The Role of Physical Image Properties in Facial Expression and Identity Perception

Mladen Sormaz

Doctor of Philosophy

University of York

Psychology

May 2016

Abstract

A number of attempts have been made to understand which physical image properties are important for the perception of different facial characteristics. These physical image properties have been broadly split in to two categories; namely facial shape and facial surface. Current accounts of face processing suggest that whilst judgements of facial identity rely approximately equally on facial shape and surface properties, judgements of facial expression are heavily shape dependent. This thesis presents behavioural experiments and fMRI experiments employing multi voxel pattern analysis (MVPA) to investigate the extent to which facial shape and surface properties underpin identity and expression perception and how these image properties are represented neurally. The first empirical chapter presents experiments showing that facial expressions are categorised approximately equally well when either facial shape or surface is the varying image cue. The second empirical chapter shows that neural patterns of response to facial expressions in the Occipital Face Area (OFA) and Superior Temporal Sulcus (STS) are reflected by patterns of perceptual similarity of the different expressions, in turn these patterns of perceptual similarity can be predicted by both facial shape and surface properties. The third empirical chapter demonstrates that distinct patterns of neural response can be found to shape based but not surface based cues to facial identity in the OFA and Fusiform Face Area (FFA). The final experimental chapter in this thesis demonstrates that the newly discovered contrast chimera effect is heavily dependent on the eye region and holistic face representations conveying facial identity. Taken together, these findings show the importance of facial surface as well as facial shape in expression perception. For facial identity both facial shape and surface cues are important for the contrast chimera effect although there are more consistent identity based neural response patterns to facial shape in face responsive brain regions.

List of Contents

List of Tables

List of Figures

Acknowledgements

This thesis is the most challenging academic and personal undertaking I have faced , as well as my own hard work this thesis is the result of a great number of contributions to my life by important people both academically and personally. Of course my first gratitude and love will always reserved for my parents Rajko and Svetlana Sormaz, brother Vladan Sormaz , grandmother Nada Nisevic . I also reserve a special thank you for Kirstie Wailes-Newson for all of your love, support and constant supply of baking.

This thesis has been a challenging but ultimately rewarding experience through which I've learned a lot and I am extremely grateful to have had this opportunity. For this reason I would like to firstly extend my deepest gratitude to my supervisors Prof. Andy Young and Prof. Tim Andrews. Their supervision and support has been excellent throughout and I feel extremely lucky to have had such clear guidance and investment. In addition their general generosity and humour throughout has meant a great deal to me and will be something I always treasure. I'd also like to thank Dr. William Smith for being part of my thesis advisory panel and providing excellent advice and support throughout my thesis. Furthermore I'd like to acknowledge the pleasure it was to work in a lab group with my colleagues Dr. David Watson, Tessa Flack, Dr. Richard Carter-Harris and Dr. Clare Sutherland. I would also like to acknowledge times where the technical support of Dr. Andrew Gouws and Dr. Mark Hymers was invaluable.

A number of personal relationships have been significant throughout my doctorate and I'd like to particularly acknowledge the friendship and support of the following: Dr. Andrew Quinn, Dr. Edward Silson, Dr. Markus van Ackeren, Catarina Teige, Dr. Emma Holmes, Dr. Adele Goman, Dr. Bruce Keefe, Dr. Scott Cairney, Dr. Daniel Baker and Prof. Pete Thompson. I'd also like to thank all of the past and present members of Wentworth graduate football club and the Heavy Rollers staff cricket club for providing an excellent escape from all of life's stresses. At various stages you've helped me delude myself into believing that if I just had one really good season I might bring back the trend of being a dual England international at both football and cricket. I'd also like to extend my thanks to all of the members of the James College welfare team who I've had the pleasure to work alongside over the last few years.

Finally, given the difficulties presented to my family by the Balkan civil wars I'd like to also extend my greatest appreciation for the help and shelter provided to us by both the Sormaz and Nisevic branches of my family and to the Parojcic family for the great help they provided for my family when settling here in Great Britain. Without their generosity and kindness I am certain I would never have had this opportunity and words will never truly express the true extent of my gratitude.

This thesis is dedicated to the memory of Gojko Nisevic (1932 -1997), Milojko Sormaz (1926 - 2011), Predrag Nisevic (1956 - 1992) and Ranko Sormaz (1951 - 1995).

Author’s Declaration

This thesis presents original work completed by the author, Mladen Sormaz, under the joint supervision of Prof. Andrew Young and Prof. Timothy Andrews. The empirical work presented in this thesis has been published or is currently under review in the following peer-reviewed journals:

Sormaz, M., Andrews, T. J., & Young, A. W. (2013). Contrast negation and the importance of the eye region for holistic representations of facial identity. Journal of Experimental Psychology. Human Perception and Performance, 39(6), 1667–77.

