate that the face cues that yield judgements of attractiveness seem to be invariant across different types of face and that preference appears in very early life. This is currently being investigated further with neonate judgements (Langlois, personal communication).

The evidence described here suggests that biological and evolutionary consideration play a greater part in attractiveness judgements and that these will transcend cultural considerations (Alley and Hildebrandt, 1988; Langlois and Roggman, 1990; Thornhill and Gangestad, 1993). This indicates the cultural independence of attractiveness judgements and supports the evolutionary hypothesis. Proponents of evolutionary theories contend that mate selection preferences are affected by the biology of reproduction and that desired physical traits and sex differences in mate preferences reflect biological differences regarding reproduction (reviewed in Barber, 1995). As such males and females are likely to use different strategies to determine what they find attractive (Trivers, 1972; Buss, 1985, 1989, 1995).

Male vs. female judgements

In the search for a theoretical basis for attractiveness judgements in humans, sociobiological studies of other animals have been explored. Males of most species increase their reproductive success by mating with as many females as possible, while females rarely benefit from multiple matings (Bateman, 1948). One hypothesis for this gender difference in judgements is the parental investment model (Trivers, 1972, 1985). This proposes that since females have a shorter reproductive life span than do males, males are said to maximise reproductive success by being sexually responsive to a larger number of partners and thereby be attracted to the opposite sex primarily by visual cues (e.g. youth, physical attractiveness) that signal reproductive health (Fisher, 1930) while females would seek to acquire a male with the means to provide and share her parental investment. This argument has been extended to suggest mating motives in humans (Buss, 1987).

It is possible that aspects of female beauty in humans, if the parental investment holds true, would correlate highly with precisely this reproductive quality (Buss, 1987; Symons, 1979) and would suggest youth, health and fertility. Parental investment theory would also hold that women would be more inclined to select as mates men with resources, status and the ability to provide for and protect potential offspring. Indeed women are said to seek non-appearance related factors (cues to status) that maximise the survival prospects of each of their children. This is supported by evidence that women report partners' physical attractiveness as less important than men (Feingold, 1990, 1992; Hatfield and Sprecher, 1986) and they value social status and ambition more than men (Buss, 1989, 1991; Ford and Beach, 1951; Kenrick *et al*, 1990). Indeed, it was found that males were rated as more physically attractive and attracted more female sexual partners if they were in positions associated with greater status and economic wealth (Townsend, 1993; Townsend, Kline and Wasserman, 1995; Townsend and Levy, 1990).

Women are purported to make more comparisons based on social appearance and fashion when selecting a romantic partner than do men (Wheller and Miyake, 1992). It is also likely that females may place less stress on physical appearance per se than do males when evaluating other-sex faces (Berscheid, Dion, Walster and Walster, 1971; Buss, 1987; Feingold, 1990; Townsend, 1989) and this pattern appears across a wide range of cultures (Buss, 1989; Ford and Beach, 1951).

These differences in approach to partner selection may reflect preferences when selecting for long-term partnerships; preferences for short-term 'extra pair' copulations may reflect attractiveness along other dimensions (Gangestad and Thornhill, 1997). Further, sex differences were appreciably larger in the paradigms that examined self-reports of partner preferences than in paradigms that examined behaviour (Feingold, 1990). Partner selection for both sexes is strongly predicted by physical attractiveness (Walster *et al*, 1966) and when asked to make judgements of purely physical attractiveness men and women have a tendency to agree in their judgements of male (Kopera, Maier and Johnson, 1971; Jankowiak, Hill and Donovan, 1992; Morse, Reis, Gruzen and Wolff, 1974; Udry, 1965) and female attractiveness (Cross and Cross, 1971; Johnston and Franklin, 1993; Morse, Reis, Gruzen and Wolff, 1974). Here, it may be that there is a halo-effect such that physically attractive people are perceived as having desirable personality traits (Dion, 1986; Eagly *et al*, 1991; Feingold, 1992a; Langlois, 1986). This aside, more attention does appear to be made to female physical appearance than male appearance cross-culturally (Ford and Beach, 1951) and both men and women consistently give higher ratings of attractiveness to female faces than to male faces (Jackson, 1992; Brewster, 1997). Women also tend to give higher ratings than men to both male and female faces (Cross and Cross, 1971; Kowner and Ogawa, 1995) and it has been argued that the onus may be on females to advertise physical attractiveness in order to obtain a high-status male, increasing the chance of survival of offspring (Barber, 1995; Buss, 1987). Indeed a woman's physical appearance is the most powerful predictor of occupational status of the man she will marry (Buss, 1994). One aspect of facial appearance which may differentially effect attractiveness judgements for females and males is apparent youth or sexual maturity.

Neonate and sexually mature features

It has been argued that youthful features may be of importance in female attractiveness as they signal high fertility and the female reproductive life span is limited by menopause, while for males more mature features may be classified as attractive as they indicate social dominance and possibly a likelihood to acquire status and resources. Youthful features may also be less important for male suitors since youth is less relevant to male fertility than the willingness and capability to invest in dependent offspring (Trivers, 1972). Jankowiak *et al* (1992) reported that men of varying ages had a greater tendency to find younger, undergraduate females to be attractive while females from a range of age groups showed a greater preference for older, postgraduate males. The women in this study further stated that many of the males they found to be unattractive looked "too young and inexperienced" while older men were "better looking because you could talk to them about serious concerns".

It has also been demonstrated by several investigators that younger or more baby-faced appearing adults (both male and female) are seen as more attractive than older, more mature appearing adults (Korthase and Trenholme, 1982; Berry and McArthur, 1985) by both male and female judges. Facial attractiveness decreases as a function of age (Jackson, 1992; Zebrowitz, Olson and Hoffman, 1993). This decline has been found to be greater for female faces than for males (Deutsch, Zalenski and Clark, 1986). More youthful faces may have been preferred because neonate features elicited positive care-taking responses. Equally these neonate features may have conveyed the appearance of youth, health and an extended period of fertility (Symons, 1979). Neonate facial features in male and female faces may also inspire feelings of naivete, honesty, kindness and warmth (Berry and McArthur, 1985, 1986; McArthur and Berry, 1987).

In men, prominent cheekbones and enlarged jaws associated with mature males may also attest to immunocompetence in males (Grammer and Thornhill, 1994) while adult faces have relatively small foreheads and eyes and a larger nose, cheekbones, jaws and chin (Enlow, 1990) than children. Guthrie (1976) and Keating, Mazur and Segal (1981) suggested that maturity features must convey an image of status, power and dominance. Keating (1995) found that females preferred drawings of males with the sexually mature and dominant features of square jaw and thick lips; they also note though that the faces with the most mature features were not rated the most attractive. A man who looks too mature and powerful may not arouse a woman's warm, care-giving emotions and may not elicit as much attraction as a man who can stimulate nurturing responses through the display of slightly more neotenus features such as large eyes (McArthur and Apatow, 1983-1984; Cunningham *et al*, 1990). Cunningham *et al* (1990) also argue that the relationship in facial features between maturity and neoteny found to be attractive in male faces may also be apparent in attractive female faces.

In addition to aiding male attractiveness through conveyance of maturity and social dominance some sexually mature facial features may contribute to female attractiveness. Responsiveness to a few mature features such as high cheekbones or cheeks which have lost some of their "baby fat" may have evolved to discourage paedophilia and ensure that advances were made only to post-pubescent females. Prominent cheekbones, narrow cheeks and large lips in women indicate the onset of oestrogen production and therefore the beginning of sexual maturity and, perhaps, immunocompetence (see Grossman, 1985; Singh, 1993, 1995) and preference for these features highlights desire for a degree of sexual maturity in women (Johnston and Franklin, 1993; Alley and Cunningham, 1991).

Johnston and Franklin (1993) find an interesting trade-off occurring such that certain youthful dimensions are found in a face which is holistically judged to be of the age of a fairly mature female. Allowing subjects to generate their ideal facial stimuli from a series of composite faces using a genetic algorithm, Johnston and Franklin were able to show that the ideal female faces generated by female and male individuals alike differed from the average of the population from which they were derived in several consistent, key ways. They describe the beautiful faces selected in their study to be "...of highly fertile age, suitable for mates, but [having] the facial characteristics of a much more youthful female" (page 196). The mean estimated age for the faces selected as beautiful was 24.9 years, roughly the same as that, 24.8 years, found to be most attractive in a survey across 37 cultures (Buss, 1989). This overall impression was over two years younger than the mean age estimation of the faces from which they were evolved (estimated mean age 27.4; genuine mean age 19.9) while mean lip size of the beautiful faces was characteristic of 14 year-old females and the nose-chin distance was shortened to that typical of an 11-year old as derived from the Farkas (1981) growth curve. Eye-chin and eye-nose distances in the generated beautiful faces were also characteristic of a more youthful face. Consistent with this finding, Henss (1991) reported that the women rated as most attractive in his study were those most underestimated in age and Jones (1995) found the faces of more attractive women to have a shape which was calculated to be neotenous relative to that typical for their age. So, it could be predicted that highly attractive faces would appear to be of an age at which fertility is near maximum while having some facial characteristics of a more youthful face. A more attractive woman may be one who is sexually mature yet whose features have had relatively little influence from the influx of adrenal androgens which lead to facial growth in women at puberty (Bermant and Davidson, 1974; Enlow, 1990).

The subtleties of such a compromise and the nature of an idealised balance in the configuration of features between maturity and neoteny have been analysed further and summarised by Cunningham (1986; Cunningham *et al,* 1990). Using black and white photos, Cunningham (1986) found males to be attracted to females possessing neonate features of large eyes, small nose area, small chin and widely spaced eyes while nose size made no difference for female preferences in men. Here, a large chin was found to be more attractive whilst a small chin was preferred in women. On a cultural note, hair and skin colour was not related to attractiveness judgements in Cunningham's studies. When faces are subjectively rated for babyishness, there is a positive correlation for attractiveness for female faces but not for male faces (Berry and McArthur, 1988). This further suggests that youthful features are found to be more attractive in females while larger chin as typified by post-pubertal growth in children is found to be more attractive in males. Thus Alley and Cunningham (1991) have argued for multiple motives in judgements of attractiveness - with some components signalling neoteny and others signalling maturity. The perceptual changes resultant from ageing are not necessarily limited to shape characters already previously described alone and ageing may also differentially affect skin complexion and other facial and body traits.

Changes in skin texture and colouration with age may influence perceived attractiveness. As women age, increasing androgen levels turn on genes for growth of body hair. Artificial removal of body hair through shaving or waxing may enhance perceived youth and this may have a bearing on perceived attractiveness. Relative body hairlessness has been found to be a sexually attractive feminine trait (van den Berghe and Frost, 1986). There is a tendency for skin pigmentation to increase with age in most racial groups (Darwin, 1871; Frost, 1988). Skin pigmentation is also slightly darker in males and female skin darkens during pregnancy and during infertile phases of the menstrual cycle (van den Berghe and Frost, 1986). Van den Berghe and Frost found that in 57 societies across Europe, Africa, Asia, the Americas and the South Pacific, 47 preferred lighter-coloured skins. Of these, 30 showed a preference for light-skinned women exclusively, 14 preferred light skin in both sexes and 3 showed exclusive preference for light-skinned men. Feinman and Gill (1978) also found that American males showed a preference for lighter complexion and eyes in women. However, Feinman and Gill relied on self-report, relying on the imagination of the subject, and did not use facial stimuli in their study. Hulse (1967) found a preference for lighter skin in Japanese men and women. The hormonal changes which occur at puberty and lead to the development of sexually mature features, primarily for males, may also affect another trait said to be connected with perceived attractiveness: symmetry.

Symmetry and pathogen resistance

The role of symmetry in attractiveness rests on the notion that one important aspect for which human attractiveness acts is sexual selection for individuals that are best at resisting pathogen and parasite infection (Hamilton and Zuk, 1982; Thornhill and Gangestad, 1993). This rests on the assumption that such resistance is heritable and thus individuals would select mates who displayed features indicating a healthy predisposition and pathogen resistance that would be passed on to any potential offspring. Across cultures, people find explicit evidence of disease and deformity unattractive (Ford and Beach, 1951; Richardson *et al*, 1961) and pathogen prevalence correlates substantially with the importance of physical attractiveness across societies (Gangestad and Buss, 1993). The theoretical basis for symmetry as an indicator of health and thereby a cue to attractiveness proposes that development of testosterone-mediated structures honestly advertises the pathogen resistance capabilities of individuals. The reasoning behind this is that only the most fit individuals can afford the handicap of compromising their immune system by growing such structures since malnutrition and parasites acting on an impeded immune system are capable of disrupting morphological development (Fölstad and Karter, 1992; Gangestad, Thornill and Yeo, 1994; Møller and Swaddle, 1997; Watson and Thornhill, 1994; Wedekind, 1992). The developmental stress imposed in growing exaggerated secondary sexual traits can lead to substantially higher levels of fluctuating asymmetry in these traits than in other morphological features (Møller and Hoglund, 1991) with highly fit individuals thereby capable of growing both larger and symmetrical display traits. Fluctuating asymmetry is defined as asymmetry resulting from errors in the development of normally symmetrical bilateral traits under such stressful conditions. Such traits are normally distributed between the left and right sides of individuals within the population with a mean of zero (Van Valen, 1962). Thus both mean size and asymmetry of secondary sexual traits may be used in mate choice (Møller, 1992, 1993). Indeed the size of secondary sexual traits is positively correlated with their symmetry while for other bodily traits this correlation is zero or negative (Watson and Thornhill, 1994). The role of symmetry in male characteristics has been demonstrated by sexual selection in favour of symmetrical features for insect (Radesäater and Halldórsdóttir, 1993) and avian (Møller, 1992; Swaddle and Cuthill, 1994).

Many human facial features undergo testosterone-mediated development, especially in males (Tanner, 1978; Enlow, 1990; Thornhill and Gangestad, 1993) who develop enlarged jaws, chins and cheekbones. If these structures do reliably reveal quality, humans seeking potential mates should assess not only the mean size of such secondary sexual traits, but also their symmetry. In line with this both trait size and facial averageness have been found to influence perceived attractiveness (Alley and Cunningham, 1991; Berry and McArthur, 1995; Cunningham, 1986; Hess, 1965, 1975; Langlois and Roggman, 1990; McAfee *et al*, 1982; Terry, 1977).

Symmetry may be perceived as being relatively more attractive in male than female faces as the 'testosterone hypothesis' predicts that deviations from facial asymmetry should be more apparent in males who preferentially gain an influx of testosterone with the development of secondary sexual features at puberty. Also, females may be more sensitive to symmetry in their assessment of mates. This follows from both the presumed higher susceptibility of males to fluctuating asymmetry in testosterone-mediated facial traits and the possible higher cost to females in making 'bad' mate-choice decisions (Andersson, 1994). Male body symmetry has been correlated with the number of sexual partners reported over the previous year (Gangestad and Thornhill, 1997; Thornhill and Gangestad, 1994; Watson and Thornhill, 1994). Gangestad *et* al (1994) found that the attractiveness of a male face predicted relative lack of bilateral asymmetry with respect to size of feet, ankles, elbows, wrists, ears and so forth. There is also some evidence that fluctuating asymmetry may also be linked to intelligence (Furlow *et al*, 1997), a trait women are likely to value in prospective male partners as an indicator of ability to acquire resources and care for offspring. Both male and female body symmetries have also been linked to reproductive success and there may be sexual selection for such physical traits (Møller, Soler and Thornhill, 1995; Singh, 1995). This indicates that facial beauty can be used as an index of phenotypic quality in the form of fecundity and heritable resistance to pathogens. It has been suggested that this could be evident in relative asymmetries in faces leading to lower attractiveness ratings.

However, studies using reflection of one half of the face along the midline to create symmetrical left-left and right-right stimuli have found that perfectly symmetrical faces of this nature are less attractive than veridical faces (Langlois *et al*, 1994; Kowner, 1996). Using a different procedure in which individual faces were manipulated to have varying degrees of asymmetry and subjectively ranked for attractiveness, Swaddle and Cuthill (1995) verified the finding that reducing the level of asymmetry also reduces the level of attractiveness. Langlois, Roggman and Musselman (1994) also found that faces which were improved with enhanced symmetry were those which were rated the least attractive to begin with. In all these studies it is difficult to determine if it is general asymmetry which is being reduced rather than fluctuating asymmetry; reducing all asymmetries, not just those linked to secondary sexual feature development, may make faces appear more unnatural and therefore less attractive. Alternatively, the lack of superior attractiveness ratings for symmetrical faces in these studies may result from technical confounds in the image processing techniques used to make faces more symmetrical in the studies. When making the faces symmetrical, not only the shape but the texture was altered such that, using the Swaddle and Cuttle (1995) technique in which the shape and texture information was averaged to varying degrees between the two sides of the face for instance, a spot originally appearing on the left side of the faces would be evident with half the original intensity on both the left and right side of the face in a fully symmetrical image. Perrett *et al* (submitted) controlled for this possibility by varying the degree of symmetry of the shape of an individual face while retaining either the original facial texture or that derived from the average of a group of faces of that gender. Using stimuli manipulated in this way, male and female symmetrical faces retaining individuals' original textures were found to be more attractive than the original faces with symmetrical faces featuring average texture information for a given gender rated more attractive still. Contrary also to Langlois *et al's* (1994) finding the effect of symmetry on attractiveness was not found to depend on the face's initial attractiveness.

Expressive features

Certain expressive features are likely to be attractive. In addition to finding males to be attracted to females with the mature features of wide cheekbones and narrow cheeks and the neonate feature of large eyes, the expressive features of highly set eyebrows, wide nostrils and pupils and a large smile have also been linked strongly to attractiveness (Cunningham, 1986; McGinley *et al*, 1978; McGinley *et al*, 1984). Raised eyebrows often signal interest, greeting and submission (Eibl-Eibesfeldt, 1970; Nakdimen, 1984). It is possible that individuals, both male and female, whose eyebrows are set high on their foreheads convey the image of a positive attitude and are seen to be more attractive. Dilated pupils have also been found to increase attractiveness ratings of faces by viewers (Hess, 1965, 1975). Jankowiak *et al* (1992) also noted that women found faces to be attractive which looked "happy" and "thoughtful". Expressive features such as large eyes, large smile and high eyebrows may again be a more vital cue to attractiveness for female faces than for males. These features are a further examples of a sexual dimorphism between males and females as males have a lower eyebrow ridge and more prominent eyebrows through the testosterone-expression and have a tendency to smile less often than females. Both these characteristics in males may correlate with aggression, status and dominance.

Gender typicality

There is also evidence that a degree of gender typicality may play a role in the perception of attractiveness in the process of sexual selection. Attractive people of both sexes show more sex-role stereotyped behaviour (Brehm, 1985; Reis *et al*, 1982) although this may be because they attract more sexual interest. There is evidence of a strong link between the morphology of apparent femininity and apparent ageing in women as a result of the sudden influx of oestrogen at puberty and the gradual increase of androgen level following this until menopause. The small chin and nose associated with neoteny could be equally suggestive of the low androgen level of a fertile female with reproduction value attributable to heritable hormone regulation. Such hormonal changes with males are less dramatic and thereby post-pubertal ageing is less likely to be associated with the relatively invariant changes in testosterone levels. In this way, changes which occur with ageing may be related to female gender typicality at least such that as women age post-puberty their reproductive life span decreases; their androgens levels, particularly for free testosterone, increase with a gradual reduction of sexual dimorphism resulting in features which are less typical of their gender and thus indicative of decreasing fertility.

It has already been suggested that women may adopt features which will directly indicate reproductive fitness. These have not been discussed fully in the prior commentary on youthfulness and sexual maturity. Several facial and body features have been investigated as possible indicators for female fecundity, but before now investigation has tended to concentrate on body measures. The measurement typically used in this type of study has been the waist-to-hip ratio which reflects body fat distribution. The distribution of fat in the body is influenced by levels of circulating sex hormones such as oestrogen and testosterone (Evans *et al*, 1983) and this provides an index of endocrine readiness for reproduction. Waist-to-hip ratios range from 0.67 to 0.80 in women and 0.85 to 0.95 in men with little overlap between genders but this sexual dimorphism is lost at menopause. Studies using this waist-to-hip ratio measurement have consistently found that line-figure representations of females are found to be most attractive with waist-to-hip ratios (WHRs) at the low end of this range while representations of males with the highest waist-to-hip ratios were found to be the most attractive (Henss, 1995; Singh, 1993, 1994). These judgements have been found to be consistent across culture (Singh, 1994; Singh and Luis, 1995;

personal communication from Singh cited in Henss, 1995). In summary, gender typical female features may serve to highlight reproductive potential and fertility; waist-to-hip ratio is such a trait and males are likely to be sexually attracted to such characteristics (Symons, 1979).

The health benefits associated with a low WHR also give further weight to the suggestion that oestregenised features, such as prominent cheekbones in the female face, may be an indicator of immunocompetence in females (Grossman, 1985; Grammer and Thornhill, 1994). Other morphological traits, some of which are detectable in the faces alone and may include pale skin (Frost, 1988), prominent cheekbones and high-set eyebrows may also indicate a high oestrogen-androgen level and therefore high fecundability. The full lips (resultant from high oestrogen) and small eye-chin, eye-nose and nose-chin distances (associated with low androgen level) in the beautiful faces reported by Johnston and Franklin (1993) are also those which would be expected to result from very high oestrogen-androgen levels in a female adult; these levels are demonstrated by low WHR.

Cunningham *et al* (1990) found differences between attractive features in male and female faces which reflected gender typicality. Large chins were found to be desirable for men whereas smaller chins were found to be desirable for women. Again, the factors which are typical of gender and contribute to facial attractiveness may not be limited to shape. Across a wide range of cultures in Europe, Asia, Australasia, Africa and the Americas women have been shown to have lighter skins than their male counterparts (Conway and Baker, 1976; Hulse, 1967; Kahlon, 1976; Kalla and Tiwari, 1970). Lighter skin is consistently found to be more attractive in females across this range of cultures (van den Berghe and Frost, 1986) and this may be through the revealing of fertility through gender typicality. Melanin production is thought to be controlled by the anterior pituitary gland which also secretes the gonadotropins which become sexually dimorphic at puberty. So it may be that light-skinned women are judged to be attractive because lighter skin tone once more reflects low levels of androgen in sexually mature females and as such these females are presumed to be highly fecund.

Recent experiments have investigated symmetry and other possible influences on attractiveness with composite faces considered to be appropriate stimuli for experimental study.

Composite faces as experimental stimuli

Many studies have been reviewed here which have shown consistency across culture for judgements in attractiveness and the theoretical issues underpinning the facial traits which may be vital to attractiveness. Recent experimental investigations have, therefore, moved away from the use of individual stimuli and begun studying facial attraction using composite stimuli. In line with the evolutionary hypothesis, Symons (1979) proposed an innate mechanism that detects the population mean of anatomical features. In the case of faces, he proposed a "beauty detecting" mechanism which averages observed faces: that the selection pressures are assumed to favour built-in preferences for these average faces over preferences for faces more distant from the mean. The mechanism hypothesised by Symons could be used in support of an evolutionary hypothesis such that we faces are average and do not deviate too much from the population mean are found to be more attractive. Alternatively, a cognitive approach could be used to suggest that we are attracted to a face type with which we are familiar and with which feel comfortable. As a result, we may find faces to be attractive which do not stray much from the configuration central to that for faces which we have experienced and are construed as normal. Langlois and Roggman (1990) tested the notion of a preference for average faces empirically by comparing attractiveness judgements for composite and individual faces, finding that composite or average faces were rated as more attractive than individuals and suggesting average facial features would be those found, on the whole, to be most attractive by viewers using photographic composites. They found that a composite was, on average, found to be more attractive than its constituent faces and that the blend was found to become more attractive as the number of faces incorporated into the composite increased. This, in line with evidence derived from the studies of caricaturing with recognition (reviewed in *Chapter 1*), may explain the finding of Light, Hollander and Kayra-Stuart (1981) that females judged to be most attractive may have such similar features as to be difficult to distinguish from another and hence be closer to the average or prototypical face (for further discussion, see *Chapter 1*). Grammer and Thornhill (1994) replicated Langlois and Roggman's (1990) finding and postulated that this supported the notion that symmetry is more attractive in faces: with an increased number of component faces the composite becomes more symmetrical. They argue that immunocompetence is generally better in heterozygotes; heterozygosity is generally associated with phenotypic averaging and therefore traits that approximate population means should generally be more average than extreme (Soule and Cuzin-Roudy, 1982; Thornhill and Gangestad, 1993; Watson and Thornhill, 1994). However, the production of facial composites alters not only symmetry but trait size and it is therefore difficult to isolate the role of symmetry in this instance. There may also be confounds with the methodology used in these studies.

The averaging process used in these studies also causes smoothing of the skin which removes blemishes and may give a youthful appearance, perhaps inadvertently providing a more key cue to attractiveness. Skin quality, which is affected by this process, has also been linked to attractiveness (Jackson, 1992). Indeed Langlois, Roggman and Musselman (1994) attest that the average face appeared younger than the individuals used in the composite and Benson and Perrett (1991a) found that applying the average skin texture to an individual's face shape made the individual appear more attractive. This is evidenced further from the long-established practice of using soft-focus in film-making to make an individual appear both younger and more attractive.

The studies of Langlois, Roggman and Musselman (1994) and others have found symmetry to detract from perceived facial attractiveness, suggesting that a more attractive configuration is non-symmetrical and, thereby, non-average. Studies by Perrett *et al* (1994) present forced-choice judgements in which the preferred face is deviated from the population average. Their results showed that the shape of a female composite derived from 15 faces rated highly for attractiveness was found to be more attractive than that derived from a blend of 60 faces, which included these 15 (*Figure 7.1*). Here both composites featured the texture information averaged across all 60 faces. Further, a 50% caricature of the shape differences between this attractive blend and the overall blend was found to be yet more attractive (*Figure 7.2*). The stimuli presented in Perrett *et al*'s study are illustrated in *Plates 7.1 and 7.2*.

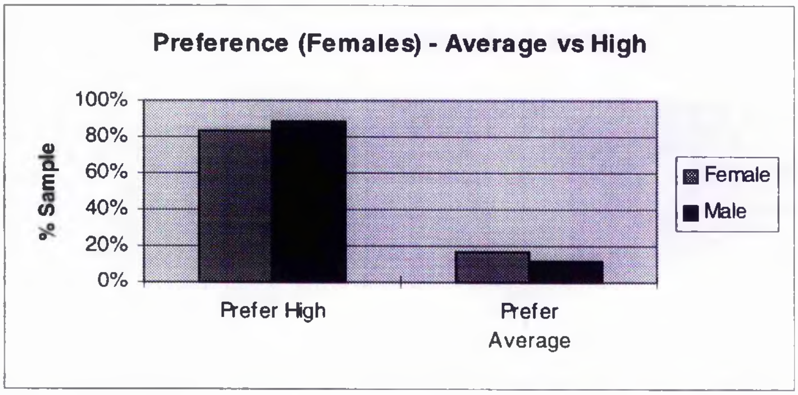


Figure 7.1. High vs. average Caucasian female shape. Male and female Caucasian subjects consistently found the shape derived from the highly attractive Caucasian female to that derived from the population. This pattern was also found with Japanese subjects viewing female Caucasian faces and both Japanese and Caucasians subjects viewing female Japanese faces.

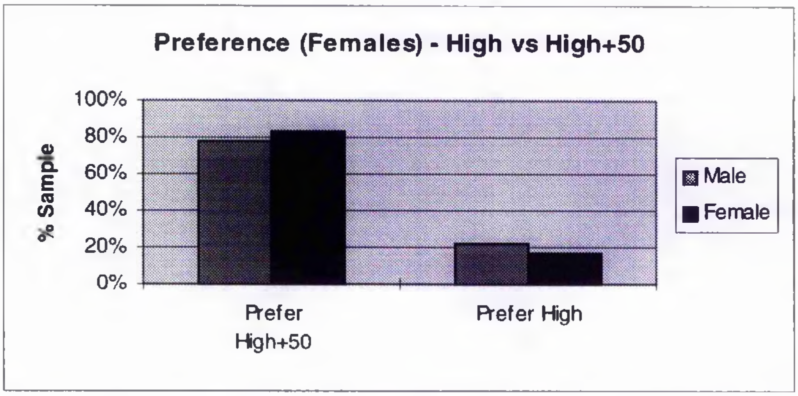


Figure 7.2. Caricatured vs. high Caucasian female shape. Male and female Caucasian subjects consistently found the shape derived from exaggerating the shape differences between the highly attractive composite and the average composite by 50% to be more attractive than the highly attractive Caucasian female. Again this pattern was consistent cross-culturally.

This indicates that the most attractive facial configuration lies along a vector which deviates from the norm. The more attractive female composites in the study by Perrett *et al* (1994) possessed the sexually mature features of higher than average cheekbones and thinner than average jaws in conjunction with the youthful features of larger than average eyes and fuller lips, associated with femininity and high oestrogen levels (*Figure 7.3 and Figure 7.4*).

Figure 7.3. From Perrett *et al* (1994). The difference between the average female Caucasian shape (a) and the highly attractive female shape (b) can be plotted (c)

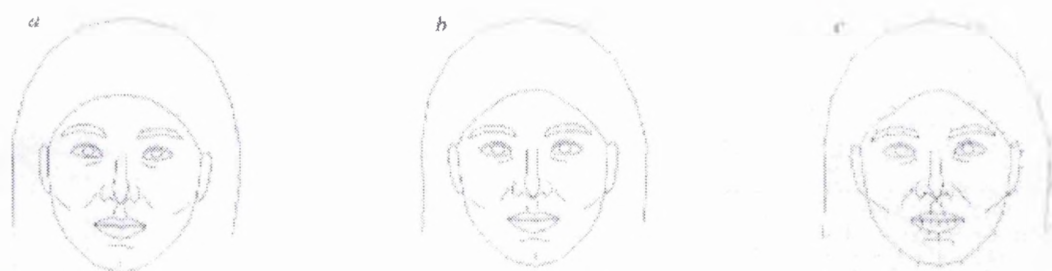


Figure 7.4. From Perrett *et al* (1994). Again the difference between the average female Japanese shape (a) and the highly attractive female shape (b) can be plotted (c). These shape differences are similar to those described for the Caucasian female.

This finding was supported cross-culturally with Japanese and Caucasian faces and Japanese and Caucasian viewers. Perrett *et al* found that this preference for non-average facial composites was not limited to female faces; the shape of a composite made from the quartile of males rated as most attractive was judged to be more attractive than an average derived from 60 individuals (*Figure 7.5*).