Sormaz, M., Watson, D. M., Smith, W. A. P., Young, A. W., & Andrews, T. J. (2016). Modelling the perceptual similarity of facial expressions from image statistics and neural responses. NeuroImage, 129, 64-71.

Sormaz, M., Young, A. W., & Andrews, T. J. (under review). Contributions of feature shapes and surface cues to the recognition of facial expressions.

Results from multiple empirical chapters have been presented at the following conferences:

Poster - Sormaz, M., Andrews, T.J., Young, A.W. (2013) *Contrast reversal supports the importance of the eye region for holistic representations of facial identity*. Poster presented at the Vision Sciences Society meeting, Naples, Florida, May 2013

Poster - Sormaz, M., Andrews, T.J., Young, A.W. (2013) *Contrast reversal supports the importance of the eye region for holistic representations of facial identity*. Poster presented at the Yorkshire Vision Network meeting, York, North Yorkshire, June 2013.

Talk – Sormaz, M., Young, A.W., Smith, W.A.P, Watson, D.M., Andrews, T.J. (2014). *Representation of facial expressions in face-responsive brain regions.* London, UK, January 2014.

Poster - Sormaz, M., Andrews, T.J., Young, A.W. (2014) *Continuous and categorical patterns of neural response to facial expressions in face-selective regions of the human brain*. Poster presented at the Vision Sciences Society meeting, St Pete Beach, Florida, May 2014

Talk – Sormaz, M., Young, A.W., Smith, W.A.P, Watson, D.M., Andrews, T.J. (2015). *Neural and perceptual responses to facial expressions track internal feature shapes*. London, UK, January 2015.

Poster - Sormaz, M., Young, A.W., Watson D.M., Andrews, T.J. (2015) *Image properties from the internal and external features of the face differentially predict patterns of neural response to expression and identity in human visual cortex.* Poster presented at the Vision Sciences Society meeting, St Pete Beach, Florida, May 2015

Poster - Sormaz, M., Young, A.W., Smith, W.A.P, Andrews, T.J. (2015) *Contributions of feature shape and surface cues to facial expression perception*. Poster presented at the European Conference of Visual Perception meeting, Liverpool, UK, August 2015.

Talk – Sormaz, M., Young, A.W., Smith, W.A.P, Watson, D.M., Andrews, T.J. (2016). *Contributions of feature shape and surface cues to facial expression perception.* London, UK, January 2016.

This work has not previously been presented for an award at this, or any other, University. All sources are acknowledged as References.

Chapter 1 - Literature Review: How facial expression and identity perception relate to image properties and the cortical face perception network

1.1 Perception of facial identity and expression

Two of the most salient facial characteristics are those that convey facial identity (a relatively invariant facial characteristic) and facial expression (a highly changeable facial characteristic). Many models of face perception (e.g. Bruce and Young, 1986; Haxby et al 2000; figure 1.2) suggest that the routes by which these characteristics are processed separates immediately after the initial processing of facial structure and physical properties (Calder & Young, 2005). This complete separation of routes for facial identity and expression processing was questioned however in a review by Calder & Young (2005) who suggested that this assumed separation is based on limited evidence. For example, this assumption is based on a small number of cases where prosopagnosic individuals who could not accurately identify facial identities could accurately identify facial expressions. Calder and Young's review suggests that there is a lack of evidence for the requisite double dissociation (examples of preserved accurate face identity perception in the absence of accurate facial expression perception, and the converse) that would be required to support a strong dissociation between identity and expression routes. Instead they suggested that a Principal Component Analysis (PCA) technique which can look at the relationship between physical image cues that convey facial identity and expression may be a useful approach to use when attempting to understand which facial cues are differentially important for the perception of facial identity or expression.

Based on the line of reasoning of Calder and Young (2005), this thesis aims to better understand the physical image characteristics that underpin the perception of facial identity and expression. Each study manipulates or measures either facial shape or surface properties in a quasi-independent manner and measures the changes in perception of facial identity or expression as a result of these manipulations.

1.2 The physical properties of face images

Many photographic face images typically used in current studies of face perception (and throughout this thesis) consist of a large number of pixels that convey specific intensities, grey levels or RGB values. In terms of image properties, face images contain physical image information at different spatial scales such as low spatial frequency information (which conveys global characteristics of an image such as orientation and proportion) and high spatial frequency information (which coveys local image detail such as edges and line boundaries) (Bar, 2004). These physical facial image properties can be systematically measured and manipulated using special computer based software.