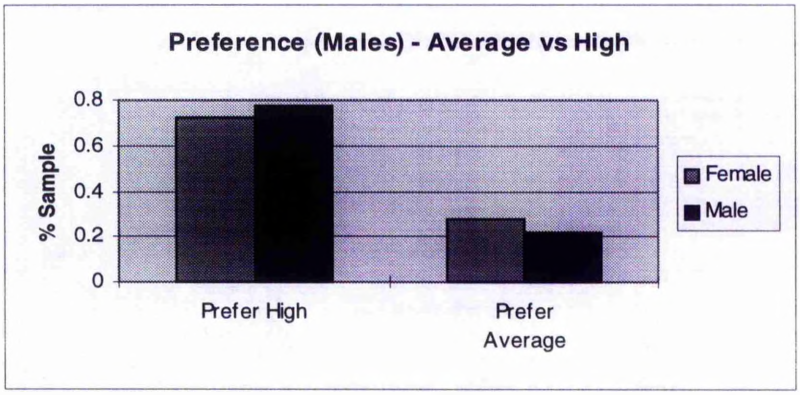
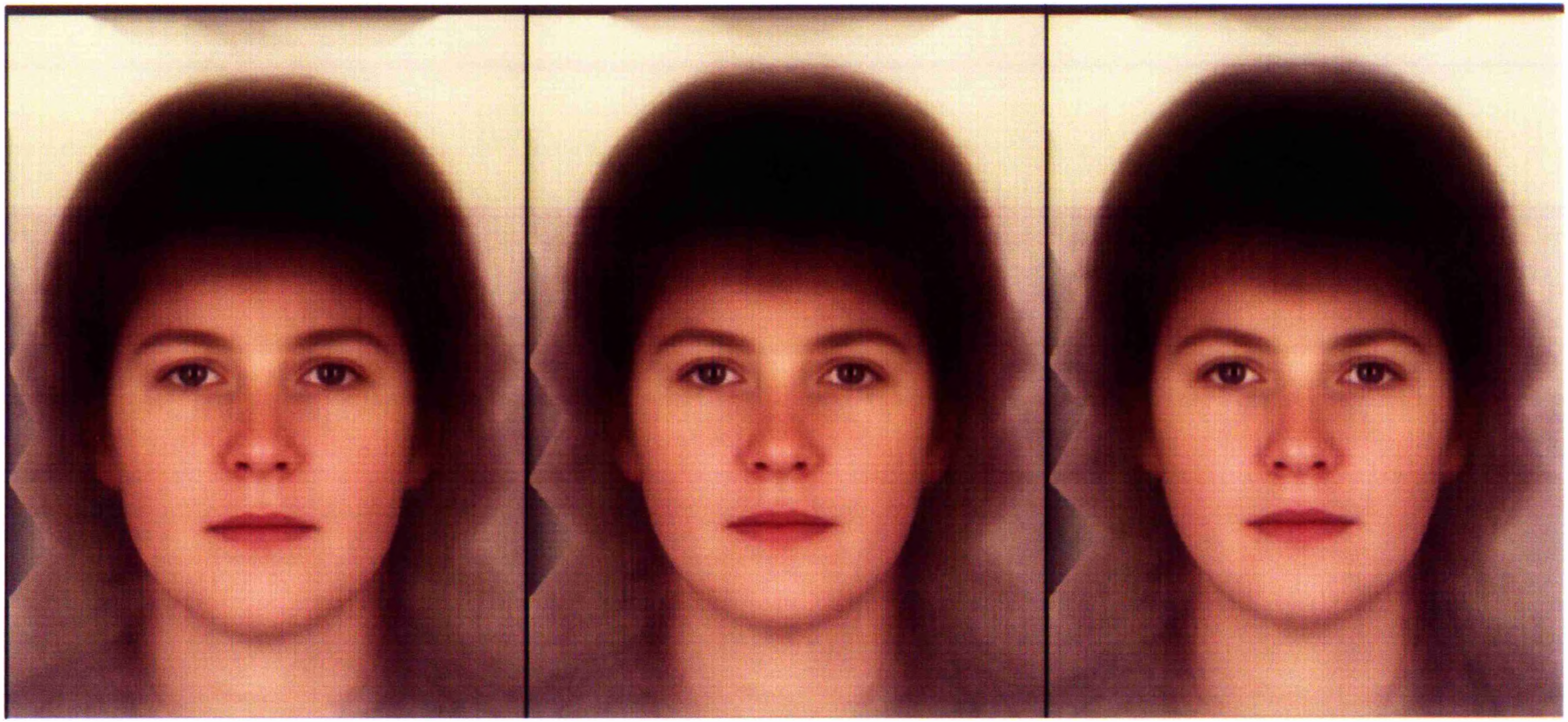


Figure 7.5. High vs. average Caucasian male shape. Male and female Caucasian subjects consistently found the shape derived from the highly attractive Caucasian male to be more attractive than that derived from the population.

Using composite portraits, Galton (1907) found beautiful features to deviate in specific ways from that which is typical of the population. These same features appear to be the common focus of cosmetic, orthodontic (Korabik, 1981) and rhinoplastic surgery (Cash and Horton, 1983). The experiments described here investigate one holistic facial factor which might affect these judgements. The variables influencing attractiveness probably include 1) neonate features; 2) mature features - it is expected that higher, wider cheekbones, and narrower cheeks would be related to greater perceived attractiveness; 3) expressive features - predicted that individuals whose features are particularly effective at signalling positive emotions, with larger smiles, higher eyebrows, and larger pupils would be seen as more attractive; 4) symmetry; 5) gender typicality. The experiments in the chapter which follows concentrate on the latter.



a b c

Plate 7.1. From Perrett *et al* (1994). The study was performed cross-culturally using colour Caucasian faces. The composite created from a blend of the quartile of female Caucasian faces rated as most attractive (b) was rated as more attractive than the composite generated from the population average. In turn the composite developed by caricaturing the difference in shape between the high attractive image and the population average (c) was rated as more attractive than the high attractive blend.

a b c

Plate 7.2. From Perrett *et al* (1994). The study was performed cross-culturally using black and white Japanese faces. The composite created from a blend of the quartile of female Japanese faces rated as most attractive (b) was rated as more attractive than the composite generated from the population average (a). In turn the composite developed by caricaturing the difference in shape between the high attractive image and the population average (c) was rated as more attractive than the high attractive blend.

Chapter 8 - Experiments Eight and Nine: Attractiveness and Effects of Manipulating Sexual Differences

Introduction

Chapter 7 described many of the attributes thought to be linked with judgements of human facial attractiveness. Amongst these were listed features associated with neoteny, sexual maturity, gender typicality, and possible indicators to health such as fluctuating asymmetry. The experiments which follow in this chapter investigate one such cue, gender typicality. This cue may relate to other aspects of attractiveness such as health and a trade-off between neoteny and sexual maturity. The affect of gender typicality in attractiveness is investigated here within- and across-gender and within- and across-culture such that both females and males living in Japan and Scotland were asked to choose faces they found to be attractive for both males and females of their own and other ethnic backgrounds. Further, in the experiment described in this chapter individuals of native African and Asian backgrounds living in South Africa were also asked to make such attractiveness judgements. From the review in Chapter 7 several predictions might be made on the outcome of the experiments to follow.

Masculinising/feminising of stimuli

It may be predicted that female face stimuli which had a more feminine shape configuration would be judged to be more attractive than those with average or more masculine features. The more attractive composite faces in Perrett *et al*'s (1994) study featured higher cheekbones, enlarged eyes, and smaller chin and nose than the average faces. Small noses and chins have also been found to be attractive in previous studies of female attractiveness and such features are associated with sexual dimorphism; at puberty males noses and chins become enlarged, while female noses and chins change by a relative small amount relative to the rest of the face. High cheekbones in women are also suggestive of high oestrogen levels. Such exhibitors of femininity may all be indicators of fertility in women. Therefore, enhancing femininity of female faces should increase attractiveness.

With male faces, strong cheekbones and features associated with masculinity and testosterone production such as large chin, wide cheekbones and nose may suggest dominance and status. Such features may be found to be attractive. Alternatively, a smaller chin and larger eyes more typical of females may be indicative of neoteny and elicit care-giving responses in potential female partners. Further, more feminine face shapes may be more expressive and such expressive traits may be found to be attractive by prospective mates. In the experiments which follow in this chapter it is predicted that for female faces, the more feminine of the stimuli would be most consistently preferred. For male faces, the predicted direction of preference is not so clear. Should subjects choose more masculine face shapes, this would support the hypothesis that masculine features are more attractive either through conveying social status and dominance or, in conjunction with the maintenance of symmetry, by indicating health through immunocompetence. Should more feminine faces be chosen this may indicate that male faces are either preferred which reflect expressive features such as intelligence or kindness or that more feminine faces are seen as less aggressive. Alternatively, feminine features may be correlated in some way with neoteny, the preference for feminised stimuli for both sexes of faces may reflect a preference for more youthful appearing individuals.

We may also expect judgements involving male faces to be more variable than those for female faces as it has been suggested that physical attractiveness plays a less important role mate selection by females (Buss, 1987, 1989, 1991). Alternatively, it has been argued that male physical attractiveness is used as a reflection of immunocompetence, status or personality traits such as intelligence and kindness. One would predict from this line of research that male faces with a more feminine shape which appear to be more expressive and neonate would be selected as they indicate such positive personality traits in the individual. These are likely to be selected for by females looking for a prospective mate.

Preference differences between male and female viewers and male and female stimuli

A prediction might be made for differences in preferences between male and female judges. It is possible that female subjects would be more divided than males over their preferences for female stimuli since males are likely to rate attractiveness which evoke sexual interest while females need not be motivated by partnership prospects. However, women may make similar judgements to men (Morse, Reis, Gruzen and Wolff, 1974). In addition to any personal preferences they might have, they are likely to have acquired some knowledge of the aspects of a female face which males find to be attractive. This could be particularly influenced in developed countries, such as Japan and the United Kingdom, by the impact of the media and commercial considerations regarding women whereby both genders are made aware of enhancement to particular facial features which increase perceived attractiveness in female faces (Etcoff, 1994). Again, in the Perrett *et al* (1994) study both men and women found the same, non-average, face to be the most attractive. Thus, the most likely outcome in the experiments described below is that males and females would agree in their attractiveness judgements.

Potential cultural effects on attractiveness judgements

A plethora of research has shown consistent cross-cultural agreement in ratings of attractiveness for individual faces. Cross-cultural agreement has also been indicated using composite stimuli by Perrett *et al* (1994). Further, children of very young ages have shown the same attractiveness preferences as adults both within- and across-culture (e.g. Langlois *et al*, 1991, Samuels *et al*, 1994). Thus it is likely that there will be cultural agreement in attractiveness judgements for the faces used in the experiments described here.

Very little rigorous experimental work has been carried out into the identification of facially attractive traits. Much of the previous research which has attempted to identify the cues which are of importance to facial attractiveness have used quasiexperimental, facialmetric designs (e.g. Cunningham, 1986; Cunningham *et al*, 1990), such that correlations were taken and there was no control for any given variables. Alternative research strategies have used methodology which can be criticised because of the comparison of composite faces with unaltered individual faces (Langlois and Roggman, 1990). The experiments which follow investigate cues underlying facial attractiveness using a more experimental paradigm comparing like with like, with manipulation being made upon one holistic trait - gender typicality - using composite images alone. In this manner any of the confounds possibly associated with the blending process (Benson and Perrett, 1991a; Pittenger, 1991) will be immaterial as all stimuli to be viewed are such blended faces. The images used for the experiments described here are also in full-colour and of much higher resolution than photographs (Cunningham, 1986; Cunningham *et al*, 1990) and images (Langlois and Roggman, 1990; Perrett *et al*, 1994) used in previous studies; this enabled delineation with greater accuracy and the production of superior, realistic blends. It is hypothesised that feminised female stimuli will be perceived to be attractive as such feminised features may reveal characteristics such as youth and attractive features such as high cheekbones which have been seen as desirable in previous studies. Further, it is hypothesised that both feminised male and female stimuli will be found to be more attractive as female faces appear to be mildly more expressive. Expressiveness in the mouth and eyes area has often been quoted as being desirable in a partner (Cunningham, 1986; McGinley *et al*, 1978; McGinley *et al*, 1984), possibly suggesting personality traits such as warmth and thoughtfulness. The changes in face shape associated with shifts exaggerating or reducing gender typicality may also incorporate the features underlying some of the other factors found in an attractive facial configuration including neoteny and sexual maturity (Cunningham, 1986; Cunningham *et al*, 1990) through indication in variations of hormone regulation. Thus increasing masculinity may age faces and increasing femininity may decrease perceived age. One part of the study described in this chapter investigates perceived age changes associated with manipulations of gender typicality. Two experiments were carried out; Experiment 8 which investigated attractiveness using static face stimuli and Experiment 9 which explored the attractiveness of stimuli using an interactive technique similar to that employed for facial identity in Chapter 6. Experiment 8 is divided into two parts - Part One involves the study of Japanese stimuli within and across-culture, while Part Two employs white British stimuli. The terms Japanese and white British used to describe the ethnic background of faces and judges follows the convention of Valentine and Endo (1992). Japanese refers to students studying in Japan or upon recent arrival to Britain with an Oriental ethnic background while white British refers to individuals who are studying in Britain with a European, or white, ethnic background.

Experiment 8 - Static Stimuli

General Experimental Methods

Apparatus. Silicon Graphics Indigo R4400 with Extreme graphics card was used for the production of stimuli. Stimuli were printed using a Mitsubishi S3410-30 sublimation full colour printer.

Part One - Japanese stimuli

Method

Subjects. 183 Japanese undergraduates at the Otemon-Gakuin University in Osaka, Japan and 17 students in a foundation course at the English Language Centre, University of St Andrews, Scotland, aged 18 to 27 (median 19); 49 female and 151 male. 153 of the students at Otemon-Gakuin University made age judgements in addition to attractiveness judgements and took part in Part Two simultaneously.

87 white British undergraduates and visitors at the University of St Andrews; 5 female and 22 male, aged 15 to 49 (median 20). Age judgements were made by 49 undergraduates at the University of St Andrews; 38 female and 11 male. 38 of the subjects also took part in Part Two simultaneously.

181 undergraduates at the University of Natal, South Africa; 132 female, aged 17 to 41 (median 19) and 49 male, aged 17 to 35 (median 20). These were divided into three groups based on ethnic background; 78 native Africans, 54 female, aged 17 to 35 (median 20) and 24 male, aged 17 to 35 (median 23); 51 Whites, 34 female, aged 17 to 41 (median 21) and 17 male, aged 17 to 23 (median 20) and 52 Asians (mostly of Indian extraction), 44 female, aged 17 to 22 (median 18) and 8 male, aged 17 to 20 (median 18). 67 of these subjects took part in Part Two at the same time.

Stimuli. Photorealistic average images of faces were created for Japanese faces. Composites were created to represent prototypical, or average, males and females. The blends were produced from 28 male (aged 20 to 23 years, mean 21.6) and 28 female (aged 20 to 22 years, mean 21.4) Japanese individuals studying at Otemon-Gakuin University. Their photographs were taken using Kodak 400 ASA 35mm print film in landscape format under standard lighting conditions with a single flash set on the camera. The photographs were stored on Kodak Photo-CD at 2048 x 3072 pixels in 24-bit colour. Each image was cropped to a resolution of 1961 x 2001, such that an individual's face occupied most of the image area, and delineated by one operator to ensure consistency across the faces. First, the average shape for each gender was derived from averaging the feature point data for each of the 28 faces in the group (see *Chapter 1* for further details). The average male shape was then normalised to have the same inter-ocular distance as the average female (*Figure 8.1*). In order to produce the average texture information for each face gender each constituent image was warped to this average shape. Each of these images was then blended to form a composite image which incorporated the average shape and texture information for Japanese male or female students (*Plate 8.1*).

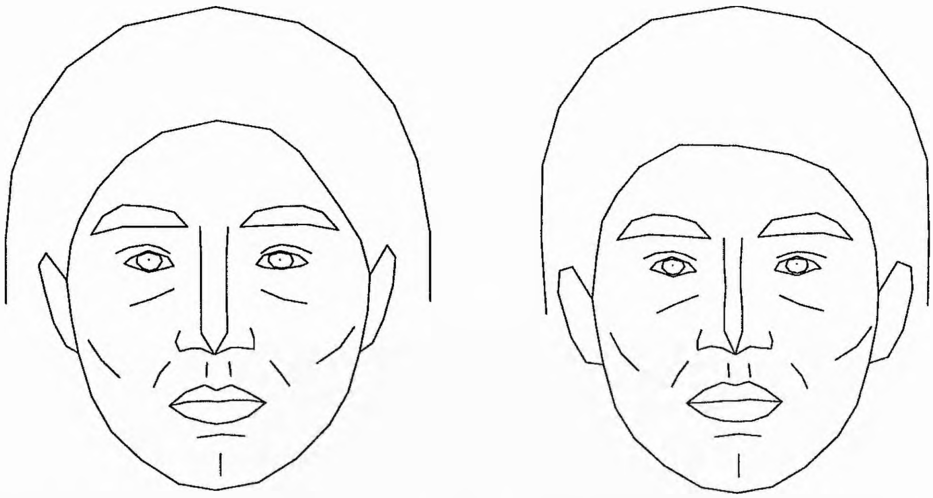


Figure 8.1. The feature-point information for the Japanese female (left) and male faces (right) was averaged in order to determine average shape information for each gender. The average male shape was then normalised to have the same inter-ocular distance as the female.

For each face gender, the face was feminised and masculinised in shape by 50% using the transform technique developed by Rowland and Perrett (1995) and Rowland *et al* (1997). Hence, for example, the Japanese female prototype was feminised by calculating the difference in shape between corresponding feature points for the Japanese female average and the Japanese male average and applying this 50% difference to the shape of the Japanese female average. The images of the constituent individuals where then warped to this feminised shape and blended to form the feminised Japanese female. Similarly, the Japanese female face was masculinised by subtracting this difference. The differences between the shape of the average female and average male faces are shown in *Figure 8.2*.

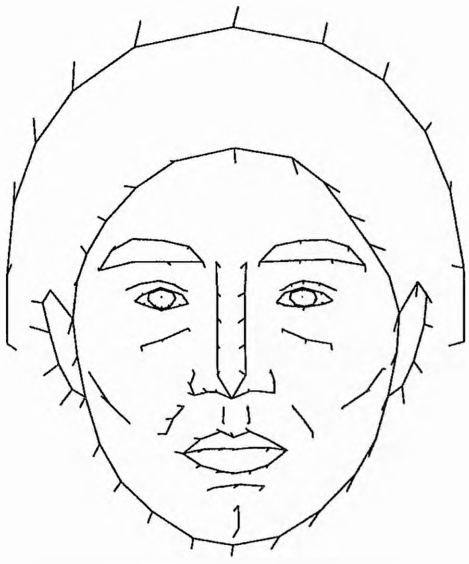


Figure 8.2. Differences between Japanese average female and male faces. This illustration represents the shape of a female face. The vectors describe the manner in which the face shape differs from the average male.

For each gender of face, the textures were kept constant such that faces were transformed in shape alone. The different blends were then printed using a dye-sublimation printer with four-colour ink roll (cyan, yellow, magenta and black).

These photographic-quality print-outs were then cropped manually so as to remove most of the ear area and the forehead. This was done so as to remove some of the distortion in the ear and hair area which is an artefact of the caricaturing or gender exaggerating/reducing process which might conceivably affect attractiveness judgements. A set of 35mm colour slides was produced using these final photographic images.

Design. Stimuli were presented on slides and shown in pairs. For each face gender, there were three different facial configurations: average, feminised and masculinised. Each subject recorded their own attractiveness preferences for each pair.

Procedure

Two pairings were produced for each gender of face, so for the female the following sets were produced: average female with masculinised female (*Plate 8.2*) and average female with feminised female (*Plate 8.3*); for male faces, the corresponding pairings were created: average male with masculinised male (*Plate 8.4*) and average male with feminised male (*Plate 8.5*). Each of the two images on these slides was labelled for identification in the judgement task to follow. Two sets of slides were created such that for each pair in one instance one member was on the left and in the other instance this same member appeared on the right. In both sets, the identifier given for a particular member in the pair remained the same. For instance, for the average female-feminised female pairing, in one case the female average would be displayed on the left with the feminised female on the right. In the alternative pairing, the feminised female would appear on the left and the average on the right.

This was performed in order to allow counterbalancing of stimulus presentation. Subjects were presented the Japanese face stimuli in random order. Slides were displayed with images as pairs. These were paired such that for each gender the following comparisons were made: female/male blend against feminised female/male blend and female/male blend against masculinised female/male blend. With each presentation subjects were asked to enter the code on the first page of a double-sided response sheet for the face within the pair they found to be the most attractive. Subjects were then instructed to turn their response sheet to the other side. The same image pairings were then repeated again in random order and subjects asked to indicate which of the pair they believed looked younger.

The experiment was performed within-culture with Japanese subjects and cross-culturally with white British subjects in Scotland and subjects from a variety of ethnic backgrounds in South Africa.

Results and analysis

The image selected as the more attractive by each subject from each pairing was determined and the data tabulated for all subjects. A similar tabulation was done separately for the age comparison judgement. This data was subjected to binomial comparison in order to determine whether one of the stimuli in each pair was consistently considered to be more attractive than the other. The same was done for the age comparison in order to determine if systematically altering the degree of gender typicality of an image influenced perceived age. Results are given separately for the female (attractiveness, *Table 8.1*; age, *Table 8.3*) and male stimuli (attractiveness, *Table 8.2*; age, *Table 8.4*).

SUBJECT		Female faces						
RACE	SEX	TOTAL	n	fem > ave	p	n	ave > masc	p
					<			<
Japanese	female	49	41	84%	0.00001	40	82%	0.00001
Japanese	male	151	117	77%	0.00001	120	79%	0.00001
	total	200	158	79%	0.00001	160	80%	0.00001
white British	female	65	50	77%	0.00001	56	86%	0.00001
white British	male	22	11	50%	n.s.	15	68%	n.s.
	total	87	61	70%	0.0005	71	82%	0.00001
South Africa								
Multirace	female	132	74	56%	n.s.	98	74%	0.00001
African	female	54	31	57%	n.s.	32	59%	n.s.
Asian	female	44	22	50%	n.s.	37	84%	0.00001
White	female	34	21	62%	n.s.	29	85%	0.00005
South Africa								
Multirace	male	49	26	53%	n.s.	32	65%	0.05
African	male	24	13	54%	n.s.	13	54%	n.s.
Asian	male	8	6	75%	n.s.	6	75%	n.s.
White	male	17	7	41%	n.s.	13	76%	0.05
South Africa								
Multirace	total	181	100	55%	n.s.	130	72%	0.00001
African	total	78	44	56%	n.s.	45	58%	n.s.
Asian	total	52	28	54%	n.s.	43	83%	0.00001
White	total	51	28	55%	n.s.	42	82%	0.00001
Pooled White								
White	female	99	71	72%	0.00005	85	86%	0.00001
White	male	39	18	46%	n.s.	28	72%	0.01
	total	138	89	64%	0.001	113	82%	0.00001
Total	female	245	165	67%	0.00001	194	79%	0.00001
	male	222	154	69%	0.00001	167	75%	0.00001
	total	467	319	68%	0.00001	361	77%	0.00001

Table 8.1. Attractiveness judgements for Japanese female stimuli. The table indicates the number of subjects which preferred the feminised female to the average and the number of subjects which found the average to be more attractive than the masculinised female.

SUBJECT		Male faces						
RACE	SEX	TOTAL	n	fem > ave	p	n	ave > masc	p
					<			<
Japanese	female	49	37	76%	0.0005	32	65%	0.05
Japanese	male	151	86	57%	n.s.	59	39%	0.01
	total	200	123	62%	0.005	91	46%	n.s.
white British	female	65	45	69%	0.005	29	45%	n.s.
white British	male	22	15	68%	n.s.	10	45%	n.s.
	total	87	60	69%	0.0005	39	45%	n.s.
South Africa								
Multirace	female	132	76	58%	n.s.	74	56%	n.s.
African	female	54	29	54%	n.s.	35	65%	0.05
Asian	female	44	27	61%	n.s.	21	47%	n.s.
White	female	34	20	59%	n.s.	18	53%	n.s.
South Africa								
Multirace	male	49	19	39%	n.s.	21	43%	n.s.
African	male	24	10	42%	n.s.	7	29%	n.s.
Asian	male	8	3	38%	n.s.	3	38%	n.s.
White	male	17	6	35%	n.s.	11	65%	n.s.
South Africa								
Multirace	total	181	95	52%	n.s.	95	52%	n.s.
African	total	78	39	50%	n.s.	42	54%	n.s.
Asian	total	52	30	58%	n.s.	24	46%	n.s.
White	total	51	26	51%	n.s.	29	57%	n.s.
Pooled White								
White	female	99	65	66%	0.05	47	47%	n.s.
White	male	39	18	46%	n.s.	18	46%	n.s.
	total	138	83	60%	0.05	65	47%	n.s.
Total	female	245	158	64%	0.00001	135	55%	n.s.
	male	222	120	54%	n.s.	90	41%	0.005
	total	467	278	60%	0.00005	225	48%	n.s.

Table 8.2. Attractiveness judgements for Japanese male stimuli - n.s. signifies that the results were not significant at $p < 0.05$. The table indicates the number of subjects which preferred the feminised male to the average and the number of subjects which found the average to be more attractive than the masculinised male.

SUBJECT		Female faces						
RACE	SEX	TOTAL	n	fem > ave	p	n	ave > masc	p
					<			<
Japanese	female	25	19	76%	0.01	17	68%	n.s.
Japanese	male	128	98	76%	0.00001	95	73%	0.00001
	total	153	117	76%	0.00001	112	73%	0.00001
white British	female	38	25	66%	n.s.	33	87%	0.00001
white British	male	12	6	50%	n.s.	9	75%	n.s.
	total	50	31	62%	n.s.	42	84%	0.00001
South Africa								
Multirace	female	132	63	47%	n.s.	84	64%	0.005
African	female	54	25	46%	n.s.	32	59%	n.s.
Asian	female	44	22	50%	n.s.	26	59%	n.s.
White	female	34	16	47%	n.s.	26	76%	0.005
South Africa								
Multirace	male	49	28	57%	n.s.	27	55%	n.s.
African	male	24	14	58%	n.s.	10	42%	n.s.
Asian	male	8	5	63%	n.s.	7	88%	0.05
White	male	17	9	53%	n.s.	10	59%	n.s.
South Africa								
Multirace	total	181	91	50%	n.s.	111	61%	0.005
African	total	78	39	50%	n.s.	42	54%	n.s.
Asian	total	52	27	52%	n.s.	33	63%	n.s.
White	total	51	25	49%	n.s.	36	71%	0.005
Pooled White								
White	female	72	41	57%	n.s.	59	82%	0.00001
White	male	29	15	52%	n.s.	19	66%	n.s.
	total	101	56	55%	n.s.	78	77%	0.00001
Total	female	195	107	55%	n.s.	134	69%	0.00001
	male	189	132	70%	0.00001	131	69%	0.00001
	total	384	239	62%	0.00001	265	69%	0.00001

Table 8.3. Relative perceived age of female Japanese stimuli. The table indicates the number of subjects which believed the feminised female looked younger than the average and the number of subjects which found the average to look younger than the masculinised female.

SUBJECT		Male faces						
RACE	SEX	TOTAL	n	fem > ave	p		ave > masc	p
					<			<
Japanese	female	25	15	60%	n.s.	7	28%	0.05
Japanese	male	128	61	48%	n.s.	31	24%	0.00001
	total	153	76	50%	n.s.	38	25%	0.00001
white British	female	38	35	92%	0.00001	33	87%	0.00001
white British	male	12	10	83%	0.05	8	67%	n.s.
	total	50	45	90%	0.00001	41	82%	0.00001
South Africa								
Multirace	female	132	64	48%	n.s.	88	67%	0.0005
African	female	54	23	43%	n.s.	35	65%	0.05
Asian	female	44	22	50%	n.s.	28	64%	0.01
White	female	34	19	56%	n.s.	25	74%	0.01
South Africa								
Multirace	male	49	31	63%	n.s.	32	65%	0.05
African	male	24	15	63%	n.s.	16	67%	n.s.
Asian	male	8	4	50%	n.s.	4	50%	n.s.
White	male	17	12	71%	n.s.	12	71%	n.s.
South Africa								
Multirace	total	181	95	52%	n.s.	120	66%	0.00001
African	total	78	38	49%	n.s.	51	65%	0.01
Asian	total	52	26	50%	n.s.	32	62%	n.s.
White	total	51	31	61%	n.s.	37	73%	0.005
Pooled White								
White	female	72	54	75%	0.00005	58	81%	0.00001
White	male	29	22	76%	0.01	20	69%	0.05
	total	101	76	75%	0.00001	78	77%	0.00001
Total	female	195	114	58%	0.05	128	66%	0.00001
	male	189	102	54%	n.s.	71	38%	0.001
	total	384	216	56%	0.05	199	52%	n.s.

Table 8.4. Relative perceived age of male Japanese stimuli. Here subjects were asked which of the stimuli they believed to be younger. The table indicates the number of subjects that believed the feminised male looked younger than the average and the number of subjects that found the average to look younger than the masculinised male.

Effect of Gender Manipulation on Attractiveness

Female Stimuli

Attractiveness judgements for the female stimuli were also analysed. Across the subject pool there was a very strong preference for the more feminine of the test stimuli and this was apparent for both male and female subjects. Representative of this, the feminised female was rated as more attractive than average by Japanese subjects. The average female was preferred to the masculinised version by Japanese subjects. These findings were found to be consistent across gender of observer.

The feminised female was rated as more attractive than average by white British subjects. The average female was preferred to the masculinised version by such subjects. Again, this was only found to be significant with female subjects.

The feminised female shape was not significantly preferred to the average female by viewers in South Africa while the average female shape was preferred to the masculinised female.

A survey of the results for South African subjects based on ethnic background indicates an impact of culture on attractiveness preferences. Most of the racial subgroups gave the same preferences as the group as a whole. However, subjects with a native African ethnic background parted with the group as a whole by finding no preference between the average male or the masculinised male. It should be noted that due to the low numbers of subjects in these racial subgroups any speculation should be withheld until further study.

In summary, the more feminine of the stimuli were selected most often in both comparisons by the pooled White viewers: feminised faces were chosen over the average and average faces were preferred to the masculinised shape.

Male Stimuli

Over the subject pool, the femininised male was consistently preferred to the average but this was limited to female subjects. The masculinised shape was also preferred to the average by male subjects while female subjects indicated no consistent preference. The feminised male was consistently preferred to the average by Japanese subjects with this limited again to cross-sex judgements. The average male was found to be less attractive than the maculinised male by Japanese male subjects and it is this group which appears to account for this finding across male subjects on the whole. Females indicated a preference for the average shape over the masculinised male.

The feminised male was also consistently preferred to the average by white British subjects. Again this preference was only significant for female viewers. No preference was found for the average over the masculinised male.

There was no preference for the male when feminised in shape over the average or for the average shape over the masculinised image by viewers in South Africa. Female native African subjects were the only group within the South African study population to indicate a preference, indicating the average faces as preferable to those which had been masculinised.

Pooling the White subjects between those viewing the stimuli in Scotland and those viewing the stimuli in South Africa, the feminised male was preferred to the average. This finding only applied to cross-sex judgements such that females consistently selected the feminised male while males were inconsistent in this choice. There was no preference for the average male over the masculinised shape by either sex of subject. Due to the unmatched number of female and male subjects viewing the stimuli in each race group and thus the overall difference in numbers between race of viewer it is not possible to make a valid comparison of judgements between race or to look at any interaction between race and sex of subject. Indeed the majority of male subjects were present in the Japanese sample. From the data available the feminised male appears to be preferred, particularly within culture and in cross-sex judgements.

Effect of Gender Manipulation on Apparent Age

The majority of viewers perceived the feminised female to be younger than the average configuration (*Table 8.3*) but this finding was limited to males. This finding seems to be accountable to the preferences of Japanese subjects who had the largest male subject population. No other group showed any consistent discrimination for this comparison. Japanese males and white British females found the average shape appeared younger than the masculinised version.

From Table 8.4 it can be noted that subjects made a consistent finding that the feminised Japanese male shape appeared younger than the average but this was limited to female subjects. The majority of female subjects across the subject pool also noted that the average male shape looked younger than the masculinised shape while males found the contrary to be the case. There appears to be some weak observable relationship between increased masculinity and apparent ageing detectable by female subjects but, with the exception of Japanese female subjects (within-culture), this does not appear to have any association with attractiveness judgements.

Again, in the feminised vs. average comparison, with the exception of within-culture judgements by Japanese viewers there was no relationship between attractiveness and apparent age judgement. There was a weak relationship between masculinising female faces and apparent ageing with white British female viewers finding the masculinised Japanese female face to appear both older and less attractive than the average.

Across comparisons with Japanese face stimuli female subjects appear to make a stronger association between the shape changes associated with increased masculinity and the ageing process.

Overall Discussion

The feminised shape male and female faces were found to be more attractive than the average configuration. This was further qualified by sex and cultural background of rater and sex of face. The average female was also preferred to the masculinised female while there was poor agreement in the choice of average male versus masculinised male faces.

With female Japanese stimuli, the feminised shape was preferred to the average and the average to the masculinised shape by both male and female Japanese subjects. This was also true for white British subjects, but only with female viewers; in Africa both sexes of White subjects preferred the average shape to the masculinised without showing any preference for the feminised over the average. In South Africa, with the exception of native African subjects, the average female was preferred to the masculine shape while there was no preference for the feminised shape over the average configuration. Outwith white British subjects this preference was limited to female observers.

The feminised male Japanese shape was consistently found to be more attractive than the average in cross-sex judgements alone. This was found with both Japanese and white British subjects, but not with subjects in South Africa. There, African females preferred the average male to the masculine while males preferred the more masculine shape to the average. Japanese males also found a more masculine shape to be preferred to the average male in same sex judgements.

These findings indicate that there is some influence by both the immediate environment and general surrounding environment on attractiveness judgements and that attractiveness preferences are also dependent on gender of observer. Here, immediate environment refers to the family and peers with which one is likely to have the greatest experience particularly in early life (e.g. Black Africans with Black Africans in their family and neighbourhood in South Africa) while surrounding environment refers to the more diverse variety of people, and therefore facial-types, one is likely to experience in later life (e.g. the variety of ethnic groups found in South Africa). Females generally tend to have a greater preference for a more feminine configuration in a forced-choice paradigm in both cross-sex and same-sex judgements while males are less consistent in their judgements This was particularly true for cross-cultural judgements with white British observers.

From the tables it can also be seen that there is little consistency in the relationship between apparent youth and judged attractiveness. Here, for instance, subjects who choose more feminine female stimuli as more attractive often do so despite not finding the stimulus to look either younger or older than its rival.

Part Two - White British stimuli

Introduction. The technique employed in Part One was replicated using white British stimuli. This was done in order to determine whether the configuration found to be attractive in Japanese stimuli would agree with that found to be attractive in white British faces. Cross-cultural judgements were made again in order to determine if judgements were consistent regardless of cultural background. In Part One there was a consistent preference, both within and across culture, for the feminised female and male face over the average face by female viewers. In the masculinised vs. average shape comparison there were few instances in which a significant preference was found between the male stimuli and these were not consistent with preferences being made in either direction. In contrast, with female stimuli the average face was consistently preferred to the masculinised shape. For this reason only the feminised vs. average shape presentation was performed for male stimuli in Experiment 9.

Method

Subjects. 38 white British visitors to an open day at the University of St Andrews, 28 females, aged 15 to 49 (median 16) and 10 males, aged 15 to 49 (median 17) took part. These subjects also took part in Part One.

153 Japanese undergraduates at the Otemon Gakuin University, 128 males, aged 18 to 27 (median 19) and 25 females, aged 18 to 24 (median 19). All these subjects were also involved in Part One.

67 undergraduates at the University of Natal, 45 female, aged 19 to 41 (median 21) and 22 male, aged 20 to 35 (median 23). These were divided into three groups based on ethnic background; 24 Africans, 13 female, aged 19 to 39 (median 21) and 11 male, aged 23 to 35 (median 24); 32 Whites, 21 female, aged 20 to 41 (median 21) and 11 male, aged 20 to 23 (median 21); 11 Asians (mostly of Indian extraction), all female, aged 20 to 22 (median 20). All these subjects were also involved in Part One.

Stimuli. Photorealistic average images generated from white British faces were created as for Japanese faces. Composites were created to represent prototypical or average males and females. The shape of the white British blends were produced from 25 male (aged 19 to 23 years, mean 21.0) and 51 female (aged 19 to 26, mean 20.6) white British individuals studying at the University of St Andrews - the average texture for the white British female was derived from a subset of 30 female faces (aged 19 to 22, mean 20.6) and the white British male from a subset of 12 faces (aged 19 to 23, mean 21.1), all of which had been photographed in standard lighting conditions. All the faces were photographed using Fuji Chrome film with ASA 100 and were taken in portrait format with a flash to the left and right side of the face in order to minimise shading. The photographs were stored as images on Kodak Photo-CD and delineated at the fullest size possible - 2048 pixels x 3072 pixels - by one operator to ensure both accuracy and consistency across the faces. The photorealistic blend for the female and male were produced in the same manner as for the Japanese stimuli. Both this set of images and the Japanese faces used in Experiment 8 were delineated in the same manner in order to allow consistency between the two sets. Before blending the average female shape was normalised to have the same inter-ocular distance as the average male (*Figure 8.3*). The final composites incorporated the average shape and texture information for a white British female and male student (*Plate 8.6*).

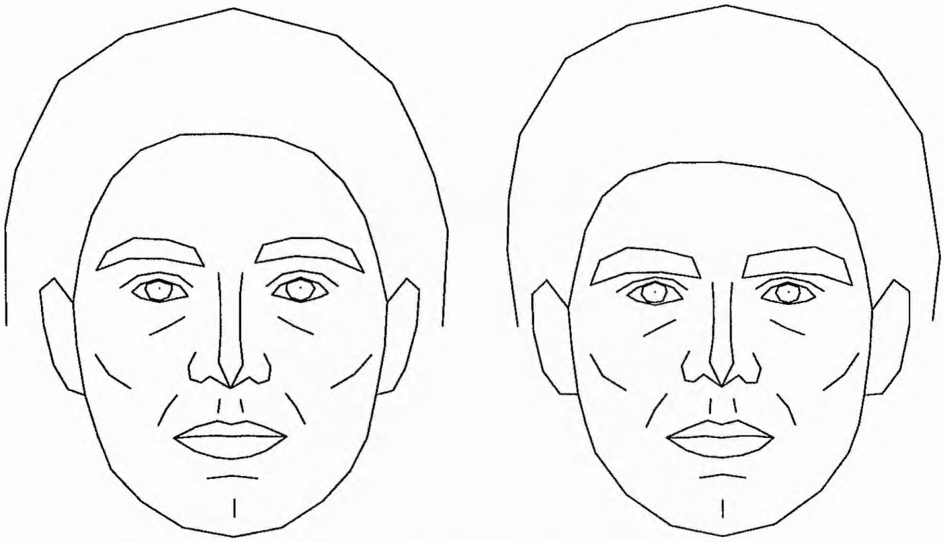


Figure 8.3. The feature-point information for the white British female (left) and male (right) faces was averaged in order to determine average shape information for each gender. The average female shape was then normalised to have the same inter-ocular distance as the male.

As with the Japanese stimuli, the female composite face was feminised and masculinised in shape by 50%. The white British female prototype was feminised by calculating the difference in shape between corresponding feature points for the white British female average and the white British male average and applying this 50% difference to the shape of the white British female average. The images of the constituent individuals where then warped to this feminised shape and blended to form the feminised white British female. Similarly, the white British female face could be masculinised by subtracting this difference. The differences between the shape of the average female and average male faces are shown in *Figure 8.4*.

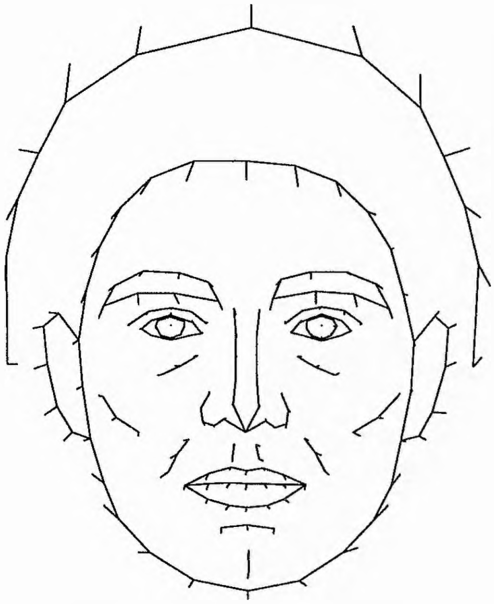


Figure 8.4. Differences between white British average female and male faces. This illustration represents the shape of a female face. The vectors describe the manner in which the face shape differs from the average male.

Hence the textures were kept constant such that faces were transformed in shape alone. The white British male was also feminised in a manner analogous to that for the female. The different blends were then printed using a dye-sublimation printer. These photographic-quality print-outs were then cropped to remove most of the ear area and the hair to the forehead. A set of 35mm colour slides was produced using these final photographic images in the same manner as with the Japanese stimuli in Part One.

Design. Stimuli were presented on slides and shown in pairs. For female face stimuli, there were three different facial configurations: average, feminised and masculinised. For male face stimuli, subjects made attractiveness judgements between two stimuli (one judgement), feminised or average. Subjects attractiveness preferences for each pair were recorded.

Procedure

Two pairings were produced for the female: average female with feminised female (*Plate 8.7*) and average female with masculinised female (*Plate 8.8*) as in Part One. One was produced for the male: average male with feminised male (*Plate 8.9*). Subjects were presented the white British face stimuli in random order. Slides were displayed with images as pairs. These were paired such that the following comparisons were made: female/male blend against feminised female/male blend and female blend against masculinised blend. With each presentation subjects were asked to write the code for the face within the pair they found to be the most attractive. Subjects in Studies 2 and 3 were then presented with stimuli again in random order and asked to indicate which of the stimuli they judged to be the youngest.

Results and analysis

The image selected, by each subject, from each pairing as the more attractive by each subject from each pairing was determined and the data tabulated for all subjects. This data was subjected to binomial comparison in order to determine whether one of the stimuli in each pair was consistently considered to be more attractive than the other. Results for perceived attractiveness and apparent youth are given separately for the female stimuli (Table 8.5; Table 8.7) and male stimuli (Table 8.6; Table 8.8).

SUBJECT		Female faces						
RACE	SEX	TOTAL	n	fem > ave	p	n	ave > masc	p
					<			<
Japanese	female	25	9	36%	n.s.	16	64%	n.s.
Japanese	male	128	35	27%	0.00001	63	49%	n.s.
	total	153	44	29%	0.00001	79	52%	n.s.
white British	female	28	21	75%	0.01	24	86%	0.0005
white British	male	10	5	50%	n.s.	5	50%	n.s.
	total	38	26	68%	0.05	29	76%	0.005
South Africa								
Multirace	female	45	33	73%	0.005	33	73%	0.005
African	female	13	10	77%	n.s.	10	77%	n.s.
Asian	female	11	7	64%	n.s.	7	64%	n.s.
White	female	21	16	76%	0.05	16	76%	0.05
South Africa								
Multirace	male	22	17	77%	0.05	16	73%	0.05
African	male	11	7	64%	n.s.	6	55%	n.s.
White	male	11	10	91%	0.01	10	91%	0.01
South Africa								
Multirace	total	67	50	75%	0.0001	49	73%	0.0005
African	total	24	17	71%	0.05	16	67%	n.s.
Asian	total	11	7	64%	n.s.	7	64%	n.s.
White	total	32	26	81%	0.0005	26	81%	0.0005
Pooled White								
White	female	49	37	76%	0.0005	40	82%	0.00005
White	male	21	15	71%	0.05	15	71%	0.05
	total	70	52	74%	0.0005	55	79%	0.00001
Total	female	98	63	64%	0.005	73	75%	0.00001
	male	160	57	36%	0.0005	84	53%	n.s.
	total	258	120	47%	n.s.	157	61%	0.0005

Table 8.5. Attractiveness preferences for white British female stimuli. This table indicates the number of subjects that preferred the feminised female stimulus to the average shape and those who preferred the average female stimulus to the masculinised female shape.

SUBJECT		Male faces			
RACE	SEX	TOTAL	n	fem > ave	p
					<
Japanese	female	25	18	72%	0.05
Japanese	male	128	92	72%	0.00001
	total	153	110	72%	0.00001
White British	female	28	15	54%	n.s.
white British	male	10	4	40%	n.s.
	total	38	19	50%	n.s.
South Africa					
Multirace	female	45	27	60%	n.s.
African	female	13	4	31%	n.s.
Asian	female	11	8	73%	n.s.
White	female	21	15	71%	0.05
South Africa					
Multirace	male	22	9	41%	n.s.
African	male	11	4	36%	n.s.
White	male	11	5	45%	n.s.
South Africa					
Multirace	total	67	36	54%	n.s.
African	total	24	8	33%	n.s.
Asian	total	11	8	73%	n.s.
White	total	32	20	63%	n.s.
Pooled White					
White	female	49	30	61%	n.s.
White	male	21	9	43%	n.s.
	total	70	39	56%	n.s.
Total	female	98	60	61%	0.05
	male	160	105	66%	0.0001
	total	258	165	64%	0.00001

Table 8.6. Attractiveness preferences for white British male stimuli. This table indicates the number of subjects that preferred the feminised male stimulus to the average male shape.

SUBJECT		Female faces						
RACE	SEX	TOTAL	n	fem > ave	p	n	ave > masc	p
					<			<
Japanese	female	25	5	20%	n.s.	15	60%	n.s.
Japanese	male	128	30	23%	0.00001	56	44%	n.s.
	total	153	35	23%	0.00001	71	46%	n.s.
South Africa								
Multirace	female	45	33	73%	0.005	35	78%	0.0005
African	female	13	10	77%	n.s.	10	77%	n.s.
Asian	female	11	8	73%	n.s.	8	64%	n.s.
White	female	21	15	71%	0.05	17	81%	0.005
South Africa								
Multirace	male	22	13	59%	n.s.	13	59%	n.s.
African	male	11	6	55%	n.s.	7	64%	n.s.
White	male	11	7	64%	n.s.	6	55%	n.s.
South Africa								
Multirace	total	67	46	69%	0.005	48	72%	0.0005
African	total	24	16	67%	n.s.	17	71%	0.05
Asian	total	11	8	73%	n.s.	8	73%	n.s.
White	total	32	22	69%	0.05	23	72%	0.05
Total	female	70	38	54%	n.s.	50	71%	0.0005
	male	150	43	29%	0.00001	69	46%	n.s.
	total	220	81	37%	0.0001	119	54%	n.s.

Table 8.7. Relative perceived age of white British female stimuli. This table indicates the number of subjects that found the feminised female stimulus to look younger than the average shape and those who found the average female stimulus to look younger than the masculinised female shape.

SUBJECT		Male faces			
RACE	SEX	TOTAL	n	fem > ave	p
					<
Japanese	female	25	20	80%	0.005
Japanese	male	128	103	80%	0.00001
	total	153	123	80%	0.00001
South Africa					
Multirace	female	45	7	16%	0.00001
African	female	13	4	31%	n.s.
Asian	female	11	1	9%	0.01
White	female	21	2	10%	0.0005
South Africa					
Multirace	male	22	4	10%	0.005
African	male	11	3	27%	n.s.
White	male	11	1	9%	0.01
South Africa					
Multirace	total	67	11	16%	0.00001
African	total	24	7	29%	0.05
Asian	total	11	1	9%	0.01
White	total	32	3	9%	0.00001
Total	female	70	27	39%	n.s.
	male	150	107	71%	0.00001
	total	220	134	61%	0.005

Table 8.8. Relative perceived age of white British male stimuli. This table indicates the number of subjects that found the feminised male stimulus to appear younger than the average male shape.

Effect of Gender Manipulation on Apparent Attractiveness

Female Stimuli

Overall, the feminised female was selected as more attractive than the average by female subjects while the opposite was found by male subjects. Female subjects also preferred the average to the masculinised shape. The feminised female was rated as more attractive than the average by white British subjects. The average female was also significantly preferred to the masculinised version. This was only found for female subjects with males divided over which faces they found to be the most attractive.

The average female was rated as more attractive than the feminised version by male Japanese subjects alone accounting for this preference in the overall male population. There was no real difference in preference in any other judgement.

With South African viewers, the more feminine of each shape was chosen most often as the more attractive in both the comparisons. This was found with female and male subjects. Analysis within the South African group is spurious given the low and inconsistent number of subjects within each racial subgroup. Race and sex comparison could not be made because of low and inconsistent numbers of individuals between subject samples. Indeed Japanese males accounted for 80% of the overall male population and thus had a tendency to bias any results.

Male Stimuli

The feminised male face was consistently preferred to the average across the subject pool. White British subjects showed no preferences between the male stimuli. Both male and female Japanese subjects found the feminised male shape to be more attractive than the average. As with the white British Subjects, the South African viewers showed no general preferences for male stimuli. Within the South African viewing group, female White subjects preferred the feminised male shape to the average while males showed no preference. African and Asian subjects showed the same lack of preference between stimuli as the general South African test population.

When the results for White subjects were pooled, no preferences for male stimuli with white British viewers were found. Race and sex of subject could not be compared due to low and inconsistent subject numbers between each subject sample.

Effect of Gender Manipulation on Apparent Age

The feminised white British female face was seen to be younger than the average by Japanese males and Whites tested in South Africa. White and native African subjects tested in South Africa also found the average face appeared younger than the masculinised shape. There was a stronger but not completely consistent relationship here such that Japanese males and Whites in South Africa found the feminised female looked younger and was more attractive and white British subjects found the average face appeared both younger and more attractive than the masculinised configuration.

The feminised white British male face was seen to be younger than the average by the majority of subjects tested across culture (*Table 8.8*) in contrast with the finding that only Japanese subjects and White females tested in South Africa consistently identified the more feminine face to be more attractive.

Overall Discussion

There was poor consistency in attractiveness preferences between the average and feminised face shape. The feminised shaped male face was found to be more attractive than the average configuration. The average female face was found to be more attractive than the masculinised shape. These findings were further qualified by sex and cultural background of rater and sex of face.

The average white British female face was preferred to the feminised shape by Japanese subjects while white British subjects showed a preference for the feminised shape. South African subjects also reflected this white British opinion primarily in the White subgroup. The average female was also preferred to the masculinised female but this was mainly isolated to judgements by female viewers with males rarely showing a strong preference.

Attractiveness preferences between feminised and average shape male stimuli were limited to Japanese observers and female White subjects in an environment where individuals of another ethnic background predominated. With white British stimuli there was no clear link between apparent age and attractiveness. Only Japanese and White South African subjects gave a significant attractiveness preference for male stimuli, selecting the feminised face; while the Japanese group rated this face as younger than the average shape all other groups rated the face as appearing significantly older. There was a stronger, but not entirely consistent, relationship between attractiveness and apparent youth for white British female faces.

General Discussion

Male Japanese viewers also showed a preferences for the average female to the feminised female shape but not for the average over the masculinised female. In within-culture judgements, white British females also considered both the feminised female to be more attractive than the average and the average to be more attractive than the masculinised configuration. White males also showed this within-culture preference very strongly as a sub-group of the South African viewers.

In summary, the findings suggest an overall preferences for males and females with a more feminine facial shape configuration. The relationship between attractiveness and apparent youth using these stimuli seems to be limited but there is a possibility that the changes contingent with femininity may also correlate with those found in a more youthful face. Male subjects appear to be less consistent in their judgements than females with white British males displaying a general indifference and Japanese males showing a preference for the average female to the feminised version in cross-race judgements. Having a more balanced number of male and female subjects and a more balanced number of subjects within the different viewing cultures would allow for a more sensitive analysis of these differences.

One might predict from the pattern of results that using an approach where people were able to select the face they found most attractive interactively using a technique similar to that used in Experiments 6 and 7 (*Chapter 6*) would result in individuals opting for an image which was consistently different from average. The pattern of results found here suggest that stimuli more feminine in shape to some degree would be found to be the most attractive with Japanese and white British stimuli and both male and female sexes. The findings also indicate, somewhat unsurprisingly, that the preferred level of feminisation would be smaller for male faces than female faces. It would also appear that judgements would be similar between Japanese and white British subjects with smaller deviations from the average to be found by African subjects. It may also be hypothesised from these findings that Japanese subjects would select female faces with a greater degree of feminisation within their own race than in cross-race judgements and vice-versa. Japanese male subjects consistently preferred feminised females with Japanese stimuli while showing the opposite preference with White stimuli. Conversely, pooled White male subjects consistently found the feminised white British female face shape to be more attractive than the average while indicating no preference with Japanese female stimuli. This may be because individual perceptual histories are such that white British viewers and Japanese viewers use slightly different configural cues when judging attractiveness of faces. One would therefore hypothesise that Whites and Japanese subjects would select own-race faces which had been feminised more relative to the average than other-race faces using such an interactive technique.

Experiment 9 - Interactive Manipulation

The findings in Experiment 8 were produced by using a forced-choice paradigm. The outcome of these experiments demonstrated that viewers generally find facial configurations that are more feminine to be more attractive with both male and female Japanese faces and with female white British faces. However, using this paradigm there were some instances were average faces were seen to be more attractive than their feminised versions and occasions when it was not possible to detect a consistent preference. This may be because the average face was indeed found to be in a more attractive configuration than the feminised version. However, it is perhaps more likely that the use of stimuli with a 50% gender caricature was too extreme to provide an accurate measure of the preferred facial configuration. In order to cater for this possibility a more sensitive procedure was designed to employ the interactive techniques introduced in *Chapter 6*. This allows the viewer to manipulate the images presented online, effectively altering the degree of feminisation or masculinisation of the face in order to determine the point at which he or she finds a particular face image to be the most attractive. In this manner it is possible to identify the precise degree of deviation from the norm which is found by viewers, on the whole, to display the most attractive configuration.

Further, in Experiment 8 sample sizes were biased such that the numbers of subjects within each sex or race group were not equated, making between-race and between-gender analysis impractical. Here a similar number of Japanese and white British subjects viewed the stimuli with numbers of males and female viewers balanced within each race group.

The hypothesis that faces are recognised by comparison to some culturally specific norm which is developed for use in the processing of faces and dependent on an individual's perceptual history (*Chapter 1*) might suggest that the degree of femininity/masculinity chosen as more attractive would vary between other- and same-race viewed stimuli and that as a result an individual would select faces as more attractive which showed a greater degree of change in some direction from the average of their own race than the average of the other race.

Method

Apparatus. A Silicon Graphics R4400 Indigo maximum impact computer with 4MB TRAM texture memory was used for stimulus creation and presentation in 24-bit and real-time with monitor set at medium contrast (factory setting).

Subjects. 50 white British undergraduates and postgraduates at the University of St Andrews, Scotland: 25 females, ages 19 to 30 (median 20) and 25 males, ages 19 to 31 (median 21); 42 Japanese subjects recruited from Doshisha University and tested at ATR, Japan: 19 females and 23 males.

Stimuli. Photorealistic average images of faces were created for both Japanese and white British faces in the manner described for Experiment 8. Within each race, blends were created to represent prototypical or average males and females (Japanese, *Plate 8.1*; white British, *Plate 8.6*). The Japanese blends were produced from 28 male and 28 female Japanese individuals studying at Otemon-Gakuin University and taken in standard lighting conditions. Each of these pictures was cropped so as remove the background, hair and ears. This technique was similar to that performed on the stimuli for the experiments described in *Chapter 2* through *Chapter 6*. The background and external features was replaced with a black background and the border between the facial outline and the new background was convolved, or blurred over a gradient, so as to not have an abrupt disjoint between the face and the black background. This procedure was carried out in order to prevent distortion of the background and blurring of the hair, ears and neckline - an artefact of the blending and gender-manipulation process (see *Plates 8.1* and *8.6*) - from affecting subjects' attractiveness judgements. Thus the end results produced cropped images of male and female Japanese blends (*Plate 8.10*). The shape of the white British blends were produced from 25 male and 51 female white British individuals studying at the University of St Andrews - the average texture for the white British individuals was derived from a subset of 12 male and 30 female faces which had been taken in standard lighting conditions. Lighting conditions and film used differed between the white British and Japanese groups. Each of these white British images was cropped so as remove the background, hair and ears. Thus the end results produced cropped images of male and female white British blends (*Plate 8.11*).

For each face gender, the resultant cropped face composite was feminised and masculinised in shape by 25% and 50%. For example, the Japanese female prototype was feminised by calculating the difference in shape between corresponding feature points for the Japanese female average and the Japanese male average and applying this 25% or 50% difference to the shape of the Japanese female average. Similarly, the Japanese female face was masculinised by subtracting this difference. For each gender of face, the textures were kept constant such that faces were transformed in shape alone. An interactive morph was created utilising the Silicon Graphics hardware. This used the resultant five images for each face gender and race (average, feminised by 50%, feminised by 25%, masculinised by 25% and masculinised by 50%) to enable movement between a feminised version and a masculinised version that could be made in real-time. In order to do this the original images were first stretched to be 512 x 512 pixels size. Feature point data were also stretched by the same factor to maintain registration. The Delaunay refinement algorithm (Ruppert, 1995) creates a triangular tessellation from the feature point data which projects the texture into the desired shape. To do this, the original (unstretched) version of the feature point data is used to map the texture map triangles on screen in the correct dimensions. The amount of texture which is present in the displayed image is controlled using an alpha component. This specifies the transparency, how the values of two overlaid pixels are combined, between two images in the continuum. All five images were used in order to ensure smooth continuity across the sequence using four discrete morphs. The five boundaries for this sequence are illustrated for the female (*Plate 8.12*) and male white British images (*Plate 8.13*) and for the female (*Plate 8.14*) and male Japanese images (*Plate 8.15*).

Design. For each face gender, there were two independent variables being manipulated in this experiment: race of face and race of viewer. The faces being displayed were formed from blending either Japanese or white British faces. These faces were viewed either by white British subjects in St Andrews or Japanese subjects from Kyoto. The dependent variable was the precise degree of feminisation/masculination perceived by an individual subject to be the most attractive (varied across an infinite continuum from a 50% feminised blend to a 50% masculinised blend).

Procedure

Stimuli were presented in the centre of an 800 x 800 pixel window on a Silicon Graphics computer in 24 bit colour. Both Japanese and white British subjects were instructed to move the mouse from left to right until they found the point at which they believed the face presented to be most attractive. This process was done twice for each type of face: Japanese male and female, white British male and female. In each of these trial pairs, the faces were reversed so that the end-points were at opposing ends. Hence, in one instance the feminised Japanese female would be to the right end of the scale with masculinised Japanese female on the left and in the other the feminised version would appear at the left end of the scale with masculinised version on the right. In all instances, the average (average) face would appear at the midpoint of the scale.

Results and analysis

Two three-way ANOVAs were performed on the data - the female and male stimuli being analysed separately with subject gender (two levels) and subject race (two levels) as between-subject factors and race of stimulus face as a within-subject factor (two levels: Japanese male and white British male; Japanese female and white British female, respectively).

Female stimuli

For female stimuli there was no main effect of subject gender ($F_{1,88}$=1.58, p=0.21, mse=0.14), subject race ($F_{1,88}$=0.32, p=0.57, mse=0.003) or race of stimulus face ($F_{1,88}$=1.42, p=0.24, mse=.005). An interaction was found between race of subject and race of stimulus face ($F_{1,88}$=17.06, p<0.001, mse=0.54). No other interactions were found. Further analysis was carried out to interpret the interaction between race of subject and race of face. The data for both male and female subjects were combined in this analysis.

White British subjects

All results here were tested against a hypothesised mean of 0% change in masculinity/femininity, that is masculinising or feminising the face shape would have no effect on perceived attractiveness and subjects would prefer the average shape. The most attractive faces shape for Japanese faces was on average feminised by 10% which is significantly different from average (p=0.027, t=2.3, df=49) as was the white British face which, on average, was feminised by 24% (p<0.001, t=7.6, df=49). The distribution of faces selected as most attractive is illustrated in *Figure 8.5.*

Japanese Subjects

The pattern of results with Japanese subjects was generally found to be similar for white British subjects. A feminine Japanese face was selected with an average degree of feminisation of 23% (p<0.001, t=7.6, df=41) while, on average, a white British face which had been feminised by 15% was selected as the most attractive (p<0.001, t=4.5, df=41; *Figure 8.5*).