Over the last 40 years computer image manipulation software such as Psychomorph (Tiddeman, Burt, & Perrett, 2001) has been developed to allow the manipulation of facial image properties in a number of different ways. This process typically involves placing a number of fiducial points (see figure 1.1 left image panel) on facial landmarks common to all facial images in a process known as delineation. The position of these fiducial points serves two purposes; firstly, the points define the 2D shape of the face image, secondly, they also allow tessellation (the placing of an mesh on the image, see figure 1.1 right panel) so that the image can be manipulated in a controlled, accurate manner (Bruce & Young, 2012). By delineating image properties in this way a number of image transformations can be carried out such as; facial averaging (e.g. averaging a facial expression from many examples of a facial expression) , facial morphing (increasing or decreasing intensity of a facial cue along a continuum) or facial caricaturing (e.g. exaggerating image differences).

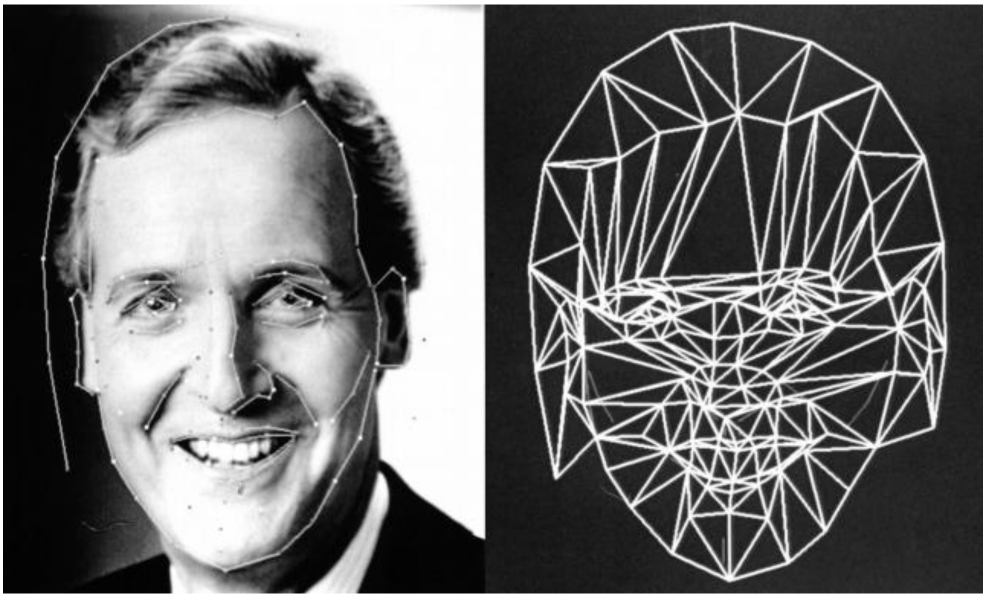

Figure 1.1: Image taken from Bruce and Young (2012), originally developed by David Perrett and the University of St. Andrews. (left) shows an example of a delineated face with major facial landmarks marked by white dots superimposed onto the image. (right) shows what the underlying tessellated version of the image in the left panel looks like. This mesh like image segregates the image into small triangles allowing specific and accurate manipulation of each pixel contained within the tessellation.

Computer manipulation software such as Psychomorph therefore typically decomposes face images into two basic types of visual information: 2D shape (feature positions, edges) and surface properties (surface reflectance, pigmentation) (Bruce & Young, 1998). The fiducial positions therefore represent the image's shape, and its surface properties can be defined as the pattern of pigmentation and brightness contained in the tessellated regions of a face image.

This separation of shape and surface properties is achieved at the level of the 2D image. In practice, faces have a three-dimensional shape that includes information about depth and shading produced by relative occlusion and shape from shading cues; for example the eyes are generally more occluded than the tip of the nose. So there is usually covariation between the shape and surface properties of a facial image, but within this constraint Psychomorph allows facial shape and surface cues to be manipulated quasi-independently from each other in each of the aforementioned facial image transformations and has been successfully applied in a number of areas.

This thesis uses this approach to facial image manipulation and measurement to understand the relative roles of shape and surface information in different aspects of face recognition; namely the perception of facial expression and facial identity. There is clear evidence that shape and surface cues contribute to the perception of a number of other facial characteristics including gender (O'Toole et al., 1998), age (Burt & Perrett, 1995), race (O'Toole, Deffenbacher, Valentin, & Abdi, 1994), attractiveness (Vernon, Sutherland, Young, & Hartley, 2014), and dominance (Torrance, Wincenciak, Hahn, DeBruine, & Jones, 2014). The following introduction will outline and analyse the contribution shape and surface properties are thought to make to facial expression and identity judgements and whether this is reflected in the neural representations of these image properties.

1.2 Facial cues to expressions of emotion

1.2.1 Early evidence of the importance of shape as a cue to facial expression

One aim of this thesis is to clarify the importance of shape and surface information in facial expression perception. Until now expression perception has been thought to depend heavily on shape information (Bruce & Young, 2012; 1998). This viewpoint has arisen from the lack of a substantial effect of contrast reversal (conversion into photo negatives) on accurate identification of expressions (White, 2001; Harris, Young, & Andrews, 2012) when compared to the markedly deleterious effect it has on perception of identity (e.g. Galper, 1970; Kemp, Pike, White, & Musselman, 1996). This is thought to happen because the facial feature edges and shapes that that convey facial expression are spared whilst the surface information crucial to facial identity is impaired (Bruce & Young (2012). This contrast negation paradigm has been widely used to understand which physical cues are vital to specific facial judgements. For example the decrease in facial identity recognition seen in photographic negatives suggests that the facial surface cues which are disrupted during contrast negation (i.e. pigmentation pattern and brightness) are vital to the recognition of identity.

This well replicated contrast negation effect has also been utilised to study expression perception. It provides a method of testing the extent to which facial expression perception depends on surface pigmentation and brightness. White (2001) tested the effect of contrast negation on matching accuracy of a series of simultaneously presented facial expressions. In a series of expression matching experiments participants responded to pairs of simultaneously presented contrast positive or contrast negated faces that differed in either identity, expression, both, or not at all. Interestingly, although distinguishing expression in contrast negated faces occurred more slowly and slightly (but not significantly) less accurately, there were far more errors and longer reaction times for identity judgements. An additional important subsidiary finding of the study was that changes in identity modulated the ability to detect expressions; this interaction suggests that although the primary cues conveying expression and identity may be different, the two processes may not be completely independent of each other.

Numerous replications (e.g. Bruce & Young, 1998; Magnussen, Sunde, & Dyrnes, 1994; Pallett & Meng, 2013; Harris, Young, & Andrews, 2014a) have supported White's (2001) findings by showing that correct expression identification is possible when face images are contrast negated. In a similar vein, prior research investigating how recognisable line drawings (which provide only edge based shape information) of facial expression are, found that facial expression could be identified from line drawings (Etcoff & Magee, 1992; Kirita & Endo, 1995; Magnussen, Sunde, & Dyrnes, 1994; Mckelvie, 1973). More recently White & Li (2006) found that blurring an image of facial expression and therefore removing high spatial frequency edge information significantly impaired judgements of facial expression. This convergence of evidence has been taken to indicate that edge based face information that reflects shape changes is the key visual cue for accurate expression identification.

These findings are important as they show two main things; firstly, the difficulty in recognising the identities of contrast negated faces doesn't arise from inability to detect their expressions, as had been previously suggested by Galper (1970) and Galper & Hochberg (1971). Secondly, that surface properties thought to be key in conveying identity information don't seem to be as crucial in conveying facial expression. White's results neatly demonstrate that the effects of contrast negation seem to be largely confined to facial identity rather than expression judgements and that different image property cues may be used for facial identification or expression processing.

1.2.2 The relative contributions of shape and surface properties to facial expression perception

Image measurement based methods have shown that expression perception appears to be heavily reliant on facial shape properties. The PCA method mentioned earlier allows the extraction and separation of different facial cues that might code for different facial judgements (Calder & Young, 2005). It achieves this by reducing the dimensionality of a complex number of information points in a number of samples to a simpler set of components that code for systematic variations in the data. For example if there are 10000 pixels with corresponding pixel intensities that represent a facial image, the aim of the PCA is to reduce the number of underlying dimensions whilst still being able to accurately represent the important information in the face image (Burton, Bruce, & Hancock, 1999). This technique was originally successfully adopted by Kirby & Sirovich (1990) and Turk & Pentland (1991) in the computer vision face recognition literature. Subsequently Burton, Bruce & Hancock (1999) found that it provides an elegant method of linking physical image properties to perceptual responses by human observers to face images. An additional advantage of this method is that the separation of cues is carried out in an unsupervised manner based on objective measurements and statistical analyses of the face image (Calder & Young, 2005).