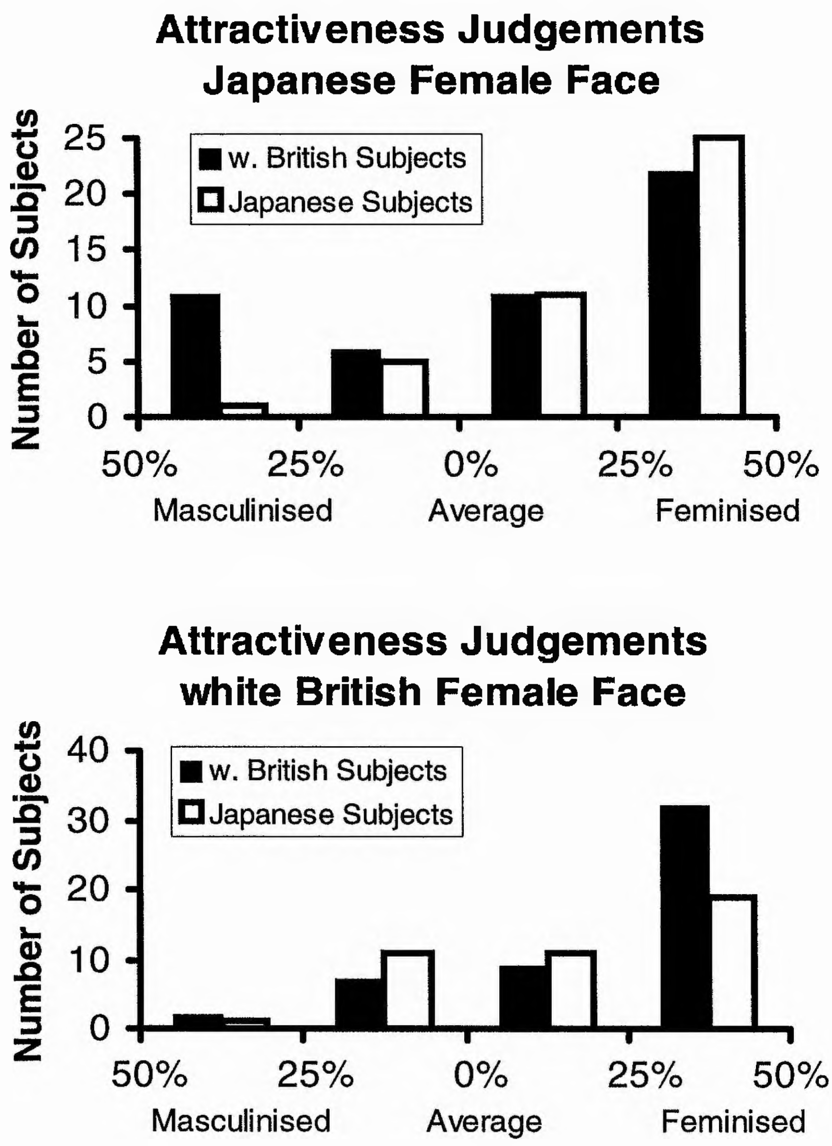


Figure 8.5. Attractiveness judgements for female faces. Japanese and white British subjects both showed a tendency to select more face shapes which were more feminine than average for Japanese and white British female faces.

Cross-race comparison

Despite the similar pattern in results for female stimuli between white British and Japanese subjects it would appear that, although feminisation of females for attractiveness occurs with other-race faces, a markedly more feminine face shape is preferred to an average in same-culture attractiveness judgements (*Figure 8.4* and *Figure 8.6*). The summary of the results for each race of viewer showing this interaction is described in *Figure 8.6*.

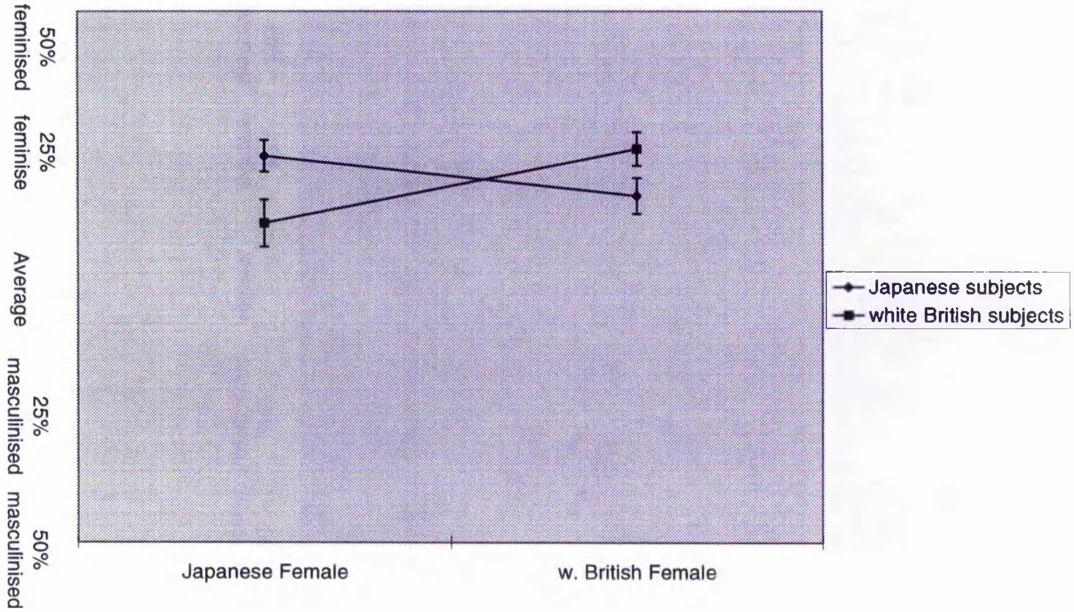


Figure 8.6. A summary of the results across race of stimulus face and race of viewer. The pattern of results was similar for Japanese and white British female faces with a more feminised face shape consistently selected as more attractive. An interaction was found between race of subject and race of faces such that with female stimuli a greater degree of feminisation was produced within-race when viewers were selecting the most attractive face.

<u>Male Stimuli</u>

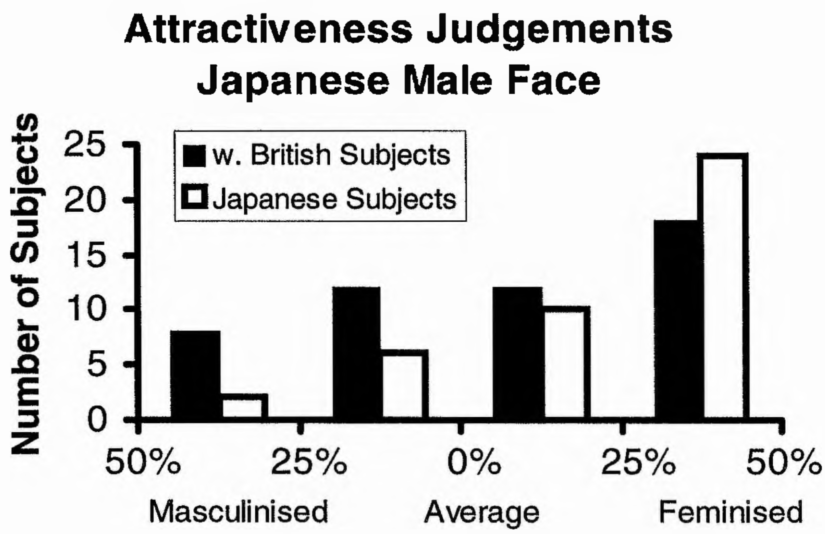
For male stimuli there was no main effect of subject gender ($F_{1,88}$=0.18, p=0.67, mse=0.002), subject race ($F_{1,88}$=2.94, p=0.09, mse=0.25) or race of stimulus face ($F_{1,88}$=0.018, p=0.89, mse=0.0007). No interactions were found. Results are, however, summarised for both white British and Japanese subjects.

white British subjects

Results were analysed in a similar manner to the female stimuli. All results were tested against a hypothesised mean of 0% change in femininity/masculinity. That is feminising or masculinising the face shape would have no effect on perceived attractiveness and subjects would prefer the average shape. Feminised versions of males were also the most preferred. On average a Japanese male face which had been feminised in shape by 9% was selected ($p=0.03$, $t=-2.2$, $df=49$) while a white British male face was feminised, on average, by 15% ($p<0.001$, $t=-4.2$, $df=49$). This is illustrated in *Figure 8.7.*

Japanese subjects

As indicated by the ANOVA Japanese subjects repeated the preference pattern found with white British viewers. More feminine versions of both Japanese (feminised in shape by 20%; $p<0.001$, $t=-6.5$, $df=41$) and white British (feminised in shape by 17%, $p<0.001$, $t= -4.8$, $df=41$) male faces were also preferentially selected (*Figure 8.7*).



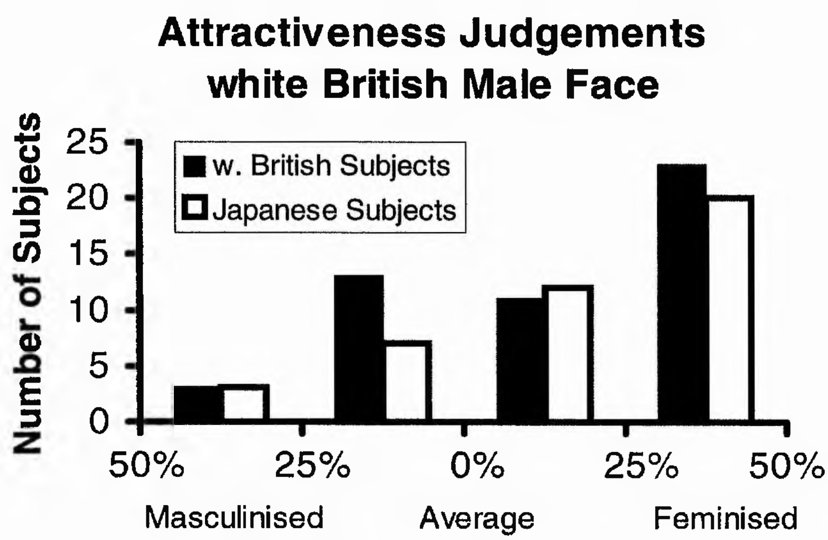


Figure 8.7. Attractiveness judgements of male faces. Both Japanese and white British subjects selected feminised faces to be more attractive with Japanese and white British male stimuli.

Cross-race comparison

As indicated through no significant interactions in the ANOVA, the distribution of male stimuli selected as the most attractive was the same for both Japanese and white British viewers and a similar degree of feminisation was chosen for both Japanese and white British faces.

Discussion

In this study facial shapes that were more feminine were found to be more attractive in both male and female stimuli. There were no differences between male or female subjects in attractiveness judgements. For male stimuli, there was no difference within or across-race in the pattern of feminisation for preferred faces. Within-race, female images feminised more highly in shape were selected as attractive than across-race for both white British and Japanese viewers of both sexes. The use of an interactive technique to investigate the impact of femininity/masculinity on attractiveness for faces proved to be a very sensitive measure. It was able to detect subtle differences in attractiveness preferences between Japanese and white British subjects when looking at faces derived from their own or another culture.

Conclusion

The findings of the studies described in this chapter reinforce the hypothesis that the average configuration, as stipulated by Langlois and Roggman (1990), is not the most attractive. Previous work (Perrett *et al*, 1994) showed that attractive faces differed from the population average in shape. This research did not draw a conclusion as to which dimension of non-averageness was important for attractiveness although measurements for the more attractive shape (e.g. large eyes and small chin) did correlate with the differences between females and males. The studies here confirm the importance of masculinity/femininity as a dimension critical to attractiveness. In these studies faces which are more feminine than the average configuration have been consistently found to be more attractive, both for female and male faces. This finding appears to persist across both culture of face stimuli and culture of viewer.

The finding that more feminine female faces shapes are consistently found to be attractive for female stimuli concurs with evolutionary hypotheses for sexual selection. Such a facial configuration in females may hold cues to health and fertility through a display of gender typicality and apparent youth (Symons, 1979). The smaller nose and chin and fuller lips inherent in the more feminine face shape (*Figures 8.2* and *8.4*) is suggestive of the presence of high oestrogen/androgen ratio for circulating hormones. The implication through indicating such a high oestrogen level is that women with the resultant configuration may appear to be more healthy and to have a high reproductive potential. It need not be that the viewer overtly assesses the face to appear more healthy or to reflect high fecundity as evolutionary theory does not require observers to realise they are selecting for reproductive potential. It is sufficient that viewers find such features, which may reflect such potential, attractive. This is supported by the finding that women with low waist-to-hip ratios (WHRs), a result of selective deposition of adipose tissue resulting from high oestrogen levels, are less at risk for diabetes, gall bladder disease, cancers and disorders of the circulatory system (Barber, 1995). Women with a low WHR also tend to have their first child at an earlier age than those for whom this indicator of circulation oestrogen level is higher (Kaye *et al*, 1990). Such a configuration may also be indicative of related factor such as youth (indicating long-term reproductive fitness) and sexual maturity (indicating onset of the reproductive life-span). In future research it should be possible to establish the relationship between WHR and an attractive female face shape.

The configuration of the Japanese and White female face shapes found to be more attractive revealed many of the deviations from the average found in the more attractive female face composites described by Perrett *et al* (1994). Larger eyes, a smaller nose and chin and higher cheekbones were present in the more attractive faces in both previous (Cunningham, 1986; Johnston and Franklin, 1993; Perrett *et al, 1994*) and present studies. This indicates that a female face shape may be preferred with both youthful features (e.g. large eyes, small chin and full lips) and some sexually mature features (e.g. high, narrow cheeks) found in a face which is feminised to a limited degree. The female face shapes judged here as the more attractive were also not necessarily seen as younger, suggesting that it may indeed be the case that faces are found to be attractive which appear to be sexually mature but have some features which are more typical of a younger individual (Henss, 1991; Johnston and Franklin, 1993; Jones, 1995). Such a facial configuration would be evident for those individuals which had received relatively little influence from the influx of adrenal androgens which lead to facial growth in women at puberty (Enlow, 1990).

The traits found to be attractive in the female faces need not necessarily be selected as indicative of fertility and thus the female face configurations selected as most attractive need not necessarily be explained by the evolutionary hypothesis. In terms of cognitive theory, there may be a strong degree of correlation between Japanese and British standards of beauty and fashion as a result of a greater amount of shared media exposure between the two cultures. The distortions from average which are found to be attractive here need not necessarily reflect evolutionarily desirable qualities such as fertility but may display shape-characteristics highlighted by popular cosmetics.

It is possible that many of the features which relate to attractiveness and are found in this feminine component also elicit the reactions given to neonatal and expressive stimuli. Sexual dimorphism in faces is such that the female face shape has many features in common with neonatal faces (females have smaller faces than males, larger eyes relative to the size of the face and less protuberant chins). Female faces may, as a rule, exhibit these neonate and expressive features which have been associated with more attractive faces since such features tend to be sexually dimorphic and more evident in female than male faces (Enlow, 1990; Nakdimen, 1984). Baby-faced individuals are generally found to be more attractive to viewers (Korthase and Trenholme, 1982; McArthur and Berry, 1987). Their faces may elicit feelings of care-giving to the viewer. This finding may relate not only to males judging females in this study but equally to females judging males.

Male faces were also selected as attractive if they had a more feminine shape configuration. In addition to dominant, and possibly intimidating, features women are found to be attracted to males who are likely to more sensitive, warm, understanding and sincere (Berry, 1991; Jankowiak, Hill and Donovan, 1992). Females appear to possess more youthful features than males (Barber, 1995) and as such feminising male faces may induce them to appear more neonate and therefore more attractive. The elements of youth emphasised by feminisation could be leading to greater attractiveness of the male faces judged as attractive in the studies here. Such male faces with a more feminine configuration also have a more-raised eyebrow ridge and a larger smile. The eyebrow and mouth shape changes may lead to greater expressiveness in the face. They may be seen as more "happy" or "thoughtful": character traits which women have noted to be more attractive in men (Jankowiak *et al*, 1992).

The preference pattern for judging female attractiveness was also fairly consistent in a cross-cultural paradigm. However, selection for attractiveness of faces may be mediated somewhat by cultural factors. Indeed, in Experiment 9 subjects using the interactive paradigm introduced a greater degree of feminisation for the female faces they selected for attractiveness for same-race faces than other-race faces. Through our experience with the outside world it is likely that we develop some kind of norm by which we are able to recognise faces (see *Chapter 1*). This process may also be linked to a mechanism relating to the configuration of faces and mediating the degree of deviation from some norm which eventually becomes stable in adult life (in line with the thoughts of Symons, 1979 who proposed an innate mechanism which could detect the population mean of anatomical features). Symons (1979), however, argued that any deviation from this norm would result in a face being seen as less attractive. The findings here indicate that there may a be deviation detecting mechanism which is important biologically, providing information which indicates a balance between youthful features (indicating health or long-term reproductive success) and sexually mature features. As a result, faces which lie at a certain distance along a specific vector from this norm would be seen as the most attractive. This is suggested by the findings here. This information may be implicit in a "feminine" face to some degree. When applied to a face from another culture the amount of deviation from this norm which is deemed to be attractive may be attenuated in some way in order to counteract differences in mean facial configuration between cultures. Here this seems to be revealed by the finding that the level of feminisation for other-race female faces selected as most attractive is less than for own-race faces when the subject is interactively producing faces which they see as most attractive. It is perhaps this innate, personalised norm building mechanism which allows us through experience to select for optimal attractiveness and we then select faces for attractiveness which differ from this norm in varying ways. These differences may be those acted upon by sexual selection to indicate health, youth or reproductive success through gender typicality as shown for female faces here. This may help to explain why preferences for a feminised female face shape were stronger within-culture than across-culture.

Unlike that for female faces, the configuration of male faces found to be attractive was consistent across race and sex of face stimuli and race and sex of subject. The difference in this finding for male faces and female faces may highlight the differences in strategy for mate selection between women and men. Previous research suggests that across a wide range of cultures men consistently use physical cues as their main criteria in selecting a female mate. These cues invariably attest to the fertility, health, and reproductive potential of the prospective mate. Conversely, men with the means to invest in potential offspring are most often sought and this need not be implicit in physical attributes. In modern, developed society face cues which reflect a warm, positive, caring expression in the individual or elicit care-giving responses from the viewer may be sufficient for such a person to be seen as an attractive possibility as a mate (Buss, 1987, 1989; Berry, 1991).

Studies of individuals with damage to their facial recognition system have provided evidence in line with the explanation suggested above for the cross-cultural differences found for the ideal facial configuration in female stimuli. RP, a patient suffering from prosopagnosia, a deficit which prevents an individual from being able to identify familiar faces, was found to give facial attractiveness ratings which differed strongly from both age-matched and younger raters (Brewster, 1997). RP's ratings were consistent across repeated trials. This finding indicates that the same deficit which disrupts face recognition may also alter our perception of attractiveness for people's faces. Exaggerating facial deviations from a digitally produced average face, or norm, to produce caricatures has been found to improve recognition performance (Rhodes *et al*, 1987; this thesis, *Chapter 2* and *Chapter 4*). This suggests that facial identification is supported by mechanisms equivalent to those described for attractiveness. It may be that brain damage in prosopagnosia prevents access to mechanisms through which faces are identified through comparison to a norm. If attractiveness judgements are also mediated by such a process then they too will be affected in prosopagnosia suffers.

The development of the "norm" engaged in this process is, of course, mediated further by social factors such as status or clothing and cognitive factors such as stereotyping which differ across individuals and cultures (Kowner and Ogawa, 1993). For instance, gender roles in a given society may affect the degree to which apparent femininity or masculinity is beneficial to facial attractiveness. Bernstein, Lin and Mclellan (1982) noted that various ethnic groups use slightly different standards of attractiveness, but they are capable of applying these rules to other groups as well as their own. Further, cultural factors such as pathogen prevalence will also influence the role of physical attractiveness to mate selection (Gangestad and Buss, 1993). To this extent personal attractiveness judgements surely contain a subjective component.

In linking the study of attractiveness to previous studies of identification one can turn to the potential role of colour. The role of colour in perceived attractiveness could also be examined using similar techniques. Colour has already been shown to be an important aspect of faces in recognition (*Chapters 3 and 4*) and likeness judgements (*Chapter 6*) and perception of age (Burt and Perrett, 1995). It may also be that colour plays an important role in attractiveness for example as an expression of gender typicality,. It has been shown that fair skin is found to be attractive in females and that female skin-colour is generally lighter than males (Kalla and Tiwari, 1970; van den Berghe and Frost, 1986) and mediated by oestrogen production such that skin-colour even varies during pregnancy and phases of the menstrual cycle (van den Berghe and Frost, 1986). It may be that dark-skinned males are, on the whole, found to be more attractive as there is anecdotal evidence that women prefer men who are "tall, dark and handsome" (e.g. Feinman and Gill, 1978). Again, dark colouration in males may be linked to testosterone production (van den Berghe and Frost, 1986). Techniques are becoming available which would allow a viewer to manipulate images in two dimensions so both colour and texture could be varied until the subject found the ideal configuration of shape and colour. This would give information as to the viewer's idealised preference, for say gender typicality, as indicated by the final position in both dimensions as these judgements are likely to be made together in real-life.

Plate 8.1. Composites were formed by warping the texture of all the Japanese female (left) and male (right) faces into the average shape for that gender and blending the pixel values together. The Japanese male was normalised so as to have the same inter-ocular distance as the female.

Plate 8.2. Subjects were presented with this stimuli slide such that the Japanese female composite which had been masculinised in shape by 50% (I1) was displayed alongside the veridical female average (I2). Subjects were asked to indicate on their response sheets which of the two images they preferred. In 50% of the trials the images were reversed so that face I2 appeared on the left and face I1 appeared on the right. Both images retain the colour of the average female face.

Plate 8.3. Subjects were presented with this stimuli slide such that the veridical Japanese female average (J1) was displayed alongside the female composite which was feminised in shape by 50% (J2). Subjects were asked to indicate on their response sheets which of the two images they preferred. In 50% of the trials the images were reversed so that face J2 appeared on the left and face J1 appeared on the right. Both images retain the colour of the average female.

Plate 8.4. Subjects were presented with this stimuli slide such that the veridical Japanese male average (K1) was displayed alongside the male composite which had been masculinised in shape by 50% (K2). Subjects were asked to indicate on their response sheets which of the two images they preferred. In 50% of the trials the images were reversed so that face K2 appeared on the left and face K1 appeared on the right. Both images retain the colour of the average male.

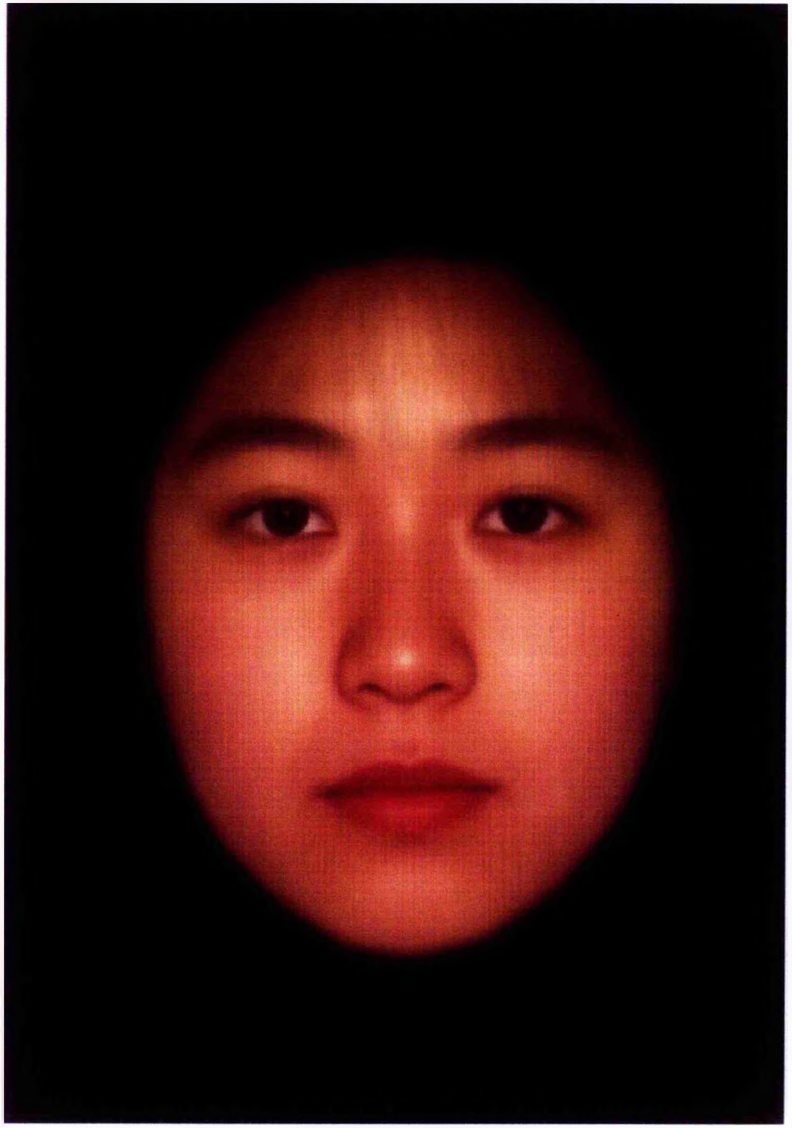
Plate 8.5. Subjects were presented with this stimuli slide such that the Japanese male composite which had been feminised in shape by 50% (L1) was displayed alongside the average male composite (L2). Subjects were asked to indicate on their response sheets which of the two images they preferred. In 50% of the trials the images were reversed so that face L2 appeared on the left and face L1 appeared on the right. Both images retain the colour of the average male.

Plate 8.6. Composites were formed by warping the texture of all the white British female (left) and male (right) faces into the average shape for that gender and blending the pixel values together. The white British female was normalised so as to have the same inter-ocular distance as the male.

Plate 8.7. Subjects were presented with this stimuli slide such that the white British female composite which had been feminised in shape (C1) was displayed alongside the female veridical average (C2). Subjects were asked to indicate on their response sheets which of the two images they preferred. In 50% of the trials the images were reversed so that face C2 appeared on the left and face C1 appeared on the right. Both images retain the average female colour.

Plate 8.8. Subjects were presented with this stimuli slide such that the veridical white British female average composite (D1) was displayed alongside the female composite which had been masculinised in shape (D2). Subjects were asked to indicate on their response sheets which of the two images they preferred. In 50% of the trials the images were reversed so that face D2 appeared on the left and face D1 appeared on the right. Both images retain the average female colour.

Plate 8.9. Subjects were presented with this stimuli slide such that the white British male composite which had been feminised in shape (A1) was displayed alongside the male veridical average composite (A2). Subjects were asked to indicate on their response sheets which of the two images they preferred. In 50% of the trials the images were reversed so that face A2 appeared on the left and face A1 appeared on the right. Both images retain the average male colour.



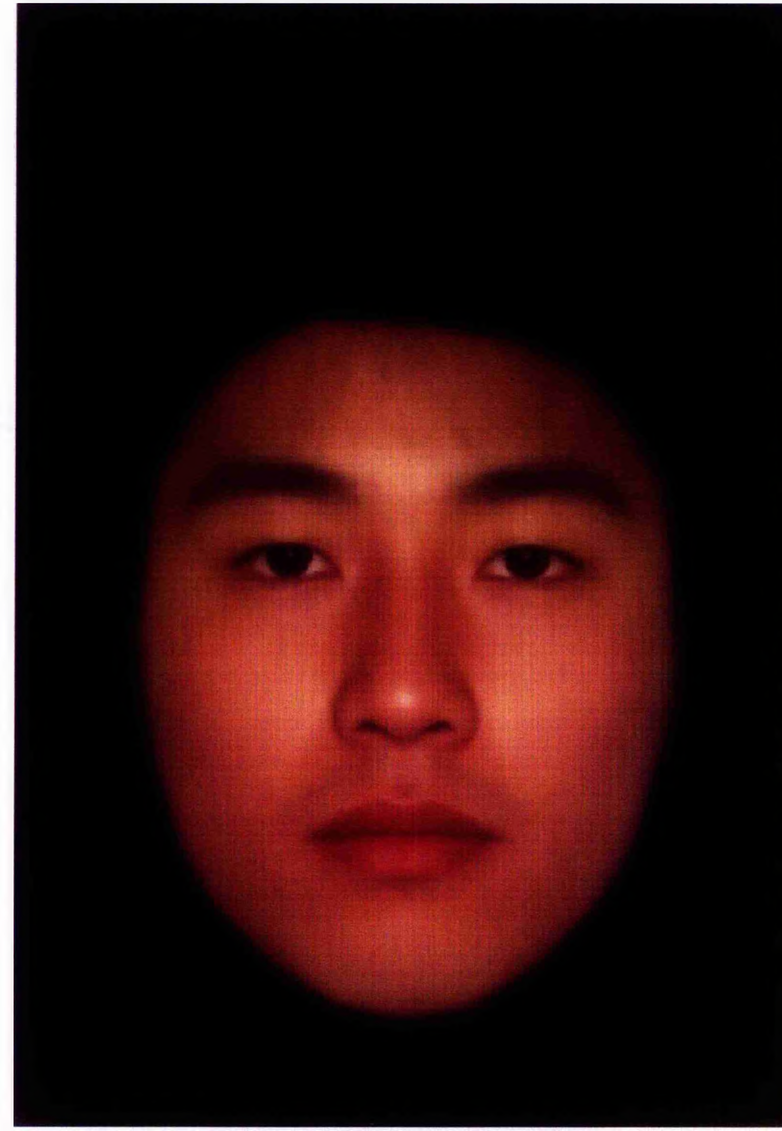


Plate 8.10. The Japanese male and female blends were cropped to remove external features such has hair and ears and the neck area.

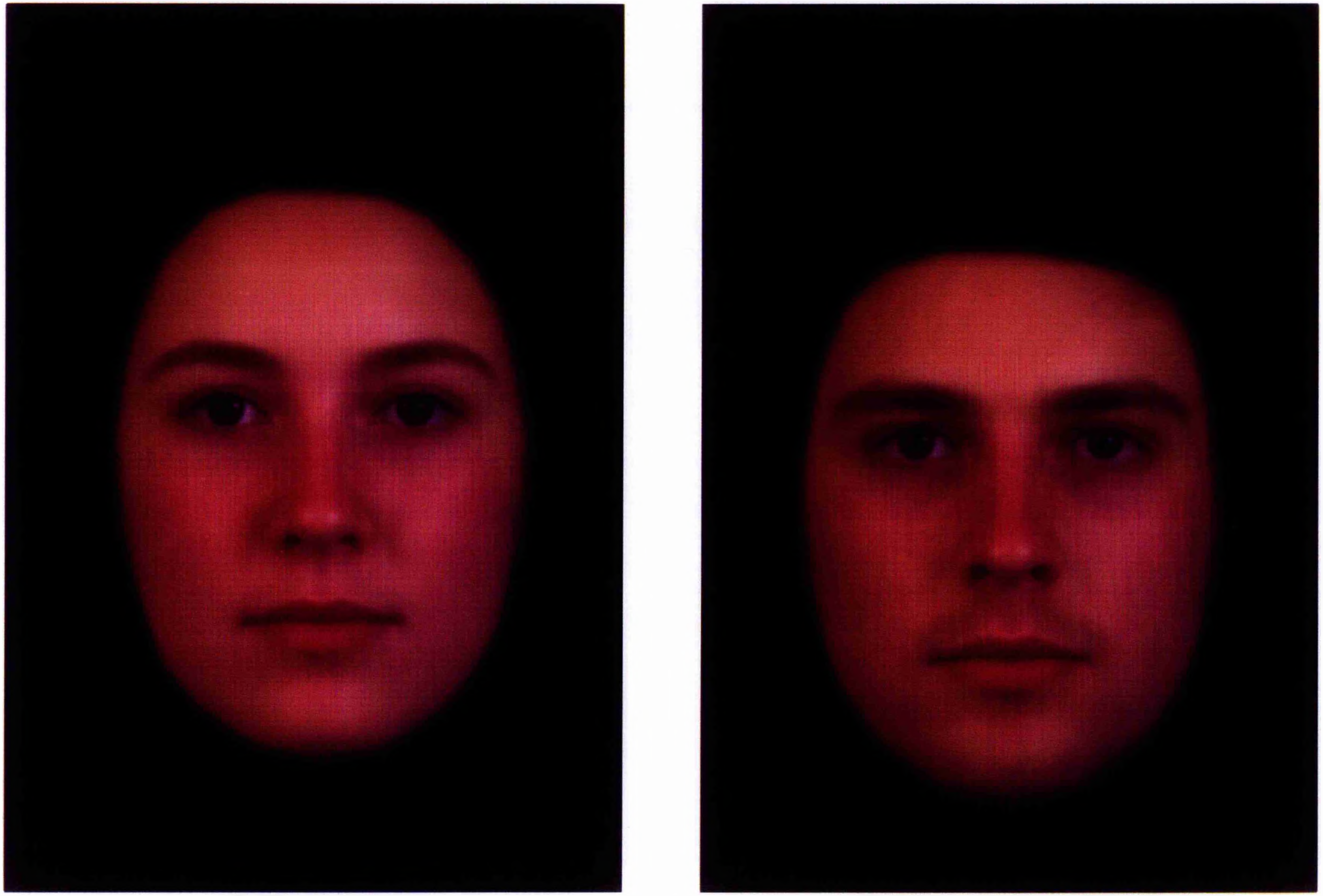
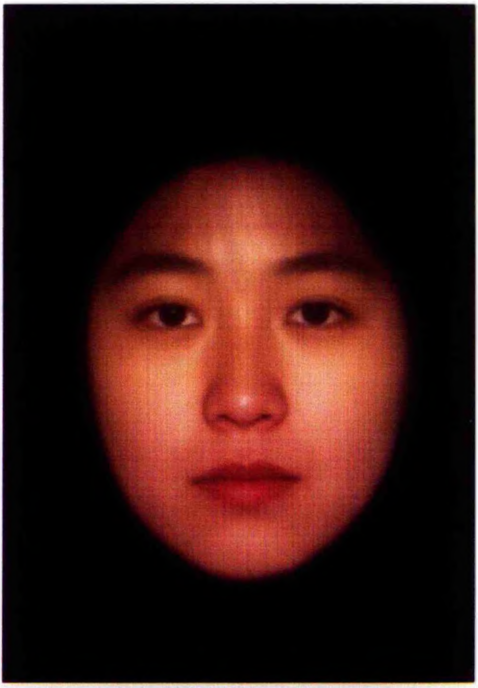
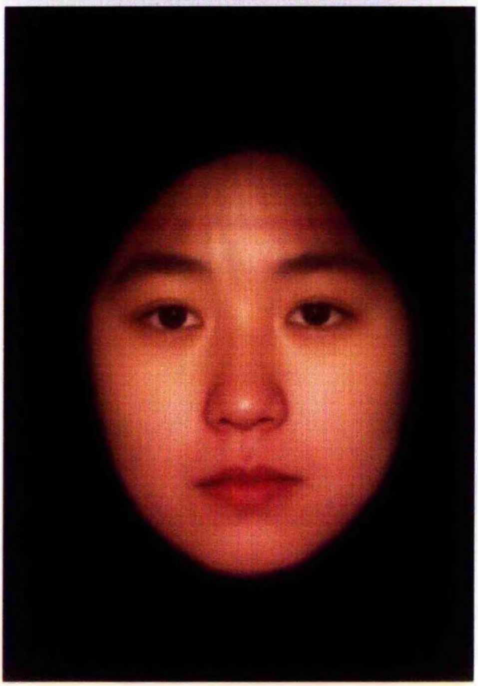


Plate 8.11. The white British male and female blends were cropped to remove external features such has hair and ears and the neck area.

Plate 8.12. These images illustrate the range of choice the subject is given to select which he/she believes to be the most attractive Japanese female. The top left image is masculinised in shape by 50%, the top middle image is masculinised in shape by 25% and the top right image describes the average Japanese female shape. The bottom left and bottom right images have been feminised by 25% and 50% respectively. All images retain the texture for the average Japanese female. When individuals manipulate the image on-screen it moves fluidly between these five images and all the possible configurations in-between.





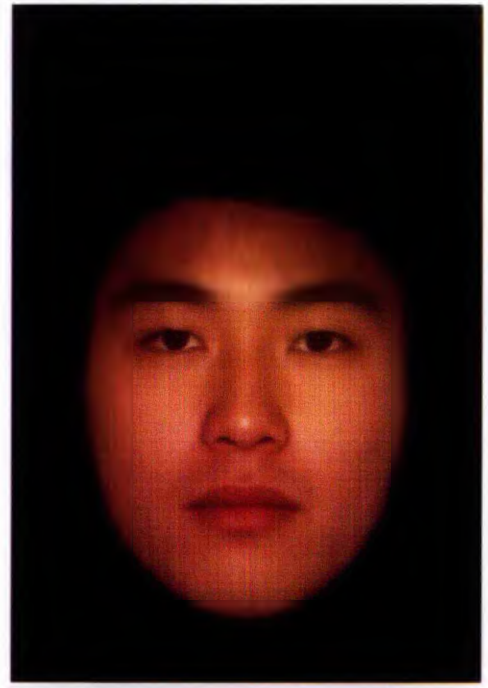
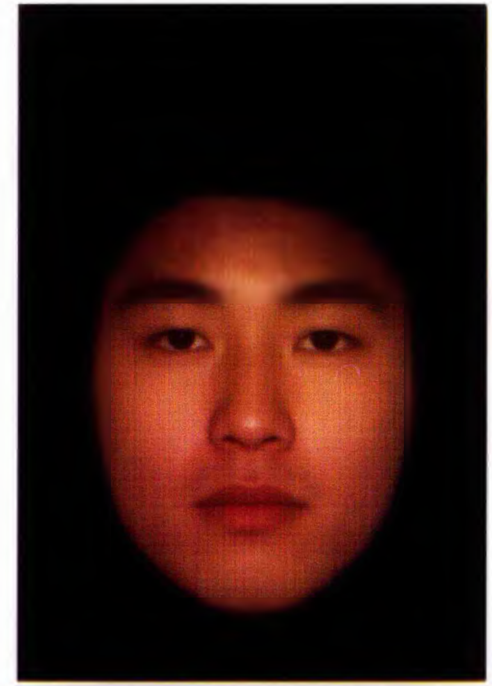
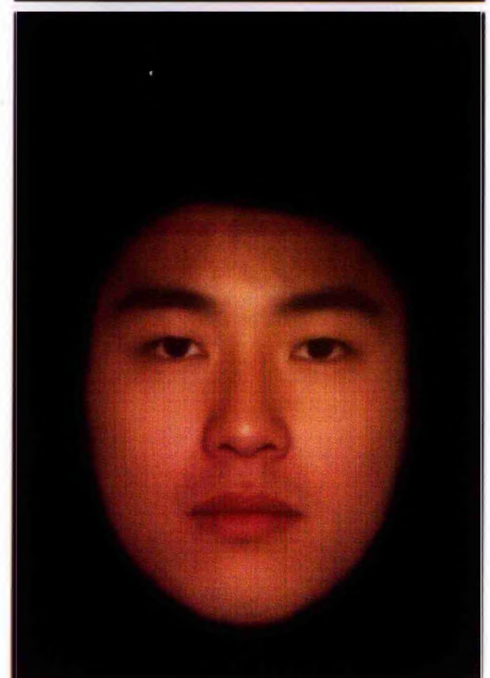
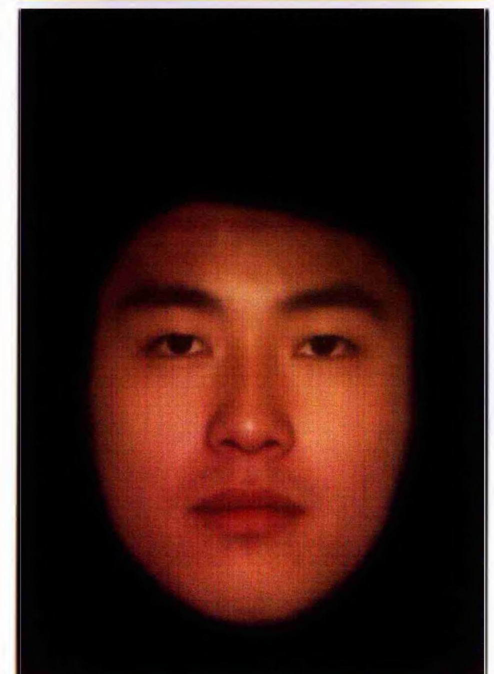
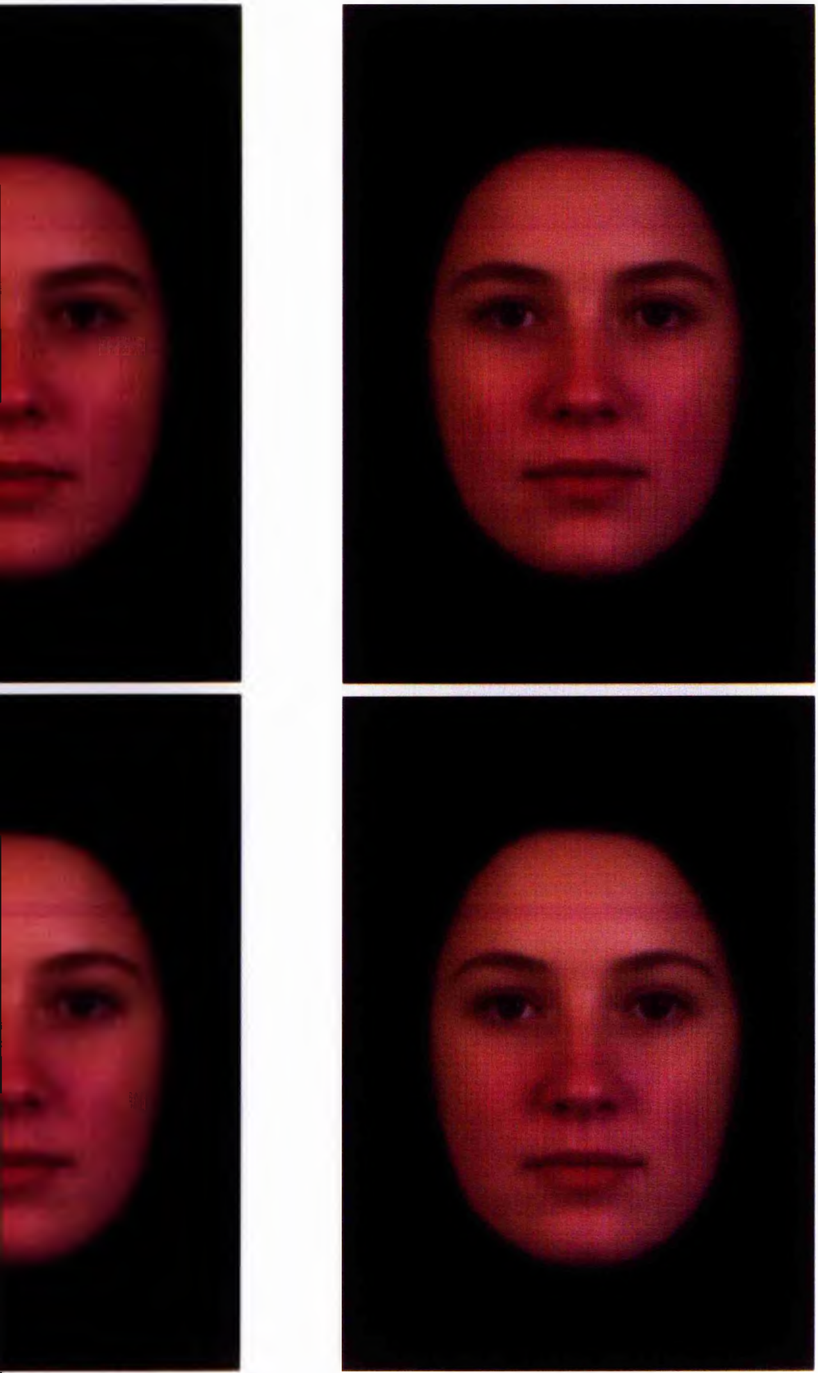






Plate 8.13. These images illustrate the range of choice the subject is given to select which he/she believes to be the most attractive Japanese male. The top left image is feminised in shape by 50%, the top middle image is feminised in shape by 25% and the top right image describes the average Japanese male shape. The bottom left and bottom right images have been masculinised by 25% and 50% respectively. All images retain the texture for the average Japanese male. When individuals manipulate the image on-screen it move fluidly between these five images and all the possible configurations in-between.

Plate 8.14. These images illustrate the range of choice the subject is given to select which he/she believes to be the most attractive white British female. The top left image is masculinised in shape by 50%, the top middle image is masculinised in shape by 25% and the top right image describes the average white British female shape. The bottom left and bottom right images have been feminised by 25% and 50% respectively. All images retain the texture for the average white British female.



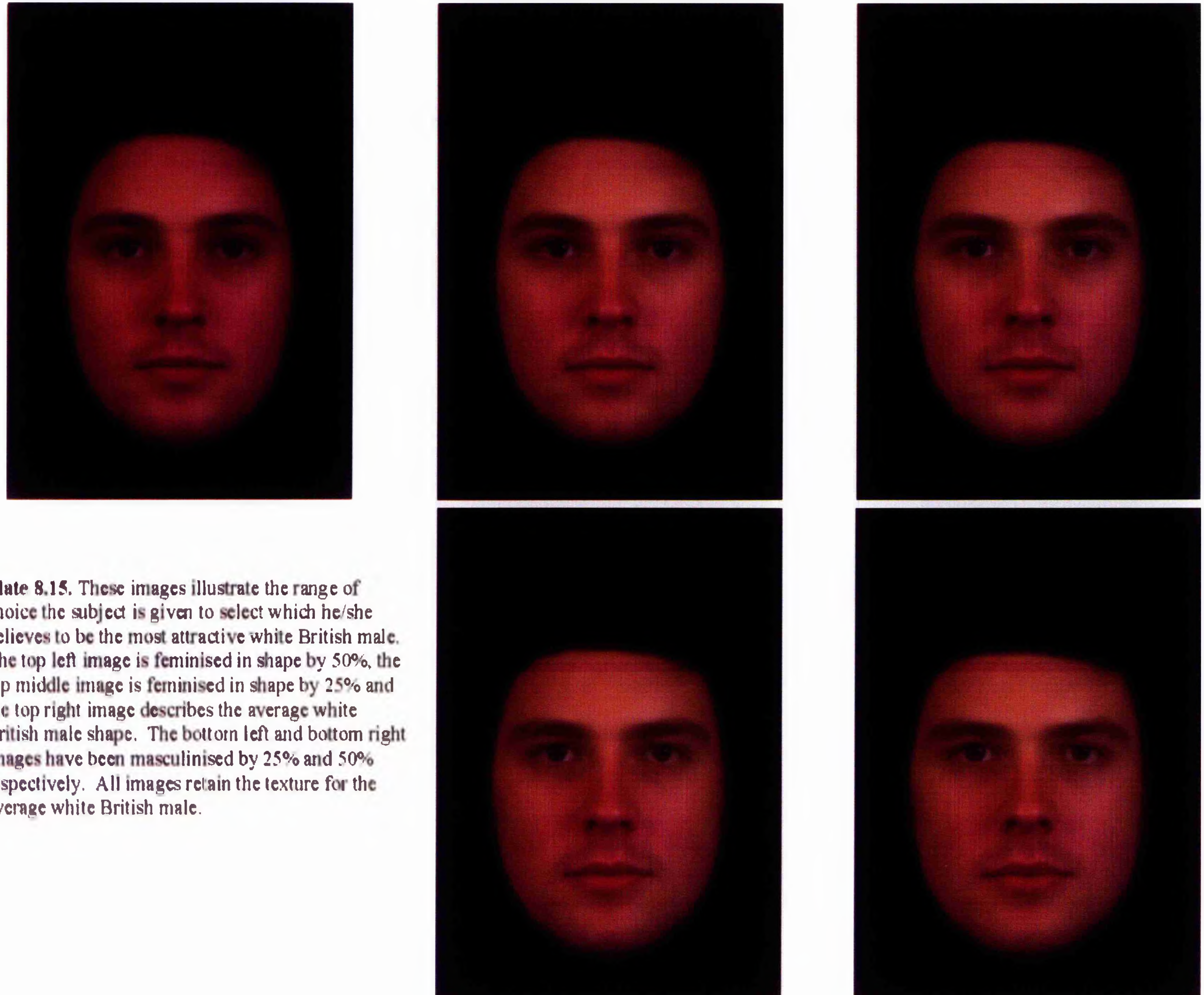


Plate 8.15. These images illustrate the range of choice the subject is given to select which he/she believes to be the most attractive white British male. The top left image is feminised in shape by 50%, the top middle image is feminised in shape by 25% and the top right image describes the average white British male shape. The bottom left and bottom right images have been masculinised by 25% and 50% respectively. All images retain the texture for the average white British male.

Chapter 9 - Conclusions and Lines for Future Research

General Conclusion

In summary, this thesis describes the use of digital manipulation to study perception of facial identity and attractiveness. A presentation time paradigm was developed by which stimuli could be presented for a range of brief display periods. Using this paradigm, subjects recognised photo-realistic target faces caricatured in shape with greater accuracy than veridical images consistent with the facilitation of reaction time with caricatures in prior studies. Subjects were further required to identify colour representations of famous faces which were either veridical, caricatured in colour space or had enhanced colour saturation and intensity contrast (as contrast controls). Recognition accuracy was greater when viewing the colour-caricatured stimuli than either the veridical images or the contrast controls. The removal of colour to produce grey-scale images also decreased accuracy of face recognition indicating that colour information aids facial identification.

Caricaturing of faces, therefore, can be extended to the colour domain and, as with shape caricaturing, enhancement of distinctive information can produce a recognition advantage for famous faces. Subjects were also asked to identify the best-likeness for individuals using photo-realistic stimuli and an interactive paradigm with shape-caricature, colour-caricature and contrast-control varied by the user in real-time. The best-likeness with shape manipulation was a slight anti-caricature while with colour-caricature and contrast-control images a mildly exaggerated image was selected as the best-likeness. Thus although images caricatured substantially in colour or shape (+40%) induce superior recognition compared to veridical images such substantial exaggerations are not seen as best-likenesses under prolonged exposure.

The use of composite images was applied to attractiveness judgements of faces. The gender typicality for white British and Japanese composite faces was manipulated and subjects were presented with the images using both a static forced-choice paradigm and an interactive paradigm. Male and female face shapes with enhanced feminine features were consistently found to be more attractive than average. Preference for enhanced femininity for female faces was greater for subjects making within- than cross-cultural judgements.

All the findings described lend support to the hypothesis that through our perceptual development some process occurs by which our lifetime experience of other faces mediates our subsequent recognition and judgements of attractiveness. It is suggestive of a process through which, when we become very familiar with a face, the information regarding the colour and shape that is distinctive to that particular face is encoded. When we later see that person the information which distinguishes their face triggers recognition. These findings suggest that this occurs by identifying a face through the set of features for this individual which deviate most from what has been learnt from the general population of faces previously encountered and thus make it unique. Through our vast experience of faces our understanding of the information which describes a face becomes quite concrete in the manner of a prototype (Rhodes *et al*, 1987; Valentine and Bruce, 1986, 1986a) and atypical faces, even those that are very familiar to us, which lie at a distance in some "face space" are likely to be seen as unusual.

The conclusion from the studies of identification described here is that exaggerating features which are distinctive to an individual improves recognition. In a parallel fashion, although it is widely accepted that average faces are attractive it has been shown here that exaggerating feminine features enhances judgements of attractiveness. The thesis has therefore shown how computer graphics can be applied to the study of two different facial dimensions (identity and attractiveness). For both dimensions the application of computer graphics has revealed that movements in "face space" away from a central facial norm can facilitate visual processing and produce faces which are perceived to be more attractive.

Future Research

Through the use of improved technology and more rigorous methodology the techniques described here can be extended further to allow more in-depth study of both facial identity and attractiveness. It is possible for many of the presentation-time experiments described in this thesis to be replicated using masks which are more appropriate to the stimuli which they accompany (e.g. a black and white mask with black and white stimuli and a colour mask with colour stimuli in a replication of Experiment 2). Further, consistently full and complete counterbalancing would assure that the findings are shown conclusively.

The experiments described in this thesis touch on the relationship between factors such as the relationship between distinctiveness of faces, recognition accuracy and the degree of best-likeness selected. These relationships could be explored further by obtaining viewers' perceptions of the distinctiveness of the face shape of different individuals. Interactive best-likeness experiments, for instance, could then be performed to determine the degree of caricature selected for best-likeness for each individual and the relationship between distinctiveness and degree of caricature. It may be that the most distinctive, or identifiable, faces are those which are given the greatest anti-caricatures in such a paradigm. Similarly, using a presentation time paradigm it could be determined whether faces seen as more distinctive are recognised with greater accuracy than more typical faces and if caricaturing gives greater benefit to recognition with decreased distinctiveness. Similarly, if colour and texture information could be isolated, the relationship between typicality and recognition or degree of caricature selected for best-likeness could be investigated in these separate domains.

Techniques for both the manipulation of shape information and colour detail are described in this thesis. The refinement of such techniques would allow the determination of the relative role of shape and colour in identification of faces. One related line of research, for instance, could be to show subjects randomised faces of celebrities in which the stimulus displayed is either an average, composite shape containing the colour and texture information for the individual or a composite, blended face containing colour information across a group of faces warped into the shape of an individual. Again, the reduction of information available to recognition by either removing relevant shape information or relevant colour and texture information serves to degrade the amount of information available to the viewer for subsequent identification even with prolonged display times. These stimuli could either be displayed using a presentation-time paradigm or the degradation may be sufficient such that even with extended viewing the subjects ability to recognise certain individuals from the relevant stimulus is impaired. Through showing subjects the veridical representations of the target individuals, the experimenter would be able to determine whether the pictures containing all the information available for each given face were recognisable by viewers. Relative recognition accuracy for images containing only relevant shape or colour information would indicate whether colour or shape gives equal or a biased contribution to the processing required for facial identification.

The separation of colour and shape information could also be employed to extend the attractiveness experiments described in this thesis. As noted in Chapter 8, if the presence of facial hair could be controlled more effectively on male facial stimuli, colour of both male and female stimuli could be modified such that facial colour was either more or less typical for the given gender. Further, the technology is now available to allow viewers to manipulate images in both the colour and shape dimensions until an ideal face is arrived at. These manipulations could be extended to look at further facial attributes such as systematic variation of youth as evident in faces and its contribution to perceived attractiveness.

It may be that hormonal changes which influence body development are also reflected in perceptions of attractiveness. If waist-to-hip ratio is an index of attractiveness, such body information could be obtained and compared with attractiveness ratings. In this instance, for example, face shape and colour information could be captured for individuals with varying WHRs and the associated composites produced. Stimuli could be presented with either colour or shape information manipulated in isolation or in tandem and using either static stimuli or the interactive technique described here. Such a manipulation would enable the determination of whether the shape and/or colour information present in faces which are viewed as attractive is correlated with the lower WHRs found in attractive women or the high WHRs found to be attractive in men. If there is indeed no correlation between WHR and facial attractiveness this could be shown in such a study.

The ability to separate colour and shape in stimulus presentation in this way could allow us not only to gain an understanding of their respective contributions to identity and attractiveness, but also to other facial contributions such as gender or race. The study of this cross-culturally and with varying age-groups, utilising viewers of differential experience of different types of face would allow for a much more detailed understanding of the development and expression of face processing.

References

Alicke M D, Smith R H, Klotz M L, 1987 "Judgments of physical attractiveness: the role of faces and bodies" *Personality and Social Psychology Bulletin* **12** 381-389

Alley T R, 1983 "Age-related changes in body proportions, body size, and perceived cuteness" *Perceptual and Motor Skills* **56** 615-622

Alley T R, Hildebrandt K A, 1988 "Determinants and consequences of facial aesthetics", in *Social and applied aspects of processing faces* Ed T R Alley (Hillsdale, NJ: Lawrence Erlbaum) pp 101-140

Alley T R, Cunningham M R, 1991 "Average faces are attractive, but very attractive faces are not average" *Psychological Science* **2** 123-125

Andersson M B, 1994 *Sexual Selection* (Princeton, NJ: Princeton University Press)

Barber N, 1995 "The evolutionary psychology of physical attractiveness: sexual selection and human morphology" *Ethology and Sociobiology* **16** 395-424

Bartlett J C, Hurry S, Thorley W, 1984 "Typicality and familiarity of faces" *Memory and Cognition* **12** 219-228

Bartlett J C, Searcy J, 1993 "Inversion and configuration of faces" *Cognitive Psychology* **25** 281-316

Bateman A J, 1948 "Intrasexual selection in Drosophilia" *Heredity* **2** 349-368

Benson P J, Perrett D I, 1991 "Perception and recognition of photographic quality facial caricatures: implications for the recognition of natural images" *European Journal of Psychology* **3** 105-135

Benson P, Perrett D, 1991a "Computer averaging and manipulation of faces", in *Photovideo: Photography in the age of the computer* Ed P Wombell (London: Rivers Oram Press) pp 32-38

Benson P J, Perrett D I, 1991b "Synthesizing continuous-tone caricatures" *Image and Vision and Computing* **9** 123-129

Benson P, Perrett D, 1992 "Face to face with the perfect image" *New Scientist* 32-35

Benson P J, Perrett D I, 1993 "Extracting prototypical facial images from exemplars" *Perception* **22** 257-262

Benson P J, Perrett D I, 1994 "Visual processing of facial distinctiveness" *Perception* **23** 75-93

Benson P J, Perrett D I, Unpublished "Recognition of famous-face photographic-quality caricatures from their internal facial-features"

Bermant G, Davidson J M, 1974 *Biological bases of sexual behavior* (New York: Harper & Row)

Bernstein I H, Lin T, McLellan P, 1982 "Cross- vs. within-racial judgments of attractiveness" *Perception and Psychophysics* **32** 495-503

Berry D S, McArthur L Z, 1985 "Some components and consequences of a babyface" *Journal of Personality and Social Psychology* **48** 312-323

Berry D S, McArthur L Z, 1986 "Perceiving character in faces: the impact of age-related craniofacial changes on social perception" *Psychological Bulletin* **100** 3-18

Berry D S, McArthur L Z, 1988 "What's in a face? Facial maturity and the attribution of legal responsibility" *Personality and Social Psychology Bulletin* **14** 23-33

Berry D S, 1991 "Attractive faces are not all created equal: joint effects of facial babyishness and attractiveness on social perception" *Personality and Social Psychology Bulletin* **17** 523-531

Berscheid E, Dion K, Walster E, Walster G W, 1971 "Physical attractiveness and dating choice: A test of the matching hypothesis" *Journal of Experimental Social Psychology* **7** 173-189

Biederman I, 1987 "Recognition-by-components: A theory of human image understanding" *Psychological Review* **94** 115-147

Biederman I, Ju G, 1988 "Surface versus edge-based determinants of visual recognition" *Cognitive Psychology* **20** 38-64

Björntorp P, Suppl. "The associations between obesity, adipose tissue distribution and disease" *Acta Medica Scandinavica* **723** 121-134

Brehm S S, 1985 *Intimate Relationships* (New York: Random House)

Brennan S E, 1982 *Caricature generator* PhD, MIT, Cambridge, MA

Brennan S E, 1985 "The caricature generator" *Leonardo* **18** 170-178

Brewster T, 1997 Honours, University of St Andrews, St Andrews

Brodie E E, Wallace A M, B S, 1991 "Effect of surface characteristics and style of production on naming and verification of pictorial stimuli" *American Journal of Psychology* **104** 517-545

Brown E, Perrett D I, 1993 "What gives a face its gender?" *Perception* **22** 829-840

Bruce V, Healey P, Burton A M, Doyle T, 1991 "Recognising facial surfaces" *Perception* **20** 755-769

Bruce V, Burton A M, Dench N, 1994 "What's distinctive about a distinctive face?" *The Quarterly Journal of Experimental Psychology* **47A** 119-141

Bruce V, Langton S, 1994 "The use of pigmentation and shading information in recognizing the sex and identities of faces" *Perception* **23** 803-822

Burson N, Carling R, Kramlich D, 1986 *Composites* (New York: Beech Tree Books, William Morrow & Co.)

Burt D M, Perrett D I, 1995 "Perception of age in adult Caucasian male faces: computer graphic manipulation of shape and colour information" *Proceedings of the Royal Society of London, Series B* **259** 137-143

Buss D M, 1985 "Human mate selection" *American Scientist* **73** 47-51

Buss D M, 1987 "Sex differences in human mate selection criteria: an evolutionary perspective", in *Sociobiology and Psychology: Ideas, Issues, and Applications* Eds C Crawford, M Smith and D Krebs (Hillsdale, NJ: Lawrence Erlbaum) pp 335-352

Buss D M, 1989 "Sex differences in human mate preferences: evolutionary hypotheses tested in 37 cultures" *Behavioral and Brain Sciences* **12** 1-49

Buss D M, 1991 "Do women have evolved mate preferences for males with resources?" *Ethology and Sociobiology* **12** 401-408

Buss D M, 1994 "The strategies of human mating" *American Scientist* **42** 238-249

Buss D M, 1995 "Psychological sex differences - origins through sexual selection" *American Psychologist* **50** 164-168

Calder A J, Young A W, Benson P J, Perrett D I, 1996 "Self priming from distinctive and caricatured faces" *British Journal of Psychology* **87** 141-162

Carey S, Diamond R, 1977 "From piecemeal to configurational representation of faces" *Science* **195** 312-314

Carey S, 1992 "Becoming a face expert" *Philosophical Transactions of the Royal Society of London, Series B* **335** 95-103

Carroo A W, 1986 "Other race recognition: A comparison of Black Americans and African subjects" *Perceptual and Motor Skills* **62** 135-138

Carroo A W, 1987 "Recognition of faces as a function of race, attitudes and reported cross-racial friendship" *Perceptual and Motor Skills* **64** 319-325

Cash T F, Derlega V J, 1978 "The matching hypothesis: Physical attractiveness among same-sexed friends" *Personality and Social Psychology Bulletin* **4** 240-243

Cash T F, Horton C E, 1983 "Aesthetic surgery - effects of rhinoplasty on the social-perception of patients by others" *Plastic and Reconstructive Surgery* **72** 543-548

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

Chance J E, Turner A L, Goldstein A G, 1982 "Development of differential recognition of own and other-race faces" *Journal of Psychology* **112** 29-37

Chiroro P, Valentine T, 1995 "An investigation of the contact hypothesis of the own-race bias in face recognition" *The Quarter Journal of Experimental Psychology* **48A** 879-894

Cohen M E, Carr W J, 1975 "Facial recognition and the von Restorff effect" *Bulletin of the Psychonomic Society* **6** 383-384

Conway D L, Baker P T, 1976 "Skin reflectance of Quechua Indians, the effects of genetic admixture, sex and age" *American Journal of Physical Anthropology* **36** 267-282

Costen N, Craw I, Robertson G, Akamatsu S, 1996 " Automatic face recognition: what representation?" *Lecture Notes in Computer Science* **1064** 504-513

Cross J F, Cross J, Daley J, 1971 "Sex, race, age, and beauty as factors in recognition of faces" *Perception and Psychophysics* **10** 393-396

Cross J F, Cross J, 1971 "Age, sex, race, and the perception of facial beauty" *Developmental Psychology* **5** 433-439

Cunningham M, 1986 "Measuring the physical in physical attractiveness: quasi-experiments on the sociobiology of female facial beauty" *Journal of Personality and Social Psychology* **50** 925-935

Cunningham M, Barbee A P, Pike C L, 1990 "What do women want? Facialmetric assessment of multiple motives in the perception of male facial physical attractiveness" *Journal of Personality and Social Psychology* **59** 61-72

Cunningham M R, Roberts A R, Barbee A P, Druen P B, Wu C-H, 1995 ""Their ideas of beauty are, on the whole, the same as ours": Consistency and variability in the cross-cultural perception of female physical attractiveness" *Journal of Personality and Social Psychology* **68** 261-279

Darwin C, 1871 *The Descent of Man and Selection in Relation to Sex* (London: Murray)

Davidoff J B, Ostergaard A L, 1988 "The role of colour in categorical judgements" *The Quarterly Journal of Experimental Psychology* **40A** 533-544

Davies G M, Ellis H D, Shepherd J W, 1978 "Face recognition accuracy as a function of mode of presentation" *Journal of Applied Psychology* **63** 180-187

Deutsch F M, Zalenski C M, Clark M E, 1986 "Is there a double standard of aging?" *Journal of Applied Social Psychology* **9** 771-785

Diamond R, Carey S, 1986 "Why faces are and are not special: An effect of expertise" *Journal of Experimental Psychology: General* **115** 107-117

Dion K, Berscheid E, Walster E, 1972 "What is beautiful is good" *Journal of Personality and Social Psychology* **24** 285-290

Dion K K, 1973 "Young children's stereotyping of facial attractiveness" *Developmental Psychology* **9** 183-188

Dion K K, 1986 "Stereotyping based on physical attractiveness. Issues and conceptual perspectives", in *Physical appearance, stigma, and social behavior* Eds C P Herman, M P Zana and E T Higgins (Hillsdale, NJ: Lawrence Erlbaum)

Dodd J, Perrett D, Unpublished

Eagly A H, Ashmore R D, Makhijani M G, Longo L C, 1991 "What is beautiful is good, but. . .: A meta-analytic review of research on the physical attractiveness stereotype" *Psychological Bulletin* **110** 109-128

Eibl-Eibesfeldt, 1970 *Ethology, the biology of behavior* (New York: Holt, Rinehart and Winston)

Elliott E S, Willis E J, Goldstein A G, 1973 "The effects of discrimination training on the recognition of white and oriental faces" *Bulletin of the Psychonomic Society* **2** 71-73

Ellis H D, Deregowski J B, Shepherd J W, 1975 "Descriptions of white and black faces by white and black subjects" *International Journal of Psychology* **10** 119-123

Ellis H D, 1975 "Recognizing Faces" *British Journal of Psychology* **66** 409-426

Ellis H D, Shepherd J W, Davies G M, 1979 "Identification of familiar and unfamiliar faces from internal and external features: some implications for theories of face recognition" *Perception* **8** 431-439

Ellis H D, 1981 "Theoretical aspects of face recognition", in *Perceiving and remembering faces* Eds G M Davies, H D Ellis and J Shepherd (London: Academic Press)

Ellis H D, Shepherd J W, 1987 "Analysis of mechanisms by which human observers acquire knowledge of faces", report to the *Economic and Social Research Council* **C00232260**

Enlow D H, 1990 *Facial growth* (Philadelphia: Harcourt Brace Jovanovich)

Etcoff N, 1994 "Beauty and the beholder" *Nature* **368** 186-187

Evans D J, Hoffmann R G, Kalkhoff R K, Kissebah A H, 1983 "Relationship of androgenic activity to body fat topography, fat cell morphology, and metabolic aberrations in premenopausal women" *Journal of Clinical Endocrinology* **57** 304-310

Farkas L G, 1981 *Anthropometry of the face and head in medicine* (New York: Elsevier North-Holland)

Feingold A, 1988 "Matching for attractiveness in romantic partners and same-sex friends: a meta-analysis and theoretical critique" *Psychological Bulletin* **104** 226-235

Feingold A, 1990 "Gender differences in effects of physical attractiveness on romantic attraction: a comparison across five research paradigms" *Journal of personality and social psychology* **59** 981-993

Feingold A, 1992 "Gender selection in mate selection preferences: a test of the parental investment model" *Psychological Bulletin* **112** 125-139

Feingold A, 1992a "Good looking people are not what we think" *Psychological Bulletin* **111** 304-341

Feinman S, Entwisle D R, 1976 "Children's ability to recognise other children's faces" *Child Development* **47** 506-510

Feinman S, Gill G W, 1978 "Sex difference in physical attractiveness preferences" *Journal of Social Psychology* **105** 43-52

Fisher R A, 1930 *The Genetical Theory of Natural Selection* (Oxford: Clarendon Press)

Flin R H, 1985 "Development of face recognition: An encoding switch?" *British Journal of Psychology* **76** 123-134

Folstad I, Karter A J, 1992 "Parasites, bright males, and the immunocompetence handicap" *The American Naturalist* **139** 603-622

Ford C S, Beach F A, 1951 *Patterns of Sexual Behaviour* (New York: Harper)

Franks J J, Bransford J D, 1971 "Abstraction of visual patterns" *Journal of Experimental Psychology* **90** 65-74

Frost P, 1988 "Human skin colour: a possible relation between its sexual dimorphism and its social perception" *Perspectives in Biology and Medicine* **32** 39-58

Furlow F B, Armijow-Prewitt T, Gangestad S W, Thornhill R, 1997 "Fluctuating asymmetry and psychometric intelligence" *Proceedings of the Royal Society of London Series B* **264** 823-829

Galton F J, 1878 "Composite portraits" *Nature* **18** 97-100

Galton F, 1879 "Composite portraits, made by combining those of many different persons into a single resultant figure" *Journal of the Anthropological Institute* **8** 132-144

Galton F, 1907 *Inquiries into human faculty and its development* (New York: Elsevier North-Holland)

Gangestad S W, Buss D M, 1993 "Pathogen prevalence and human mate preferences" *Ethology and Sociobiology* **14** 89-96

Gangestad S W, Thornhill R A, Yeo R A, 1994 "Facial attractiveness, developmental stability, and fluctuating asymmetry" *Ethology and Sociobiology* **15** 73-85

Gangestad S W, Thornhill R, 1997 "The evolutionary psychology of extrapair sex: The role of fluctuating asymmetry" *Evolution and Human Behavior* **18** 69-88

Going M, Read D, 1974 "Effects of uniqueness, sex of subject and sex of photographs on facial recognition" *Perceptual and Motor Skills* **39** 109-110

Goldman M, Hagen M A, 1978 "The forms of caricature: Physiognomy and political bias" *Studies in the Anthropology of Visual Communication* **5** 30-36

Goldstein A G, 1975 "Recognition of inverted photographs of faces in children and adults" *Journal of Genetic Psychology* **127** 109-123

Goldstein A G, Chance J, 1979 "Do "foreign" faces really look alike?" *Bulletin of the Psychonomic Society* **13** 111-113

Goldstein A G, 1979 "Race-related variation of features: anthropometric data II" *Bulletin of the Psychonomic Society* **13** 191-193

Goldstein A G, 1979a "Facial feature variation: Anthropometric data II" *Bulletin of the Psychonomic Society* **13** 191-193

Goldstein A G, Chance J E, 1980 "Memory for faces and schema theory" *Journal of Psychology* **105** 47-59

Grammer K, Thornhill R, 1994 "Human (Homo Sapiens) facial attractiveness and sexual selection: the role of symmetry and averageness" *Journal of Comparative Psychology* **108** 233-242

Grossman C J, 1985 "Interactions between the gonadal steroids and the immune system" *Science* **227** 257-261

Guthrie R D, 1976 *Body Hot Spots* (New York: Van Nostrand Reinhold)

Hagen M A, Perkins D, 1983 "A refutation of the hypothesis of the superfidelity of caricatures relative to photographs" *Perception* **12** 55-61

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

Hamilton W D, Zuk M, 1982 "Heritable true fitness and bright birds: a role for parasites?" *Science* **218** 384-387

Hancock P J B, Burton A M, Bruce V, 1996 "Face processing: human perception and principal components analysis" *Memory and Cognition* **24** 26-40

Hatfield E, Sprecher S, 1986 *Mirror, Mirror...The Importance of Looks in Everyday Life* (Albany, NY: State University of New York Press)

Henss R, 1991 "Perceiving age and attractiveness and facial photographs" *Journal of Applied Social Psychology* **21** 933-946

Henss R, 1995 "Waist-to-hip ratio and attractiveness. Replication and extension" *Personality and Individual Differences* **19** 479-488

Hess E H, 1965 "Attitude and pupil size" *Scientific American* **212** 46-54

Hess E H, 1975 "The role of pupil size in communication" *Scientific American* **233** 110-119

Hietanen J K, Perrett D I, Oram M W, Benson P J, Dittrich W H, 1992 "The effects of lighting conditions on responses of cells selective for face views in the temporal cortex" *Experimental Brain Research* **89** 157-171

Hill H, Bruce V, Akamatsu S, 1995 "Perceiving the sex and race of faces: the role of shape and colour" *Proceedings of the Royal Society of London Series B* **261** 367-373

Horvath T, 1981 "Physical attractiveness in the selection of certain torso parameters" *Archives of Sexual Behavior* **10** 21-24

Hulse F S, 1967 "Selection for skin color among the Japanese" *American Journal of Physical Anthropology* **27** 143-156

Humphrey G K, Goodale M A, Jakobson L S, Servos P, 1994 "The role of surface information in object recognition: studies of a visual form agnosic and normal subjects" *Perception* **23** 1457-1481

Iliffe A H, 1960 "A study of preferences in feminine beauty" *British Journal of Psychology* **51** 267-273

Jackson L A, 1992 *Physical Appearance and Gender: Sociobiological and Sociocultural Perspectives* (Albany, NY: State University of New York Press)

Jankowiak W R, Hill E M, Donovan J M, 1992 "To effect of sex and sexual orientation on attractiveness judgments: an evolutionary interpretation" *Ethology and Sociobiology* **113** 73-85

Johnson R W, Dannenbring G L, Anderson N R, Villa R E, 1993 "How different cultural and geographic groups perceive the attractiveness of active and inactive feminists" *Journal of Social Psychology* **119** 111-117

Johnston A, Hill H, Carman N, 1992 "Recognising faces: Effects of lighting direction, inversion, and brightness reversal" *Perception* **21** 365-375

Johnston V S, Franklin M, 1993 "Is beauty in the eye of the beholder" *Ethology and Sociobiology* **14** 183-199

Jones D, 1995 "Sexual selection, physical attractiveness, and facial neoteny" *Current Anthropology* **36** 723-748

Kagan J, Klein R F, 1973 "Cross-cultural perspectives on early development" *American Psychologist* **28** 947-961

Kahlon D P S, 1976 "Age variation in skin colour, a study of Sikh immigrants in Britain" *Human Biology* **48** 419-428

Kalla A K, Tiwari S C, 1970 "Sex differences in skin colour in man" *Acta Geneticae Medicae et Gemellologiae* **19** 472-476

Kaye S A, Folsam A R, Prineas R J, Potter J, Gapstur S M, 1990 "The association of body fat distribution with lifestyle and reproductive factors in a population study of post-menopausal women" *International Journal of Obesity* **14** 583-592

Keating C F, Mazur A, Segall M H, 1981 "A cross-cultural exploration of physiognomic traits of dominance and happiness" *Ethology and Sociobiology* **2** 41-48

Keating C, 1985 "Gender and the physiognomy of dominance and attractiveness" *Social Psychology Quarterly* **48** 61-70

Kemp R, McManus C, Pigott T, 1990 "Sensitivity to the displacement of facial features in negative and inverted images" *Perception* **19** 531-543

Kenrick D T, Sadalla E K, Groth G, Thost M R, 1990 "Evolution, traits, and the stages of courtship: Qualifying the parental investment model" *Journal of Personality* **58** 97-116

Klatzky R L, Martin M G, Kane R A, 1982 "Semantic interpretation effects on memory for faces" *Memory and Cognition* **10** 195-206

Klatzky R L, Martin M G, Kane R A, 1982a "Influence of social-category activation on processing of visual information" *Social Cognition* **1** 95-109

Klatzky R L, Forrest F H, 1984 "Recognizing familiar and unfamiliar faces" *Memory and Cognition* **12** 60-70

Kopera A A, Maier R A, Johnson J E, 1971 "Perception of physical attractiveness: the influence of group interaction and group coaction on ratings of the attractiveness of photographs of women" *Proceedings, 79th Annual Convention, APA* 317-318

Korabik K, 1981 "Changes in physical attractiveness and interpersonal attraction" *Basic and Applied Psychology* **2** 59-65

Korthase K M, Trenholme I, 1982 "Perceived age and perceived physical attractiveness" *Perceptual and Motor Skills* **54** 1251-1258

Kowner R, Ogawa T, 1993 "The contrast effect of physical attractiveness in Japan" *The Journal of Psychology* **127** 51-64

Kowner R, 1995 "The effect of physical attractiveness comparison on choice of partners" *Journal of Social Psychology* **135** 153-165

Kowner R, Ogawa T, 1995 "The role of raters sex, personality, and appearance in judgments of facial beauty" *Perceptual and Motor Skills* **423** 363-369

Kowner R, 1996 "Facial asymmetry and attractiveness judgment in developmental perspective" *Journal of Experimental Psychology: Human Perception and Performance* **22** 662-675

Langlois J, 1986 "From the eye of the beholder to behavioral reality: The development of social behaviors and social relations as a function of physical attractiveness", in *Physical appearance, stigma, and social behavior: The Ontario Symposium* Ed M P Z C P Herman, E T Higgins (Hillsdale, NJ: Lawrence Erlbaum) pp 23-51

Langlois J H, Roggman L A, Casey R J, Ritter J M, Rieser-Danner L A, Jenkins V Y, 1987 "Infant preferences for attractive faces: rudiments of stereotype?" *Developmental Psychology* **23** 363-369

Langlois J H, Roggman L A, 1990 "Attractive faces are only average" *Psychological Science* **1** 115-121

Langlois J H, Roggman L A, Reiser-Danner L A, 1990 "Infants' differential social responses to attractive and unattractive faces" *Developmental Psychology* 153-159

Langlois J H, Ritter J M, Roggman L A, Vaughn L S, 1991 "Facial diversity and infant preferences for attractive faces" *Developmental Psychology* **27** 79-84

Langlois J H, Roggman L A, Musselman L, 1994 "What is average and what is not average about attractive faces?" *Psychological Science* **5** 214-219

Leder H, 1996 "Line drawings of faces reduce configural processing" *Perception* **25** 355-366

Lee K J, 1992 *A comparison of inspection time for familiar faces, unfamiliar faces, and simple stimuli* Honours, University of Edinburgh, Edinburgh

Lewis M B, Johnston R A, 1997 "The Thatcher illusion as a test of configural disruption" *Perception* **26** 225-227

Light L L, Kayra-Stuart F, Hollander S, 1979 "Recognition memory for typical and unusual faces" *Journal of Experimental Psychology: Human Learning and Memory* **5** 222-228

Light L L, Hollander S, Kayra-Stuart F, 1981 "Why attractive people are harder to remember" *Personality and Social Psychology Bulletin* **7** 269-276

Livingstone M S, 1988 "Art, illusion and the visual system" *Scientific American* **258** 68-75

Lott D F, 1979 "A possible role for generally adaptive features in mate selection and sexual stimulation" *Psychological Reports* **45** 539-546

Malpass R, Kravitz J, 1969 "Recognition for faces of own and other race" *Journal of Personality and Social Psychology* **13** 330-334

Maret S M, 1983 "Attractiveness ratings of photographs of Blacks by Cruzans and Americans" *The Journal of Psychology* **115** 113-116

Maret S M, Harling C A, 1985 "Cross-cultural perceptions of physical attractiveness: Ratings of photos of Whites by Cruzans and Americans" *Perceptual and Motor Skills* **60** 163-166

Marr D, Nishihara H K, 1978 "Representation and recognition of the spatial organization of three-dimensional shapes" *Proceedings of the Royal Society of London Series B* **200** 269-294

Mauro R, Kubovy M, 1992 "Caricature and face recognition" *Memory and Cognition* **20** 433-440

May K A, Perrett D I, McLeod C, Estachy C, submitted "The effects of caricaturing on recognition of unfamiliar faces"

McAfee L J, Fox R A, Hicks R A, 1982 "Attributions of male college students to variations in facial features in the line drawings of a woman's face" *Bulletin of the Psychonomic Society* **19** 143-144

McArthur L Z, Apatow K, 1983-1984 "Impressions of baby-faced adults" *Social Cognition* **2** 315-342

McArthur L Z, Berry D S, 1987 "Cross-cultural agreements in perceptions of babyfaced adults" *Journal of Cross-Cultural Psychology* **18** 165-192

McGinley H, McGinley P, Nicholas K, 1978 "Smiling, body position, and interpersonal attraction" *Bulletin of the Psychonomic Society* **12** 21-24

McGinley H, Blau G, Takai M, 1984 "Attraction effects of smiling and body position: a cultural comparison" *Perceptual and Motor Skills* **58** 915-922

Morse S J, Reis H, Gruzen J, Wolff E, 1974 "The "eye of the beholder": Determinants of physical attractiveness judgments in the U.S. and South Africa" *Journal of Personality* **42** 528-542

Morse S J, Gruzen J, Reis H, 1976 "The "eye of the beholder": A neglected variable in the study of physical attractiveness?" *Journal of Personality* **44** 209-225

Møller A P, Hoglund J, 1991 "Pattern of flunctuating asymmetry in avian feather ornaments - implications for models of sexual selection" *Proceedings of the Royal Society of London Series B* **245** 1-5

Møller A P, 1992 "Female swallow preferences for symmetrical male sexual ornaments" *Nature* **357** 238-240

Møller A P, 1993 "Patterns of fluctuating asymmetry in sexual ornaments predict female choice" *Journal of Evolutionary Biology* **6** 481-491

Møller A P, Soler M, Thornhill R, 1995 "Breast asymmetry, sexual selection, and human reproductive success" *Ethology and Sociobiology* **16** 207-219

Møller A P, Swaddle J P, 1997 *Developmental stability and evolution* (Oxford: Oxford University Press)

Nakdimen K A, 1984 "The physiognomic basis of sexual stereotyping" *American Journal of Psychiatry* **141** 499-503

Nash H, 1969 "Recognition of body-surface regions" *Genetic Psychology Monographs* **79** 297-340

Nettlebeck T, 1982 "Inspection time: an index for intelligence?" *The Quarterly Journal of Experimental Psychology* **34A** 299-312

Ostergaard A L, Davidoff J B, 1985 "Some effects of color on naming and recognition of objects" *Journal of Experimental Psychology: Learning, Memory, and Cognition* **11** 579-587

Perrett D I, May K A, Yoshikawa S, 1994 "Facial shape and judgements of female attractiveness" *Nature* **368** 239-242

Perrett D I, Burt D M, Lee K J, Rowland D A, Edwards R, submitted "Fluctuating asymmetry in human faces: symmetry is beautiful"

Pinker S, 1985 "Visual Cognition", (Cambridge, MA: MIT Press)

Pittenger J B, 1991 "On the difficulty of averaging faces: Comments on Langlois and Roggman" *Psychological Science* **2** 351-353

Pond C M, 1978 "Morphological aspects and the ecological and mechanical consequences of fat deposition in wild vertebrates" *Annual Review of Ecology and Systematics* **9** 519-570

Price C J, Humphreys G W, 1989 "The effects of surface detail on object categorization and naming" *The Quarterly Journal of Experimental Psychology* **41A** 797-828

Radesäater T, Halldórsdóttir H, 1993 "Fluctuating asymmetry and forceps size in earwigs, Forficula auricuaria" *Animal Behaviour* **45** 626-628

Reis H T, Wheeler L, Spiegl N, Kevnis M H, Nezlek J, Perri M, 1982 "Physical attractiveness in social interaction: II. Why does appearance affect social experience?" *Journal of Personality and Social Psychology* **43** 979-996

Rhodes G, Brennan S, Carey S, 1987 "Identification and ratings of caricatures: Implications for mental representations of faces" *Cognitive Psychology* **19** 473-497

Rhodes G, 1988 "Looking at faces: First-order and second-order features as determinants of facial appearance" *Perception* **17** 43-63

Rhodes G, Tan S, Taylor K, 1989 "Race sensitivity in face recognition: An effect of different encoding processes", in *Cognition in Individual and Social Contexts* Eds A F Bennett and K M McConkey (Amsterdam: North-Holland)

Rhodes G, McLean I G, 1990 "Distinctiveness and expertise effects with homogeneous stimuli: Towards a model of configural coding" *Perception* **19** 773-794

Rhodes G, Tremewan T, 1994 "Understanding face recognition: Caricature effects, inversion, and the homogeneity problem" *Visual Cognition* **1** 275-311

Rhodes G, Byatt G, Tremewan T, Kennedy A, 1996 "Facial distinctiveness and the power of caricatures" *Perception* **25** 207-223

Richardson S A, Goodman N, Hastorf A H, Dornbusch S M, 1961 "Cultural uniformity in reaction to physical disabilities" *American Sociological Review* **26** 241-247

Rolls E T, Baylis C G, 1986 "Size and contrast have only small effects on the responses to faces of neurons in the cortex of the superior temporal sulcus of the macaque monkey" *Experimental Brain Research* **65** 38-48

Rowland D A, Perrett D I, 1995 "Manipulating facial appearance through shape and color" *IEEE Computer Graphics and Applications* **15** 70-76

Rowland D A, 1997 PhD, University of St Andrews, St Andrews

Rowland D A, Perrett D I, Burt D M, Lee K J, Akamatsu S, 1997 "Transforming facial images in 2 and 3-D", in *Imagina* (Monaco) pp 159-176

Ruppert J, 1995 "A delaunay refinement algorithm for quality 2-dimensional mesh generation" *Journal of algorithms* **18** 548-585

Samuels C A, Ewy R, 1985 "Aesthetic perception of faces during infancy" *British Journal of Developmental Psychology* **3** 221-228

Samuels C A, Butterworth G, Roberts T, Graupner L, Hole G, 1994 "Facial aesthetics: babies prefer attractiveness to symmetry" *Perception* **23** 823-831

Scapinello K F, Yarmey A D, 1970 "The role of familiarity and orientation in immediate and delayed recognition of pictorial stimuli" *Psychonomic Science* **21** 329-330

Sergent J, 1984 "An investigation into component and configural processes underlying face perception" *The British Journal of Psychology* **75** 221-242

Shapiro B A, Hazan C, Haith M, 1984 "Do infants differentiate attractiveness and expressiveness in faces?", in *Meeting of the Southwestern Society for Research in Human, Development* (Denver, CO)

Shapiro P N, Penrod S, 1986 "Meta-analysis of facial identification studies" *Psychological Bulletin* **100** 139-156

Shepard R N, 1984 "Ecological constraints on internal representation: Resonant kinematics of perceiving, imagining, thinking, and dreaming" *Psychological Review* **91** 417-447

Shepherd J W, Deregowski J B, Ellis H D, 1974 "A cross-cultural study of recognition memory for faces" *International Journal of Psychology* **9** 205-211

Shepherd J W, Ellis H D, Davies G M, 1977 "Perceiving and Remembering Faces",

Shepherd J W, 1981 "Social factors in face recognition", in *Perceiving and Remembering Faces* Eds G Davies, H Ellis and J Shepherd (London: Lawrence Erlbaum) pp 55-79

Singh D, 1993 "Body shape and women's attractiveness - the critical role of waist-to-hip ratio" *Human Nature* **4** 297-321

Singh D, 1994 "Waist-to-hip ratio and judgment of attractiveness and healthiness of female figures by male and female physicians" *International Journal of Obesity* **18** 731-737

Singh D, 1995 "Female health, attractiveness, and desirability for relationships: role of breast asymmetry and waist-to-hip ratio" *Personality and Social Psychology* **65** 293-307

Singh D, Luis S, 1995 "Ethnic and gender consensus for the effect of waist-to-hip ratio judgment of women's attractiveness" *Human Nature* **6** 51-65

Snedecor G W, Cochran W G, 1980 *Statistical Methods* (Ames, Iowa: The Iowa State University Press)

Solso R L, McCarthy J E, 1981 "Prototype formation of faces: A case study of pseudo-memory" *British Journal of Psychology* **72** 499-503

Soule M E, Cuzin-Roudy J, 1982 "Allometric Variation .2. Developmental stability of extreme phenotypes" *American Naturalist* **120** 765-786

Sternberg S, 1969 "The discovery of processing stages: Extensions of Donder's method" *Acta Psychologica* **30** 276-315

Stevenage S V, 1995 "Expertise and the caricature advantage", in *Cogntive and Computational Aspects of Face Recognition* Ed T Valentine (London: Routledge) pp 24-46

Stroebe W, Insko C A, Thompson V D, Layton B D, 1971 "Effects of physical attractiveness, attitude similarity, and sex on various aspects of interpersonal attraction" *Journal of Personality and Social Psychology* **18** 79-91

Swaddle J P, Cuthill I C, 1994 "Preference for symmetric males by Zebra finches" *Nature* **367** 165-166

Swaddle J P, Cuthill I C, 1995 "Asymmetry and human facial attractiveness: symmetry may not always be beautiful" *Proceedings of the Royal Society of London, Series B* **261** 111-116

Symons D, 1979 *The evolution of human sexuality* (Oxford: Oxford University Press)

Tanaka J W, Farah M J, 1991 "Second-order relational properties and the inversion effect: Testing a theory of face perception" *Perception and Psychophysics* **50** 367-372

Tanaka J W, Simon V B, 1996 "Caricature recognition in a neural network" *Visual Cognition* **3** 301-

Tanner J M, 1978 *Education and physical growth: implications of the study of children's growth for educational theory and practice* (London: Hodder and Stoughton)

Terry R, 1977 "Further evidence on components of facial attractiveness" *Perceptual and Motor Skills* **45** 130

Thakerar J N, Iwawaki S, 1979 "Cross-cultural comparisons in interpersonal attraction of females toward males" *Journal of Social Psychology* **108** 121-122

Thompson P, 1980 "Margaret Thatcher: a new illusion" *Perception* **9** 483-484

Thornhill R, Gangestad S W, 1993 "Human facial beauty: averageness, symmetry, and parasite resistance" *Human Nature* **4** 237-269

Thornhill R, Gangestad S W, 1994 "Human fluctating asymmetry and sexual behavior" *Psychological Science* **5** 297-302

Townsend J M, 1989 "Mate selection criteria - a pilot-study" *Ethology and Sociobiology* **10** 241-253

Townsend J M, Levy G D, 1990 "Effects of potential partners' physical attractiveness and socioeconomic status on sexuality and partner selection" *Archives of Sexual Behavior* **19** 149-164

Townsend J M, 1993 "Sexuality and partner selection: sex differences among college students" *Ethology and Sociobiology* **14** 305-329

Townsend J M, Kline J, Wasserman T H, 1995 "Low investment copulation: Sex differences in motivations and emotional-reactions" *Ethology and Sociobiology* **16** 25-51

Trivers R L, 1972 "Parental investment and sexual selection", in *Sexual Selection and the Descent of Man:1871-1971* Ed B Campbell (Chicago: Aldine) pp 136-179

Trivers R, 1985 *Social Evolution* (Menlo Park, California: Benjamin/Cummings)

Tversky B, Baratz D, 1985 "Memory for faces: Are caricatures better than photographs?" *Memory and Cognition* **13** 45-49

Udry J R, 1965 "Structural correlates of feminine beauty preferences in Britain and the United States: A comparison" *Sociology and Social Research* **49** 330-342

Valentine T, Bruce V, 1986 "Recognizing familiar faces: The role of distinctiveness and familiarity" *Canadian Journal of Psychology* **40** 300-305

Valentine T, Bruce V, 1986a "The effects of distinctiveness in recognising and classifying faces" *Perception* **15** 525-535

Valentine T, Bruce V, 1986b "The effect of race, inversion and encoding activity upon face recognition" *Acta Psychologica* **61** 259-273

Valentine T, 1988 "Upside-down faces: A review of the effect of inversion upon face recognition" *British Journal of Psychology* **79** 471-491

Valentine T, 1991 "A unified account of the effects of distinctiveness, inversion, and race in face recognition" *The Quarterly Journal of Experimental Psychology* **43A** 161-204

Valentine T, Endo M, 1992 "Towards an exemplar model of face processing: The effects of race and distinctiveness" *The Quarterly Journal of Experimental Psychology* **44A** 671-703 van den Berghe P L, Frost P, 1986 “Skin color preference, sexual dimorphism and sexual selection: a case gene-culture co-evolution?” *Ethnic and Racial Studies* **9** 87-113

Van Valen L, 1962 “A study of fluctuating asymmetry” *Evolution* **16** 125-142

Vickers D, Smith P L, 1986 “The rationale for the inspection time index” *Personality and Individual Differences* **7** 609-623

Wagatsuma E, Kleinke C L, 1979 “Ratings of facial beauty by Asian-American and Caucasian females” *Journal of Social Psychology* **109** 299-300

Walster E, Aronson V, Abrahams D, Rottmann L, 1966 “Importance of physical attractiveness in dating behavior” *Journal of Personality and Social Psychology* **4** 508-516

Watson P J, Thornhill R, 1994 “Fluctuating asymmetry and sexual selection” *Trends in Ecology and Evolution* **9** 21-25

Wedekind C, 1992 “Detailed information about parasites revealed by sexual ornamentation” *Proceedings of the Royal Society of London Series B* **247** 169-174

Weisfeld G E, Weisfeld C C, Callaghan J W, 1984 “Peer and self perceptions in Hopi and Afro-American third- and sixth-graders” *Ethos* **12** 64-84

Wheller L, Miyake K, 1992 “Social comparison in everyday life” *Journal of Persnality and Social Psychology* **62** 360-373

Winograd E, 1981 “Elaboration and distinctiveness in memory for faces” *Journal of Experimental Psychology: Human Learning and Memory* **7** 181-190

Wurm L H, Legge G E, Isenberg L M, Luebker A, 1993 “Color improves object recognition in normal and low vision” *Journal of Experimental Psychology: Human Perception and Performance* **19** 899-911

Yamane S, Kaji S, Komatsu, 1988 “What facial features activate face neurons in the inferotemporal cortex of the monkey” *Experimental Brain Research* **73** 209-214

Yarmey A D, 1971 “Recognition memory for familiar "public" faces: Effects of orientation and delay” *Psychonomic Science* **24** 286-288

Yin R K, 1969 “Looking at upside-down faces” *Journal of Experimental Psychology* **81** 141-145

Young A W, Hay D C, McWeeny K H, Flude B M, Ellis A W, 1985 "Matching familiar and unfamiliar faces on internal and external features" *Perception* **14** 737-746

Zebrowitz L A, Montepare J M, Lee H K, 1993 "They don't all look alike: Individual impressions of other racial groups" *Journal of Personality and Social Psychology* **65** 85-101

Zebrowitz L A, Olson K, Hoffman K, 1993 "Stability of babyfaceness and attractiveness across the life span" *Journal of Personality and Social Psychology* **64** 453-466

Appendix I - Famous Faces Used as Test Stimuli

Experiment 1

Arnold Schwarzenegger
Harrison Ford
Jodie Foster
*Keanu Reeves
Kim Basinger
*Kylie Minogue
Madonna
Roger Moore
*Sean Connery
Sharon Stone

Experiment 2

Andie McDowell
Arnold Schwarzenegger
Emma Thompson
Harrison Ford
Jodie Foster
Madonna
Mel Gibson
Roger Moore
Rowan Atkinson
Sigourney Weaver

Experiment 3

Arnold Schwarzenegger
Harrison Ford
Jodie Foster
*Keanu Reeves
Kim Basinger
*Kylie Minogue
Madonna
Roger Moore
*Sean Connery
Sharon Stone

Experiment 4

Al Pacino
Arnold Schwarzenegger
*Gregor Fisher
Kevin Costner
Luke Perry
Roger Moore
Sean Connery
Steve Martin
Ted Danson
Val Kilmer

Experiment 5

Arnold Schwarzenegger
Bob Geldof
Clint Eastwood
*Princess Diana
Emma Thompson
Harrison Ford
Jodie Foster
Kim Basinger
Madonna
Mel Gibson
Norman Lamont
Roger Moore
Rowan Atkinson
Sharon Stone
*Sigourney Weaver

Experiment 6

Kevin Costner
Rober Deniro
Harrison Ford
Al Pacino
Christian Slater
Bruce Willis

Experiment 7

Pierce Brosnan
Johnny Depp
Michael J. Fox
Brad Pitt
Keanu Reeves
*Christopher Walken

*denotes that the face was removed from further analysis due to poor representativeness of the image

Appendix II - Feature Delineation Points

Feature delineation points from Benson and Perrett (1991, 1991a, 1991b, 1994).

No. unique points	feature
1	left pupil
1	right pupil
4	left iris
4	right iris
3	bottom of left eyelid
3	bottom of right eyelid
6	left eye
6	right eye
3	left eye line
3	right eye line
20	nose, including nostrils
8	left eye brow
8	right eye brow
22	lips (mouth always open)
5	left ear
5	right ear
26	outline of face
13	top of head (hair)
3	left smile line
3	right smile line
3	left cheekbones
3	right cheekbones
2	left upper lip line
2	right upper lip line
2	chin cleft
3	chin line
162	**Total**

Feature delineation points used to describe images in this thesis

No. unique points	feature
1	left pupil
1	right pupil
8	left iris
8	right iris
5	bottom of left eyelid
5	bottom of right eyelid
8	left eye
8	right eye
3	left eye line
3	right eye line
21	nose, including nostrils
8	left eye brow
8	right eye brow
22	lips (mouth open; 17 mouth closed)
5	left ear
5	right ear
26	outline of face
13	top of head (hair)
3	left smile line
3	right smile line
3	left cheekbones
3	right cheekbones
2	left upper lip line
2	right upper lip line
2	chin cleft
3	chin line
179 (174 mouth closed)	**Total**

Appendix III - Expanded Data Analysis

ANALYSIS OF MASK 1 VERDICAL, SHAPE CARICATURE 2-WAY ANOVA

Number of cells = 88 Number of data points = 88

STRUCTURE of ANOVA
Structure Parameters : DURATION=4 CARIC=2 SUBJECTS=-11

1	DURATION	4 levels.	FIXED
2	CARIC	2 levels.	FIXED
3	SUBJECTS	11 levels.	RANDOM

RESULTS OF ANALYSIS

Type	Source	df	SS	MS	F	p	MSE	df
MAIN EFFECT								
1 Fixed	DURATION :	3	5.91072649	1.97024216	46.934	0.0000	0.041979	30
2 Fixed	CARIC :	1	0.00003329	0.00003329	0.002	0.9663	0.017721	10
3 Random	SUBJECTS :	10	1.51566427	0.15156643	5.326	0.0002	0.028458	30
INTERACTIONS								
4:	1 2 :	3	0.29417687	0.09805896	3.446	0.0290	0.028458	30
5:	1 3 :	30	1.25937606	0.04197920	1.475	0.1463	0.028458	30
6:	2 3 :	10	0.17721424	0.01772142	0.623	0.7825	0.028458	30
7:	1 2 3 :	30	0.85374110	0.02845804				
	TOTAL	87	10.01093231					

EXPECTED MEAN SQUARES

	Source	EMS
1:	DURATION	2.7(5) + 7.3[1] + var(err)
2:	CARIC	8.0(6) + 44.0[2] + var(err)
3:	SUBJECTS	8.0(3) + var(err)
Interactions		
4:	1 2	3.7[4] + var(err)
5:	1 3	2.7(5) + var(err)
6:	2 3	8.0(6) + var(err)
7:	1 2 3	var(err)

CLASS SUMMARIES

ANALYSIS OF MASK 1 VERDICAL, SHAPE CARICATURE 2-WAY ANOVA

FACTOR: DURATION (between)
Main effect at p = 0.0000 (df=30)

	MEAN	VARIANCE	SD	S.E.M.	N
Class 1	0.4183	0.0328	0.1810	0.0386	22
Class 2	0.7248	0.0661	0.2571	0.0548	22
Class 3	1.0050	0.0569	0.2385	0.0508	22
Class 4	1.0746	0.0395	0.1988	0.0424	22

Pooled S.E.M. = 0.04 22.00 oberservations per class (harmonic mean).
Testing Differences (p = 0.05) PLSD = 0.1262
Bonferroni correction gives PLSD = 0.1745
NK = 0.13/0.15/0.17

CLASS SUMMARIES

ANALYSIS OF MASK 1 VERDICAL, SHAPE CARICATURE 2-WAY ANOVA
FACTOR: CARIC (between)
Non-significant main effect (p=0.9663)

	MEAN	VARIANCE	SD	S.E.M.	N
Class 1	0.8063	0.1476	0.3842	0.0579	44
Class 2	0.8051	0.0852	0.2919	0.0440	44

INTERACTIONS of MASK 1 VERDICAL, SHAPE CARICATURE 2-WAY ANOVA

DURATION = [1]   CARIC = [2]
Interaction significant at p = 0.0290
Means

	[2].1	[2].2
[1].1	0.330	0.507
[1].2	0.718	0.731
[1].3	1.073	0.937
[1].4	1.104	1.046

S.E. of difference between means = 0.0509
11.00 observations per cell (harmonic mean)
Testing (p = 0.050) PLSD = 0.1469
NK = 0.1470

ANALYSIS OF MALE and FEMALE Stim, COLOUR, GREY 3-WAY ANOVA

Number of cells = 144 Number of data points = 144

STRUCTURE of ANOVA

Structure Parameters : STIMGENDER=2 DURATION=3 COLOUR=2 SUBJECTS=12

1	STIMGENDER	2 levels.	FIXED
2	DURATION	3 levels.	FIXED
3	COLOUR	2 levels.	FIXED
4	SUBJECTS	12 levels.	RANDOM

RESULTS OF ANALYSIS

Type	Source	df	SS	MS	F	p	MSE	df
MAIN EFFECT								
1 Fixed	STIMGENDER :	1	0.01694147	0.01694147	0.475	0.5051	0.035680	11
2 Fixed	DURATION :	2	7.96006738	3.98003369	29.856	0.0000	0.133306	22
3 Fixed	COLOUR :	1	0.06189380	0.06189380	5.289	0.0420	0.011702	11
4 Random	SUBJECTS :	11	5.89569023	0.53597184	12.003	0.0000	0.044653	22
INTERACTIONS								
5:	1 2 :	2	0.00605139	0.00302569	0.064	0.9385	0.047525	22
6:	1 3 :	1	0.00028156	0.00028156	0.009	0.9265	0.031574	11
7:	1 4 :	11	0.39247805	0.03567982	0.799	0.6401	0.044653	22
8:	2 3 :	2	0.01241238	0.00620619	0.205	0.8164	0.030320	22
9:	2 4 :	22	2.93272425	0.13330565	2.985	0.0066	0.044653	22
10:	3 4 :	11	0.12872446	0.01170222	0.262	0.9872	0.044653	22
11:	1 2 3 :	2	0.06549693	0.03274847	0.733	0.4916	0.044653	22
12:	1 2 4 :	22	1.04555155	0.04752507	1.064	0.4425	0.044653	22
13:	1 3 4 :	11	0.34731740	0.03157431	0.707	0.7194	0.044653	22
14:	2 3 4 :	22	0.66705063	0.03032048	0.679	0.8146	0.044653	22
15:	1 2 3 4 :	22	0.98235883	0.04465267				
	TOTAL	143	20.51504031					

EXPECTED MEAN SQUARES

	Source				EMS
1:	STIMGENDER				12.0(7) + 72.0[1] + var(err)
2:	DURATION				6.0(9) + 24.0[2] + var(err)
3:	COLOUR				12.0(10) + 72.0[3] + var(err)
4:	SUBJECTS				12.0(4) + var(err)
Interactions					
5:			1	2	6.0(12) + 12.0[5] + var(err)
6:			1	3	12.0(13) + 36.0[6] + var(err)
7:			1	4	12.0(7) + var(err)
8:			2	3	6.0(14) + 12.0[8] + var(err)
9:			2	4	6.0(9) + var(err)
10:			3	4	12.0(10) + var(err)
11:		1	2	3	6.0[11] + var(err)
12:		1	2	4	6.0(12) + var(err)
13:		1	3	4	12.0(13) + var(err)
14:		2	3	4	6.0(14) + var(err)
15:	1	2	3	4	var(err)

CLASS SUMMARIES

ANALYSIS OF MALE and FEMALE Stim, COLOUR, GREY 3-WAY ANOVA
FACTOR: STIMGENDER (between)
Non-significant main effect (p=0.5051)

	MEAN	VARIANCE	SD	S.E.M.	N
Class 1	0.8354	0.1474	0.3840	0.0453	72
Class 2	0.8137	0.1413	0.3759	0.0443	72

CLASS SUMMARIES

ANALYSIS OF MALE and FEMALE Stim, COLOUR, GREY 3-WAY ANOVA

FACTOR: DURATION (between)
Main effect at p = 0.0000 (df=22)

	MEAN	VARIANCE	SD	S.E.M.	N
Class 1	0.4996	0.0769	0.2773	0.0400	48
Class 2	0.9258	0.1064	0.3262	0.0471	48
Class 3	1.0482	0.0838	0.2895	0.0418	48

Pooled S.E.M. = 0.05 48.00 oberservations per class (harmonic mean).
Testing Differences (p = 0.05) PLSD = 0.1546
Bonferroni correction gives PLSD = 0.1931
NK = 0.15/0.19

CLASS SUMMARIES

ANALYSIS OF MALE and FEMALE Stim, COLOUR, GREY 3-WAY ANOVA

FACTOR: COLOUR (between)
Main effect at p = 0.0420 (df=11)

	MEAN	VARIANCE	SD	S.E.M.	N
Class 1	0.8452	0.1470	0.3834	0.0452	72
Class 2	0.8038	0.1411	0.3756	0.0443	72

Pooled S.E.M. = 0.01 72.00 oberservations per class (harmonic mean).
Testing Differences (p = 0.05) PLSD = 0.0397
Bonferroni correction gives PLSD = 0.0397
NK = 0.04

ANALYSIS OF 2-WAY ANOVA COLOUR VS. GREY-SCALE FOR RECOGNITION BY ITEM

Number of cells = 60 Number of data points = 60

STRUCTURE of ANOVA
Structure Parameters : DURATION=3 COLOUR=2 ITEM=10

1	DURATION	3 levels.	FIXED
2	COLOUR	2 levels.	FIXED
3	ITEM	10 levels.	FIXED

RESULTS OF ANALYSIS

Type	Source	df	SS	MS	F	p	MSE	df
MAIN EFFECT								
1 Fixed	DURATION :	2	3.391951	1.6959755	69.397	0.0000	0.024439	18
2 Fixed	COLOUR :	1	0.022030	0.0220304	0.901	0.3550	0.024439	18
3 Fixed	ITEM :	9	1.488659	0.1654065	6.768	0.0003	0.024439	18
INTERACTIONS								
4:	1 2 :	2	0.016732	0.0083660	0.342	0.7146	0.024439	18
5:	1 3 :	18	0.326446	0.0181359	0.742	0.7333	0.024439	18
6:	2 3 :	9	0.058077	0.0064530	0.264	0.9768	0.024439	18
7:	1 2 3 :	18	0.439899	0.0244389				
	TOTAL	59	5.743794					

EXPECTED MEAN SQUARES

	Source	EMS
1:	DURATION	10.0[1] + var(err)
2:	COLOUR	30.0[2] + var(err)
3:	ITEM	0.7[3] + var(err)
Interactions		
4:	1 2	5.0[4] + var(err)
5:	1 3	0.1[5] + var(err)
6:	2 3	0.3[6] + var(err)
7:	1 2 3	var(err)

CLASS SUMMARIES

ANALYSIS OF 2-WAY ANOVA COLOUR VS. GREY-SCALE FOR RECOGNITION BY ITEM

FACTOR: DURATION (between)
Main effect at p = 0.0000 (df=18)

	MEAN	VARIANCE	SD	S.E.M.	N
Class 1	0.4980	0.0273	0.1652	0.0369	20
Class 2	0.9295	0.0578	0.2405	0.0538	20
Class 3	1.0525	0.0387	0.1966	0.0440	20

Pooled S.E.M. = 0.03 20.00 oberservations per class (harmonic mean).
Testing Differences (p = 0.05) PLSD = 0.1039
Bonferroni correction gives PLSD = 0.1305
NK = 0.10/0.13

CLASS SUMMARIES

ANALYSIS OF 2-WAY ANOVA COLOUR VS. GREY-SCALE FOR RECOGNITION BY ITEM
FACTOR: COLOUR (between)
Non-significant main effect (p=0.3550)

	MEAN	VARIANCE	SD	S.E.M.	N
Class 1	0.8458	0.1046	0.3235	0.0591	30
Class 2	0.8075	0.0927	0.3044	0.0556	30

CLASS SUMMARIES

ANALYSIS OF 2-WAY ANOVA COLOUR VS. GREY-SCALE FOR RECOGNITION BY ITEM

FACTOR: ITEM (between)
Main effect at p = 0.0003 (df=18)

	MEAN	VARIANCE	SD	S.E.M.	N
Class 1	0.8012	0.0609	0.2468	0.1008	6
Class 2	0.6364	0.0626	0.2502	0.1021	6
Class 3	0.7938	0.0930	0.3049	0.1245	6
Class 4	1.0842	0.1472	0.3837	0.1566	6
Class 5	0.8717	0.0710	0.2664	0.1088	6
Class 6	0.8095	0.0461	0.2146	0.0876	6
Class 7	0.5989	0.0471	0.2170	0.0886	6
Class 8	0.9455	0.1380	0.3715	0.1516	6
Class 9	1.0536	0.1047	0.3236	0.1321	6
Class 10	0.6718	0.0805	0.2838	0.1159	6

Pooled S.E.M. = 0.06 6.00 oberservations per class (harmonic mean).
Testing Differences (p = 0.05) PLSD = 0.1896
Bonferroni correction gives PLSD = 0.3497
NK = 0.19/0.23/0.26/0.27/0.29/0.30/0.31/0.32/0.32

ANALYSIS OF MASKS 1 AND 2 VERIDICAL, COLOUR, GREY 2-WAY ANOVA

Number of cells = 336 Number of data points = 336

STRUCTURE of ANOVA

Structure Parameters : MASKS=2 (SUBJECTS)=14 DURATION=4 CARIC=3

1	MASKS	2 levels. FIXED	
2	SUBJECTS	14 levels. RANDOM	NESTED WITHIN FACTOR(s) 1
3	DURATION	4 levels. FIXED	
4	CARIC	3 levels. FIXED	

RESULTS OF ANALYSIS

Type	Source	df	SS	MS	F	p	MSE	df
MAIN EFFECT								
1 Fixed	MASKS :	1	3.179769	3.179769	12.382	0.0016	0.256798	26
2 Random	(SUBJECTS	)	*** NESTED - See	interaction	1 (2) ***			
3 Fixed	DURATION :	3	21.221926	7.073975	132.536	0.0000	0.053374	78
4 Fixed	CARIC :	2	0.162467	0.081234	4.578	0.0148	0.017746	52
INTERACTIONS								
5:	1 (2):	26	6.676738	0.256798	12.531	0.0000	0.020494	156
6:	1 3 :	3	0.947293	0.315764	5.916	0.0011	0.053374	78
7:	1 4 :	2	0.055314	0.027657	1.558	0.2201	0.017746	52
10:	3 4 :	6	0.280392	0.046732	2.280	0.0388	0.020494	156
11:	1 (2) 3 :	78	4.163156	0.053374	2.604	0.0000	0.020494	156
12:	1 (2) 4 :	52	0.922787	0.017746	0.866	0.7219	0.020494	156
13:	1 3 4 :	6	0.083112	0.013852	0.676	0.6693	0.020494	156
15:	1 (2) 3 4 :	156	3.197016	0.020494				
	TOTAL	335	40.889971					

EXPECTED MEAN SQUARES

	Source	EMS
1:	MASKS	12.0(5) + 168.0[1] + var(err)
2:	SUBJECTS	NESTED - See interaction
3:	DURATION	4.0(11) + 28.0[3] + var(err)
4:	CARIC	6.0(12) + 56.0[4] + var(err)
Interactions		
5:	1 (2)	12.0(5) + var(err)
6:	1 3	4.0(11) + 14.0[6] + var(err)
7:	1 4	6.0(12) + 28.0[7] + var(err)
10:	3 4	4.7[10] + var(err)
11:	1 (2) 3	4.0(11) + var(err)
12:	1 (2) 4	6.0(12) + var(err)
13:	1 3 4	2.3[13] + var(err)
15:	1 (2) 3 4	var(err)

CLASS SUMMARIES

ANALYSIS OF MASKS 1 AND 2 VERIDICAL, COLOUR, GREY 2-WAY ANOVA

FACTOR: MASKS (between)
Main effect at p = 0.0016 (df=26)

	MEAN	VARIANCE	SD	S.E.M.	N
Class 1	0.8800	0.1201	0.3465	0.0267	168
Class 2	0.6855	0.1057	0.3252	0.0251	168

Pooled S.E.M. = 0.04 168.00 oberservations per class (harmonic mean).
Testing Differences (p = 0.05) PLSD = 0.1137
Bonferroni correction gives PLSD = 0.1137
NK = 0.11

CLASS SUMMARIES

ANALYSIS OF MASKS 1 AND 2 VERIDICAL, COLOUR, GREY 2-WAY ANOVA
FACTOR: SUBJECTS (nested within MASKS) - see corresponding interaction
CLASS SUMMARIES

ANALYSIS OF MASKS 1 AND 2 VERIDICAL, COLOUR, GREY 2-WAY ANOVA

FACTOR: DURATION (between)
Main effect at p = 0.0000 (df=78)

	MEAN	VARIANCE	SD	S.E.M.	N
Class 1	0.4255	0.0295	0.1717	0.0187	84
Class 2	0.6989	0.0803	0.2835	0.0309	84
Class 3	0.9014	0.0814	0.2853	0.0311	84
Class 4	1.1051	0.0458	0.2139	0.0233	84

Pooled S.E.M. = 0.03 84.00 oberservations per class (harmonic mean).
Testing Differences (p = 0.05) PLSD = 0.0710
Bonferroni correction gives PLSD = 0.0965
NK = 0.07/0.09/0.09

CLASS SUMMARIES

ANALYSIS OF MASKS 1 AND 2 VERIDICAL, COLOUR, GREY 2-WAY ANOVA

FACTOR: CARIC (between)
Main effect at p = 0.0148 (df=52)

	MEAN	VARIANCE	SD	S.E.M.	N
Class 1	0.7662	0.1212	0.3481	0.0329	112
Class 2	0.8138	0.1180	0.3435	0.0325	112
Class 3	0.7682	0.1277	0.3574	0.0338	112

Pooled S.E.M. = 0.01 112.00 oberservations per class (harmonic mean).
Testing Differences (p = 0.05) PLSD = 0.0357
Bonferroni correction gives PLSD = 0.0440
NK = 0.04/0.04

INTERACTIONS of MASKS 1 AND 2 VERIDICAL, COLOUR, GREY 2-WAY ANOVA

MASKS = [1] DURATION = [3]
Interaction significant at p = 0.0011
Means

	[3].1	[3].2	[3].3	[3].4
[1].1	0.460	0.856	1.043	1.161
[1].2	0.391	0.542	0.759	1.049

S.E. of difference between means = 0.0356
42.00 observations per cell (harmonic mean)
Testing (p = 0.050) PLSD = 0.1004
NK = 0.1006

DURATION = [3] CARIC = [4]
Interaction significant at p = 0.0388
Means

	[4].1	[4].2	[4].3
[3].1	0.446	0.422	0.409
[3].2	0.653	0.758	0.686
[3].3	0.875	0.977	0.853
[3].4	1.091	1.099	1.125

S.E. of difference between means = 0.0271
28.00 observations per cell (harmonic mean)
Testing (p = 0.050) PLSD = 0.0756
NK = 0.0758

ANALYSIS OF MASK 1 AND 2 VERIDICAL, COLOUR, CONTRAST 2-WAY ANOVA

Number of cells = 168 Number of data points = 168

STRUCTURE of ANOVA
Structure Parameters : MASKS=2 (ITEMS)=7 DURATION=4 CARIC=3

1	MASKS	2 levels. FIXED	
2	ITEMS	7 levels. FIXED	NESTED WITHIN FACTOR(s) 1
3	DURATION	4 levels. FIXED	
4	CARIC	3 levels. FIXED	

RESULTS OF ANALYSIS

Type	Source	df	SS	MS	F	p	MSE	df
MAIN EFFECT								
1 Fixed	MASKS :	1	1.568689	1.5686891	8.931	0.0113	0.175648	12
2 Fixed	(ITEMS)	*** NESTED - See interaction 1 (2) ***						
3 Fixed	DURATION :	3	10.970770	3.6569233	186.937	0.0000	0.019562	72
4 Fixed	CARIC :	2	0.086509	0.0432547	2.211	0.1170	0.019562	72
INTERACTIONS								
5:	1 (2):	12	2.107780	0.1756483	8.979	0.0000	0.019562	72
6:	1 3 :	3	0.442151	0.1473837	7.534	0.0002	0.019562	72
7:	1 4 :	2	0.028378	0.0141888	0.725	0.4877	0.019562	72
10:	3 4 :	6	0.114744	0.0191240	0.978	0.4468	0.019562	72
11:	1 (2) 3 :	36	0.890745	0.0247429	1.265	0.1971	0.019562	72
12:	1 (2) 4 :	24	0.565017	0.0235424	1.203	0.2688	0.019562	72
13:	1 3 4 :	6	0.039498	0.0065831	0.337	0.9154	0.019562	72
15:	1 (2) 3 4 :	72	1.408489	0.0195623				
	TOTAL	167	18.222771					

EXPECTED MEAN SQUARES

	Source	EMS
1:	MASKS	1.0[5] + 84.0[1] + var(err)
2:	ITEMS	NESTED - See interaction
3:	DURATION	14.0[3] + var(err)
4:	CARIC	28.0[4] + var(err)
Interactions		
5:	1 (2)	1.0[5] + var(err)
6:	1 3	7.0[6] + var(err)
7:	1 4	14.0[7] + var(err)
10:	3 4	2.3[10] + var(err)
11:	1 (2) 3	0.3[11] + var(err)
12:	1 (2) 4	0.5[12] + var(err)
13:	1 3 4	1.2[13] + var(err)
15:	1 (2) 3 4	var(err)

CLASS SUMMARIES

ANALYSIS OF MASK 1 AND 2 VERIDICAL, COLOUR, CONTRAST 2-WAY ANOVA

FACTOR: MASKS (between)
Main effect at p = 0.0113 (df=12)

	MEAN	VARIANCE	SD	S.E.M.	N
Class 1	0.8782	0.1055	0.3249	0.0354	84
Class 2	0.6849	0.0951	0.3084	0.0336	84

Pooled S.E.M. = 0.05 84.00 oberservations per class (harmonic mean).
Testing Differences (p = 0.05) PLSD = 0.1409
Bonferroni correction gives PLSD = 0.1409
NK = 0.14

CLASS SUMMARIES

ANALYSIS OF MASK 1 AND 2 VERIDICAL, COLOUR, CONTRAST 2-WAY ANOVA
FACTOR: ITEMS (nested within MASKS) - see corresponding interaction
CLASS SUMMARIES

ANALYSIS OF MASK 1 AND 2 VERIDICAL, COLOUR, CONTRAST 2-WAY ANOVA

FACTOR: DURATION (between)
Main effect at p = 0.0000 (df=72)

	MEAN	VARIANCE	SD	S.E.M.	N
Class 1	0.4181	0.0273	0.1652	0.0255	42
Class 2	0.6958	0.0557	0.2361	0.0364	42
Class 3	0.9041	0.0611	0.2471	0.0381	42
Class 4	1.1082	0.0328	0.1810	0.0279	42

Pooled S.E.M. = 0.02 42.00 oberservations per class (harmonic mean).
Testing Differences (p = 0.05) PLSD = 0.0608
Bonferroni correction gives PLSD = 0.0828
NK = 0.06/0.07/0.08

INTERACTIONS of MASK 1 AND 2 VERIDICAL, COLOUR, CONTRAST 2-WAY ANOVA

MASKS = [1] ITEMS = [2]
Interaction significant at p = 0.0000
Means

	[2].1	[2].2	[2].3	[2].4	[2].5	[2].6	[2].7
[1].1	0.866	0.940	0.878	0.702	0.953	1.085	0.723
[1].2	0.704	0.637	0.599	0.622	0.730	0.898	0.605

S.E. of difference between means = 0.0404
12.00 observations per cell (harmonic mean)
Testing (p = 0.050) PLSD = 0.1138
NK = 0.1140

MASKS = [1] DURATION = [3]
Interaction significant at p = 0.0002
Means

	[3].1	[3].2	[3].3	[3].4
[1].1	0.456	0.847	1.048	1.162
[1].2	0.381	0.544	0.761	1.054

S.E. of difference between means = 0.0305
21.00 observations per cell (harmonic mean)
Testing (p = 0.050) PLSD = 0.0860
NK = 0.0862

ANALYSIS OF CARICATURE, CONTRAST CONTROL, VERIDICAL 3-WAY ANOVA

Number of cells = 210 Number of data points = 210

STRUCTURE of ANOVA
Structure Parameters : DURATION=2 CARIC=3 SUBJECTS=35

1 DURATION 2 levels. FIXED
2 CARIC 3 levels. FIXED
3 SUBJECTS 35 levels. RANDOM

RESULTS OF ANALYSIS

Type	Source	df	SS	MS	F	p	MSE	df
MAIN EFFECT								
1 Fixed	DURATION :	1	0.667488	0.6674878	11.977	0.0015	0.055731	34
2 Fixed	CARIC :	2	0.183363	0.0916816	6.560	0.0025	0.013976	68
3 Random	SUBJECTS :	34	13.916675	0.4093140	20.843	0.0000	0.019638	68
INTERACTIONS								
4:	1 2 :	2	0.016419	0.0082097	0.418	0.6600	0.019638	68
5:	1 3 :	34	1.894855	0.0557310	2.838	0.0001	0.019638	68
6:	2 3 :	68	0.950370	0.0139760	0.712	0.9183	0.019638	68
7:	1 2 3 :	68	1.335366	0.0196377				
	TOTAL	209	18.964537					

EXPECTED MEAN SQUARES

	Source	EMS
1:	DURATION	6.0(5) + 105.0[1] + var(err)
2:	CARIC	3.0(6) + 35.0[2] + var(err)
3:	SUBJECTS	6.0(3) + var(err)
Interactions		
4:	1 2	17.5[4] + var(err)
5:	1 3	6.0(5) + var(err)
6:	2 3	3.0(6) + var(err)
7:	1 2 3	var(err)

CLASS SUMMARIES

ANALYSIS OF CARICATURE, CONTRAST CONTROL, VERIDICAL 3-WAY ANOVA

FACTOR: DURATION (between)
Main effect at p = 0.0015 (df=34)

	MEAN	VARIANCE	SD	S.E.M.	N
Class 1	0.7058	0.1013	0.3182	0.0311	105
Class 2	0.8185	0.0747	0.2733	0.0267	105

Pooled S.E.M. = 0.02 105.00 oberservations per class (harmonic mean).
Testing Differences (p = 0.05) PLSD = 0.0662
Bonferroni correction gives PLSD = 0.0662
NK = 0.07

CLASS SUMMARIES

ANALYSIS OF CARICATURE, CONTRAST CONTROL, VERIDICAL 3-WAY ANOVA

FACTOR: CARIC (between)
Main effect at p = 0.0025 (df=68)

	MEAN	VARIANCE	SD	S.E.M.	N
Class 1	0.8037	0.0963	0.3102	0.0371	70
Class 2	0.7455	0.0903	0.3004	0.0359	70
Class 3	0.7372	0.0857	0.2927	0.0350	70

Pooled S.E.M. = 0.01 70.00 oberservations per class (harmonic mean).
Testing Differences (p = 0.05) PLSD = 0.0399
Bonferroni correction gives PLSD = 0.0491
NK = 0.04/0.05

ANALYSIS OF INTENSITY CARIC 2=WAY ANOVA

Number of cells = 81 Number of data points = 81

STRUCTURE of ANOVA
Structure Parameters : SUBJECTS=9 DURATION=3 CARIC=3

1	SUBJECTS	9 levels. RANDOM
2	DURATION	3 levels. FIXED
3	CARIC	3 levels. FIXED

RESULTS OF ANALYSIS

Type	Source	df	SS	MS	F	p	MSE	df
MAIN EFFECT								
1 Random	SUBJECTS :	8	1.546128	0.193266	9.744	0.0000	0.019834	32
2 Fixed	DURATION :	2	4.866065	2.433033	22.478	0.0000	0.108242	16
3 Fixed	CARIC :	2	0.025476	0.012738	0.565	0.5794	0.022553	16
INTERACTIONS								
4:	1 2 :	16	1.731874	0.108242	5.457	0.0000	0.019834	32
5:	1 3 :	16	0.360849	0.022553	1.137	0.3655	0.019834	32
6:	2 3 :	4	0.157041	0.039260	1.979	0.1214	0.019834	32
7:	1 2 3 :	32	0.634686	0.019834				
	TOTAL	80	9.322119					

EXPECTED MEAN SQUARES

	Source	EMS
1:	SUBJECTS	9.0(1) + var(err)
2:	DURATION	4.5(4) + 13.5[2] + var(err)
3:	CARIC	4.5(5) + 13.5[3] + var(err)
Interactions		
4:	1 2	4.5(4) + var(err)
5:	1 3	4.5(5) + var(err)
6:	2 3	2.3[6] + var(err)
7:	1 2 3	var(err)

CLASS SUMMARIES

ANALYSIS OF INTENSITY CARIC 2=WAY ANOVA

FACTOR: DURATION (between)
Main effect at p = 0.0000 (df=16)

	MEAN	VARIANCE	SD	S.E.M.	N
Class 1	0.5488	0.0872	0.2953	0.0568	27
Class 2	0.9306	0.0308	0.1754	0.0338	27
Class 3	1.1410	0.0534	0.2311	0.0445	27

Pooled S.E.M. = 0.06 27.00 oberservations per class (harmonic mean).
Testing Differences (p = 0.05) PLSD = 0.1898
Bonferroni correction gives PLSD = 0.2394
NK = 0.19/0.23

CLASS SUMMARIES

ANALYSIS OF INTENSITY CARIC 2=WAY ANOVA
FACTOR: CARIC (between)
Non-significant main effect (p=0.5794)

	MEAN	VARIANCE	SD	S.E.M.	N
Class 1	0.8961	0.1177	0.3430	0.0660	27
Class 2	0.8528	0.1298	0.3603	0.0693	27
Class 3	0.8714	0.1101	0.3317	0.0638	27

ANALYSIS OF 2-WAY ANOVA COLOUR VS. GREY-SCALE AND CONTRAST CONTROL FOR RECOGNITION BY

Number of cells = 54 Number of data points = 54

STRUCTURE of ANOVA
Structure Parameters : DURATION=2 COLOUR=3 ITEM=-9

1	DURATION	2 levels. FIXED
2	COLOUR	3 levels. FIXED
3	ITEM	9 levels. RANDOM

RESULTS OF ANALYSIS

Type	Source	df	SS	MS	F	p	MSE	df
MAIN EFFECT								
1 Fixed	DURATION :	1	0.1683078	0.1683078	17.479	0.0031	0.0096290	8
2 Fixed	COLOUR :	2	0.0433564	0.0216782	2.071	0.1585	0.0104664	16
3 Random	ITEM :	8	1.7734161	0.2216770	72.745	0.0000	0.0030473	16
INTERACTIONS								
4:	1 2 :	2	0.0032343	0.0016172	0.531	0.5982	0.0030473	16
5:	1 3 :	8	0.0770317	0.0096290	3.160	0.0239	0.0030473	16
6:	2 3 :	16	0.1674620	0.0104664	3.435	0.0091	0.0030473	16
7:	1 2 3 :	16	0.0487568	0.0030473				
	TOTAL	53	2.2815651					

EXPECTED MEAN SQUARES

	Source	EMS
1:	DURATION	6.0(5) + 27.0[1] + var(err)
2:	COLOUR	3.0(6) + 9.0[2] + var(err)
3:	ITEM	6.0(3) + var(err)
Interactions		
4:	1 2	4.5[4] + var(err)
5:	1 3	6.0(5) + var(err)
6:	2 3	3.0(6) + var(err)
7:	1 2 3	var(err)

CLASS SUMMARIES

ANALYSIS OF 2-WAY ANOVA COLOUR VS. GREY-SCALE AND CONTRAST CONTROL FOR RECOGNITION BY ITEM

FACTOR: DURATION (between)
Main effect at p = 0.0031 (df=8)

	MEAN	VARIANCE	SD	S.E.M.	N
Class 1	0.7079	0.0345	0.1857	0.0357	27
Class 2	0.8195	0.0468	0.2163	0.0416	27

Pooled S.E.M. = 0.02 27.00 oberservations per class (harmonic mean).
Testing Differences (p = 0.05) PLSD = 0.0616
Bonferroni correction gives PLSD = 0.0616
NK = 0.06

DATA USED

-0.016	-0.108	0.033	0.000	0.179	-0.003	-0.262	-0.133	-0.153
-0.051	0.016	-0.004	0.136	-0.068	-0.213	-0.203	-0.105	0.106
-0.256	-0.306	-0.061	-0.163	-0.341	0.007	-0.051	-0.002	-0.175
-0.002	-0.274	0.023						

N = 30 Mean = -0.08 S.E.M. = 0.024
Variance = 0.02 Df = 29

Null hypothesis mean = 0.000
t = -3.394 p = 0.002 d.f. = 29

DATA USED

```
-0.120    -0.240     0.040    -0.080    -0.100     0.010

N =  6     Mean = -0.08    S.E.M. = 0.041
Variance =    0.01    Df =   5


 Null hypothesis mean =  0.000
t = -1.999     p =  0.102     d.f. =  5
```

General Linear Model

Warnings

The DESIGN subcommand is empty, so a saturated design will be generated.
The subcommand WSDESIGN is empty and will be ignored.
Since WSDESIGN was not specified, a saturated within subject design will be generated.

Within-Subjects Factors

Measure: MEASURE_1

FACE	Dependent Variable
1	DENIRO
2	FORD
3	KEVIN
4	PACINO
5	SLATER
6	WILLIS

Multivariate Tests[a]

Effect		Value	F	Hypothesis df	Error df	Sig.	Noncent. Parameter	Observed Power[b]
FACE	Pillai's Trace	.652	9.364[c]	5.000	25.000	.000	46.818	.999
	Wilks' Lambda	.348	9.364[c]	5.000	25.000	.000	46.818	.999
	Hotelling's Trace	1.873	9.364[c]	5.000	25.000	.000	46.818	.999
	Roy's Largest Root	1.873	9.364[c]	5.000	25.000	.000	46.818	.999

a. Design: Intercept
Within Subjects Design: FACE

b. Computed using alpha = .05

c. Exact statistic

Descriptives

Descriptive Statistics

	N	Mean	
	Statistic	Statistic	Std. Error
FORD	30	-.2352	4.3E-02
DENIRO	30	-.1168	4.5E-02
SLATER	30	-.1046	5.0E-02
PACINO	30	-8.E-02	4.3E-02
WILLIS	30	9.0E-03	4.8E-02
KEVIN	30	3.7E-02	4.3E-02
Valid N (listwise)	30		

T-Test

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
DENIRO	30	-.1168	.2455	4.5E-02
KEVIN	30	3.7E-02	.2332	4.3E-02
FORD	30	-.2352	.2368	4.3E-02
PACINO	30	-8.E-02	.2368	4.3E-02
SLATER	30	-.1046	.2713	5.0E-02
WILLIS	30	9.0E-03	.2613	4.8E-02

One-Sample Test

	Test Value = 0					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
DENIRO	-2.606	29	.014	-.1168	-.2084	-3.E-02
KEVIN	.861	29	.396	3.667E-02	-5.E-02	.1237
FORD	-5.441	29	.000	-.2352	-.3237	-.1468
PACINO	-1.832	29	.077	-7.92E-02	-.1676	9.2E-03
SLATER	-2.111	29	.043	-.1046	-.2059	-3.E-03
WILLIS	.188	29	.852	8.967E-03	-9.E-02	.1065

DATA USED - COLOUR CARICATURE

-0.027	0.015	0.126	0.177	0.116	0.129	0.311	0.136	0.269
0.225	0.079	0.165	0.142	-0.042	-0.068	0.187	0.413	0.409
0.330	0.299	0.059	-0.019	-0.359	0.306	0.168	0.349	-0.064
-0.025	0.381	0.141						

N = 30   Mean = 0.14   S.E.M. = 0.032
Variance = 0.03   Df = 29

Null hypothesis mean = 0.000
t = 4.560   p = 0.000   d.f. = 29

DATA USED - CONTRAST CONTROL

0.148	0.117	0.210	0.270	0.497	0.333	0.341	0.236	0.433
0.286	0.221	0.142	0.454	0.231	0.109	0.302	0.381	0.376
0.299	0.357	0.429	0.156	-0.069	0.472	0.178	0.305	0.061
0.051	0.687	0.289						

N = 30   Mean = 0.28   S.E.M. = 0.029
Variance = 0.02   Df = 29

Null hypothesis mean = 0.000
t = 9.681   p = 0.000   d.f. = 29

DATA USED - COLOUR CARICATURE

0.205 0.209 0.103 0.143 0.062

N = 5 Mean = 0.14 S.E.M. = 0.029
Variance = 0.00 Df = 4

Null hypothesis mean = 0.000
t = 5.051 p = 0.007 d.f. = 4

DATA USED - CONTRAST CONTROL

0.334 0.368 0.178 0.299 0.254

N = 5 Mean = 0.29 S.E.M. = 0.033
Variance = 0.01 Df = 4

Null hypothesis mean = 0.000
t = 8.690 p = 0.001 d.f. = 4

T-Test

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	COLOUR	.1446	5	6.400E-02	2.9E-02
	CONTRAST	.2867	5	7.378E-02	3.3E-02

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	COLOUR & CONTRAST	5	.807	.099

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	COLOUR - CONTRAST	-.1421	4.376E-02	2.0E-02	-.1965	-9.E-02	-7.262	4	.002

T-Test

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	COL_SUB	.1446	30	.1856	3.4E-02
	CON_SUB	.2867	30	.1625	3.0E-02

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	COL_SUB & CON_SUB	30	.784	.000

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	COL_SUB - CON_SUB	-.1421	.1166	2.1E-02	-.1857	-1.E-01	-6.678	29	.000

Descriptives

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
COL_SUB	30	-.32	.58	.1446	.1856
CON_SUB	30	-.04	.67	.2867	.1625
COLOUR	5	.06	.21	.1446	6.400E-02
CONTRAST	5	.18	.37	.2867	7.378E-02
Valid N (listwise)	5				

General Linear Model

Warnings

The DESIGN subcommand is empty, so a saturated design will be generated.
The subcommand WSDESIGN is empty and will be ignored.
Since WSDESIGN was not specified, a saturated within subject design will be generated.

Within-Subjects Factors

Measure: MEASURE_1

COL_CON	FACE	Dependent Variable
1	1	FAC1_COL
	2	FAC2_COL
	3	FAC3_COL
	4	FAC4_COL
	5	FAC5_COL
2	1	FAC1_CON
	2	FAC2_CON
	3	FAC3_CON
	4	FAC4_CON
	5	FAC5_CON

Between-Subjects Factors

		Value Label
HILOW	1.00	no label defined
	2.00	no label defined

Multivariate Tests[a]

Effect		Value	F	Hypothesis df	Error df	Sig.	Noncent. Parameter	Observed Power[b]
COL_CON	Pillai's Trace	.728	21.425[c]	1.000	8.000	.002	21.425	.981
	Wilks' Lambda	.272	21.425[c]	1.000	8.000	.002	21.425	.981
	Hotelling's Trace	2.678	21.425[c]	1.000	8.000	.002	21.425	.981
	Roy's Largest Root	2.678	21.425[c]	1.000	8.000	.002	21.425	.981
COL_CON * HILOW	Pillai's Trace	.008	.062[c]	1.000	8.000	.810	.062	.056
	Wilks' Lambda	.992	.062[c]	1.000	8.000	.810	.062	.056

Multivariate Tests[a]

Effect		Value	F	Hypothesis df	Error df	Sig.	Noncent. Parameter	Observed Power[b]
COL_CON * HILOW	Hotelling's Trace	.008	.062[c]	1.000	8.000	.810	.062	.056
	Roy's Largest Root	.008	.062[c]	1.000	8.000	.810	.062	.056
FACE	Pillai's Trace	.648	2.301[c]	4.000	5.000	.193	9.205	.322
	Wilks' Lambda	.352	2.301[c]	4.000	5.000	.193	9.205	.322
	Hotelling's Trace	1.841	2.301[c]	4.000	5.000	.193	9.205	.322
	Roy's Largest Root	1.841	2.301[c]	4.000	5.000	.193	9.205	.322
FACE * HILOW	Pillai's Trace	.555	1.556[c]	4.000	5.000	.316	6.224	.228
	Wilks' Lambda	.445	1.556[c]	4.000	5.000	.316	6.224	.228
	Hotelling's Trace	1.245	1.556[c]	4.000	5.000	.316	6.224	.228
	Roy's Largest Root	1.245	1.556[c]	4.000	5.000	.316	6.224	.228
COL_CON * FACE	Pillai's Trace	.298	.531[c]	4.000	5.000	.720	2.124	.105
	Wilks' Lambda	.702	.531[c]	4.000	5.000	720	2.124	.105
	Hotelling's Trace	.425	.531[c]	4.000	5.000	.720	2.124	.105
	Roy's Largest Root	.425	.531[c]	4.000	5.000	.720	2.124	.105
COL_CON * FACE * HILOW	Pillai's Trace	.462	1.075[c]	4.000	5.000	.457	4.299	.168
	Wilks' Lambda	.538	1.075[c]	4.000	5.000	.457	4.299	.168
	Hotelling's Trace	.860	1.075[c]	4.000	5.000	.457	4.299	.168
	Roy's Largest Root	.860	1.075[c]	4.000	5.000	.457	4.299	.168

a. Design: Intercept+HILOW
Within Subjects Design: COL_CON+FACE+COL_CON*FACE

b. Computed using alpha = .05

c. Exact statistic

Mauchly's Test of Sphericity[a]

Measure: MEASURE_1

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon[b]		
					Greenhouse-Geisser	Huynh-Feldt	Lower-bound
COL_CON	1.000	.000	0	.	1.000	1.000	1.000
FACE	.101	14.682	9	.110	.677	1.000	.250
COL_CON * FACE	.215	9.877	9	.376	.602	.987	.250

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

a. Design: Intercept+HILOW
Within Subjects Design: COL_CON+FACE+COL_CON*FACE

b. May be used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed in the layers (by default) of the Tests of Within Subjects Effects table.

Tests of Within-Subjects Effects

Measure: MEASURE_1
Sphericity Assumed

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Noncent. Parameter	Observed Power[a]
COL_CON	6072.826	1	6072.826	21.425	.002	21.425	.981
COL_CON * HILOW	17.503	1	17.503	.062	.810	.062	.056
Error(COL_CON)	2267.605	8	283.451				
FACE	4703.815	4	1175.954	3.480	.018	13.920	.803
FACE * HILCW	971.087	4	242.772	.718	.586	2.874	.206
Error(FACE)	10813.6	32	337.924				
COL_CON * FACE	896.154	4	224.038	1.091	.378	4.362	.303
COL_CON * FACE * HILOW	529.475	4	132.369	.644	.635	2.577	.188
Error(COL_CON*FACE)	6573.745	32	205.430				

a. Computed using alpha = .05

Tests of Within-Subjects Contrasts

Measure: MEASURE_1

Source	Transformed Variable	Type III Sum of Squares	df	Mean Square	F	Sig.	Noncent. Parameter
COL_CON	COL_CON_1	6072.826	1	6072.826	21.425	.002	21.425
COL_CON * HILOW	COL_CON_1	17.503	1	17.503	.062	.810	.062
Error(COL_CON)	COL_CON_1	2267.605	8	283.451			
FACE	FACE_1	3269.493	1	3269.493	7.221	.028	7.221
	FACE_2	7.3E-03	1	7.3E-03	.000	.996	.000
	FACE_3	1395.610	1	1395.610	8.646	.019	8.646
	FACE_4	38.705	1	38.705	.080	.784	.080
FACE * HILOW	FACE_1	156.008	1	156.008	.345	.573	.345
	FACE_2	10.504	1	10.504	.041	.845	.041
	FACE_3	465.979	1	465.979	2.887	.128	2.887
	FACE_4	338.595	1	338.595	.704	.426	.704
Error(FACE)	FACE_1	3622.132	8	452.767			
	FACE_2	2053.466	8	256.683			
	FACE_3	1291.365	8	161.421			
	FACE_4	3846.615	8	480.827			
COL_CON * FACE	COL_CON_1*FACE_1	3.965	1	3.965	.036	.854	.036
	COL_CON_1*FACE_2	58.082	1	58.082	.143	.716	.143
	COL_CON_1*FACE_3	326.861	1	326.861	2.441	.157	2.441
	COL_CON_1*FACE_4	507.245	1	507.245	2.964	.123	2.964
COL_CON * FACE * HILOW	COL_CON_1*FACE_1	463.053	1	463.053	4.235	.074	4.235
	COL_CON_1*FACE_2	42.396	1	42.396	.104	.755	.104
	COL_CON_1*FACE_3	8.921	1	8.921	.067	.803	.067
	COL_CON_1*FACE_4	15.105	1	15.105	.088	.774	.088
Error(COL_CON*FACE)	COL_CON_1*FACE_1	874.785	8	109.348			
	COL_CON_1*FACE_2	3258.701	8	407.338			
	COL_CON_1*FACE_3	1071.227	8	133.903			
	COL_CON_1*FACE_4	1369.032	8	171.129			

Tests of Within-Subjects Contrasts

Measure: MEASURE_1

Source	Transformed Variable	Observed Power[a]
COL_CON	COL_CON_1	.981
COL_CON * HILOW	COL_CON_1	.056
Error(COL_CON)	COL_CON_1	
FACE	FACE_1	.655
	FACE_2	.050
	FACE_3	.731
	FACE_4	.057
FACE * HILOW	FACE_1	.081
	FACE_2	.054
	FACE_3	.322
	FACE_4	.115
Error(FACE)	FACE_1	
	FACE_2	
	FACE_3	
	FACE_4	
COL_CON * FACE	COL_CON_1*FACE_1	.053
	COL_CON_1*FACE_2	.063
	COL_CON_1*FACE_3	.281
	COL_CON_1*FACE_4	.329
COL_CON * FACE * HILOW	COL_CON_1*FACE_1	.441
	COL_CON_1*FACE_2	.059
	COL_CON_1*FACE_3	.056
	COL_CON_1*FACE_4	.058
Error(COL_CON*FACE)	COL_CON_1*FACE_1	
	COL_CON_1*FACE_2	
	COL_CON_1*FACE_3	
	COL_CON_1*FACE_4	

a. Computed using alpha = .05

Tests of Between-Subjects Effects

Measure: MEASURE_1
Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Noncent. Parameter	Observed Power[a]
Intercept	55262.8	1	55262.8	23.249	.001	23.249	.987
HILOW	4429.124	1	4429.124	1.863	.209	1.863	.226
Error	19015.9	8	2376.983				

a. Computed using alpha = .05

Descriptives

Descriptive Statistics

HILOW		N	Minimum	Maximum	Mean	Std. Deviation
1.00	FAC1_CON	5	12.64	68.00	41.7280	19.6228
	FAC2_CON	5	41.76	56.80	47.2000	5.7433
	FAC3_CON	5	23.36	48.96	33.5040	12.0286
	FAC4_CON	5	14.08	56.00	40.8640	17.1258
	FAC5_CON	5	17.60	33.60	28.5760	6.3373
	FAC1_COL	5	29.76	60.16	38.4640	12.2995
	FAC2_COL	5	-26.40	45.60	24.1600	29.4350
	FAC3_COL	5	.64	30.88	19.6800	11.7864
	FAC4_COL	5	-11.04	48.00	19.7120	22.9075
	FAC5_COL	5	-13.76	32.48	7.7440	16.9806
	Valid N (listwise)	5				
2.00	FAC1_CON	5	13.44	74.40	35.8720	25.6103
	FAC2_CON	5	.96	49.92	24.8000	22.2837
	FAC3_CON	5	-17.12	68.16	21.6807	31.3045
	FAC4_CON	5	5.44	78.08	26.4320	29.9423
	FAC5_CON	5	-28.80	66.24	12.3520	34.2421
	FAC1_COL	5	2.08	48.48	20.2560	17.7984
	FAC2_COL	5	-33.92	23.36	-4.4480	20.7900
	FAC3_COL	5	-20.96	58.56	15.1360	28.5147
	FAC4_COL	5	-1.12	42.24	14.6240	16.7545
	FAC5_COL	5	-33.92	36.32	1.8240	27.5229
	Valid N (listwise)	5				

Experiment 1) Chapter 3

Descriptives

Descriptive Statistics

SUB_RACE		N	Minimum	Maximum	Mean	Std. Deviation
caucasia	JAP_FEM	50	.02	.97	.6017	.3149
	JAP_MALE	50	.02	.97	.4075	.2925
	JH_FEM	50	.02	.99	.7417	.2238
	JH_MALE	50	.02	.97	.3528	.2508
	Valid N (listwise)	50				
japanese	JAP_FEM	42	.23	.98	.7288	.1943
	JAP_MALE	42	.02	.92	.2860	.2151
	JH_FEM	42	.13	.97	.6531	.2209
	JH_MALE	42	.02	.86	.3264	.2348
	Valid N (listwise)	42				

General Linear Model

Warnings

The DESIGN subcommand is empty, so a saturated design will be generated.
The subcommand WSDESIGN is empty and will be ignored.
Since WSDESIGN was not specified, a saturated within subject design will be generated.

Within-Subjects Factors

Measure: MEASURE_1

FACE	Dependent Variable
1	JAP_MALE
2	JH_MALE

Between-Subjects Factors

		Value Label
SUB_GEN	female	no label defined
	male	no label defined
SUB_RACE	caucasia	no label defined
	japanese	no label defined

Multivariate Tests[a]

Effect		Value	F	Hypothesis df	Error df	Sig.	Noncent. Parameter	Observed Power[b]
FACE	Pillai's Trace	.000	.018[c]	1.000	88.000	.894	.018	.052
	Wilks' Lambda	1.000	.018[c]	1.000	88.000	.894	.018	.052
	Hotelling's Trace	.000	.018[c]	1.000	88.000	.894	.018	.052
	Roy's Largest Root	.000	.018[c]	1.000	88.000	.894	.018	.052
FACE * SUB_GEN	Pillai's Trace	.002	.141[c]	1.000	88.000	.709	.141	.066
	Wilks' Lambda	.998	.141[c]	1.000	88.000	.709	.141	.066
	Hotelling's Trace	.002	.141[c]	1.000	88.000	.709	.141	.066
	Roy's Largest Root	.002	.141[c]	1.000	88.000	.709	.141	.066
FACE * SUB_RACE	Pillai's Trace	.031	2.800[c]	1.000	88.000	.098	2.800	.380
	Wilks' Lambda	.969	2.800[c]	1.000	88.000	.098	2.800	.380
	Hotelling's Trace	.032	2.800[c]	1.000	88.000	.098	2.800	.380
	Roy's Largest Root	.032	2.800[c]	1.000	88.000	.098	2.800	.380
FACE * SUB_GEN * SUB_RACE	Pillai's Trace	.034	3.076[c]	1.000	88.000	.083	3.076	.411
	Wilks' Lambda	.966	3.076[c]	1.000	88.000	.083	3.076	.411
	Hotelling's Trace	.035	3.076[c]	1.000	88.000	.083	3 076	.411
	Roy's Largest Root	.035	3.076[c]	1.000	88.000	.083	3.076	.411

a. Design: Intercept+SUB_GEN+SUB_RACE+SUB_GEN * SUB_RACE
Within Subjects Design: FACE

b. Computed using alpha = .05

c. Exact statistic

Mauchly's Test of Sphericity[a]

Measure: MEASURE_1

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon[b]		
					Greenhouse-Geisser	Huynh-Feldt	Lower-bound
FACE	1.000	.000	0	.	1.000	1.000	1.000

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

a. Design: Intercept+SUB_GEN+SUB_RACE+SUB_GEN * SUB_RACE
Within Subjects Design: FACE

b. May be used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed in the layers (by default) of the Tests of Within Subjects Effects table.

Tests of Within-Subjects Effects

Measure: MEASURE_1
Sphericity Assumed

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Noncent. Parameter	Observed Power[a]
FACE	7.4E-04	1	7.4E-04	.018	.894	.018	.052
FACE * SUB_GEN	5.9E-03	1	5.9E-03	.141	.709	.141	.066
FACE * SUB_RACE	.117	1	.117	2.800	.098	2.800	.380
FACE * SUB_GEN * SUB_RACE	.128	1	.128	3.076	.083	3.076	.411
Error(FACE)	3.663	88	4.2E-02				

a. Computed using alpha = .05

Tests of Within-Subjects Contrasts

Measure: MEASURE_1

Source	Transformed Variable	Type III Sum of Squares	df	Mean Square	F	Sig.
FACE	FACE_1	7.4E-04	1	7.4E-04	.018	.894
FACE * SUB_GEN	FACE_1	5.9E-03	1	5.9E-03	.141	.709
FACE * SUB_RACE	FACE_1	.117	1	.117	2.800	.098
FACE * SUB_GEN * SUB_RACE	FACE_1	.128	1	.128	3.076	.083
Error(FACE)	FACE_1	3.663	88	4.2E-02		

Tests of Within-Subjects Contrasts

Measure: MEASURE_1

Source	Transformed Variable	Noncent. Parameter	Observed Power[a]
FACE	FACE_1	.018	.052
FACE * SUB_GEN	FACE_1	.141	.066
FACE * SUB_RACE	FACE_1	2.800	.380
FACE * SUB_GEN * SUB_RACE	FACE_1	3.076	.411
Error(FACE)	FACE_1		

a. Computed using alpha = .05

Tests of Between-Subjects Effects

Measure: MEASURE_1
Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Noncent. Parameter	Observed Power[a]
Intercept	21.354	1	21.354	248.268	.000	248.268	1.000
SUB_GEN	1.5E-02	1	1.5E-02	.179	.673	.179	.070
SUB_RACE	.253	1	.253	2.943	.090	2.943	.396
SUB_GEN * SUB_RACE	5.0E-02	1	5.0E-02	.578	.449	.578	.117
Error	7.569	88	8.6E-02				

a. Computed using alpha = .05

General Linear Model

Warnings

The DESIGN subcommand is empty, so a saturated design will be generated.
The subcommand WSDESIGN is empty and will be ignored.
Since WSDESIGN was not specified, a saturated within subject design will be generated.

Within-Subjects Factors

Measure: MEASURE_1

FACE	Dependent Variable
1	JAP_FEM
2	JH_FEM

Between-Subjects Factors

		Value Label
SUB_GEN	female	no label defined
	male	no label defined
SUB_RACE	caucasia	no label defined
	japanese	no label defined

Multivariate Tests[a]

Effect		Value	F	Hypothesis df	Error df	Sig.	Noncent. Parameter	Observed Power[b]
FACE	Pillai's Trace	.016	1.419[c]	1.000	88.000	.237	1.419	.218
	Wilks' Lambda	.984	1.419[c]	1.000	88.000	.237	1.419	.218
	Hotelling's Trace	.016	1.419[c]	1.000	88.000	.237	1.419	.218
	Roy's Largest Root	.016	1.419[c]	1.000	88.000	.237	1.419	.218
FACE * SUB_GEN	Pillai's Trace	.002	.157[c]	1.000	88.000	.693	.157	.068
	Wilks' Lambda	.998	.157[c]	1.000	88.000	.693	.157	.068
	Hotelling's Trace	.002	.157[c]	1.000	88.000	.693	.157	.068
	Roy's Largest Root	.002	.157[c]	1.000	88.000	.693	.157	.068
FACE * SUB_RACE	Pillai's Trace	.162	17.063[c]	1.000	88.000	.000	17.063	.983
	Wilks' Lambda	.838	17.063[c]	1.000	88.000	.000	17.063	.983
	Hotelling's Trace	.194	17.063[c]	1.000	88.000	.000	17.063	.983
	Roy's Largest Root	.194	17.063[c]	1.000	88.000	.000	17.063	.983
FACE * SUB_GEN * SUB_RACE	Pillai's Trace	.001	.060[c]	1.000	88.000	.807	.060	.057
	Wilks' Lambda	.999	.060[c]	1.000	88.000	.807	.060	.057
	Hotelling's Trace	.001	.060[c]	1.000	88.000	.807	.060	.057
	Roy's Largest Root	.001	.060[c]	1.000	88.000	.807	.060	.057

a. Design: Intercept+SUB_GEN+SUB_RACE+SUB_GEN * SUB_RACE
Within Subjects Design: FACE

b. Computed using alpha = .05

c. Exact statistic

Mauchly's Test of Sphericity[a]

Measure: MEASURE_1

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon[b]		
					Greenhouse-Geisser	Huynh-Feldt	Lower-bound
FACE	1.000	.000	0	.	1.000	1.000	1.000

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

a. Design: Intercept+SUB_GEN+SUB_RACE+SUB_GEN * SUB_RACE
Within Subjects Design: FACE

b. May be used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed in the layers (by default) of the Tests of Within Subjects Effects table.

Tests of Within-Subjects Effects

Measure: MEASURE_1
Sphericity Assumed

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Noncent. Parameter	Observed Power[a]
FACE	4.5E-02	1	4.5E-02	1.419	.237	1.419	.218
FACE * SUB_GEN	4.9E-03	1	4.9E-03	.157	.693	.157	.068
FACE * SUB_RACE	.536	1	.536	17.063	.000	17.063	.983
FACE * SUB_GEN * SUB_RACE	1.9E-03	1	1.9E-03	.060	.807	.060	.057
Error(FACE)	2.766	88	3.1E-02				

a. Computed using alpha = .05

Tests of Within-Subjects Contrasts

Measure: MEASURE_1

Source	Transformed Variable	Type III Sum of Squares	df	Mean Square	F	Sig.
FACE	FACE_1	4.5E-02	1	4.5E-02	1.419	.237
FACE * SUB_GEN	FACE_1	4.9E-03	1	4.9E-03	.157	.693
FACE * SUB_RACE	FACE_1	.536	1	.536	17.063	.000
FACE * SUB_GEN * SUB_RACE	FACE_1	1.9E-03	1	1.9E-03	.060	.807
Error(FACE)	FACE_1	2.766	88	3.1E-02		

Tests of Within-Subjects Contrasts

Measure: MEASURE_1

Source	Transformed Variable	Noncent. Parameter	Observed Power[a]
FACE	FACE_1	1.419	.218
FACE * SUB_GEN	FACE_1	.157	.068
FACE * SUB_RACE	FACE_1	17.063	.983
FACE * SUB_GEN * SUB_RACE	FACE_1	.060	.057
Error(FACE)	FACE_1		

a. Computed using alpha = .05

Tests of Between-Subjects Effects

Measure: MEASURE_1
Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Noncent. Parameter	Observed Power[a]
Intercept	85.057	1	85.057	960.949	.000	960.949	1.000
SUB_GEN	.140	1	.140	1.576	.213	1.576	.237
SUB_RACE	2.8E-02	1	2.8E-02	.320	.573	.320	.087
SUB_GEN * SUB_RACE	.188	1	.188	2.130	.148	2.130	.303
Error	7.789	88	8.9E-02				

a. Computed using alpha = .05