One such PCA analysis was carried out by Calder, Burton, Miller, Young, & Akamatsu (2001) who decomposed the Ekman & Friesen (1976) Pictures of Facial Affect into their principal components. It was found that Principal Components that coded facial shape could code aspects of both facial identity and facial expression. Surface based cues on the other hand, mainly coded facial identity. From this evidence, it appeared that facial shape is the main cue that contributes to the perception of facial expressions. However, the contribution of facial shape information to expression perception must be understood in the context of human perception.

The much studied 'composite effect' (Young, Hellawell, & Hay, 1987) has proven to be important for representations of facial expression as well as facial identity in human observers. Calder, Young, Keane, & Dean (2000) repeated the 'facial composite effect' experiment originally carried out by Young et al (1987) testing holistic perceptions of facial identity. Calder and colleagues created aligned and misaligned composite expressions of emotion by combining face top and bottom face halves of different facial expressions. Results showed that participants were slower when accurately identifying expressions in aligned composite than misaligned composite expressions. Thus suggesting that observers integrate both top and bottom halves of the composite facial expression. Complementary findings were reported by Butler, Oruc, Fox, & Barton (2008) who used perceptual adaptation to test whether participants would show expression aftereffects to facial expression features presented in an incorrect configuration as well as to those presented in the correct configuration. Results showed that expression aftereffects only occurred in conditions where schematic images of facial expressions were presented in the correct configuration. This suggests that global rather than local shape cues are important in representations of facial expression. Taken together these findings suggest that feature shape configuration is an important cue to facial expression perception.

Although feature shape is undoubtedly critical to the perception of facial expression subsequent studies have shown that the importance of facial surface information has been underestimated. Whilst White (2001) used an expression matching experimental paradigm to show that contrast negation exerts little effect on expression identification, more recently Benton (2009) shows that surface pigmentation plays an important role in expression perception. Using the face adaptation paradigm (Clifford & Rhodes, 2005; Webster, Kaping, Mizokami, & Duhamel, 2004) results showed that a significantly decreased expression adaptation effect in observers when probe and test faces were show in different contrast polarities. Similarly Pallett & Meng (2013) found that although expression matching was not impaired when images were contrast negated, there was a marked decrease in perceptual encoding (through adaptation) to contrast negated facial expressions. This tempers prior assumptions which suggested that facial expression information is carried almost exclusively by edge (shape) information, as it is clear that contrast negation (which disrupts surface properties) clearly degrades the expression adaptation effect size. Taken together, these findings suggest that there are neural populations involved in encoding representations of facial expressions that are sensitive to facial surface information.

For this reason an important aim of this thesis is to clarify whether expression judgements rely on shape information alone, or whether surface information can also be used by the observer. Chapter 3 addresses this using an expression categorisation paradigm, where observers attempted to identify the expression in images where either facial shape or surface expression information has been degraded. A subsequent experiment in Chapter 3 also investigates the presence of competing shape and surface cues from different expressions in order to understand which cue is dominant. The final experiment in Chapter 3 uses contrast negation to test claims that negation has no significant effect on expression identification. Following this, Chapter 4 looks at neural responses by testing whether correlated patterns of perceptual similarity and neural response to expression are based on shape or surface based image properties.

1.3 Image cues that convey facial identity

1.3.1 Evidence for the importance of surface image properties in facial identity perception

The PCA method discussed earlier with regard to facial expression has also been used to model the degree of independence between representations of facial identity and expression (Calder & Young, 2005). Numerous studies have shown that there is sufficient independence between principal components that separately model facial identity and facial expression (Calder et al., 2000; Calder et al., 2001), with some overlap between principal components (Calder & Young, 2005). With the independence of these principal components in mind, it is reasonable to suggest that perception of facial identity is somewhat dependent on facial cues that are not as important in facial expression perception (Calder & Young, 2005).

PCA methods have also been used to understand how perception of facial identity is dependent on the statistical properties of face images primarily conveying facial identity. Early work by Hancock, Bruce, & Burton (1998) found that ratings of similarity and memory for unfamiliar faces could be predicted successfully by either lower level image properties or a PCA of the pixel based image information. Subsequently Burton, Bruce, & Hancock, (1999) extended the PCA method to clarify which image properties are most salient in the perception of familiar faces. Burton and colleagues (1999) found that a PCA based on pixel based facial surface information (where all faces had been fitted to the same shape template) could correctly identify familiar faces successfully and more efficiently than a PCA combining both facial shape and surface information.

When considering important facial cues to facial identity, classic findings in the face processing literature such as the composite effect (Young et al., 1987) and part-whole effect (Tanaka & Farah, 1993) have been taken to show that holistic face representation (but not absolute feature position, as shown by Hole, George, Eaves, & Rasek, 2002) and feature shape convey important facial identity information. In addition to shape, the importance of surface properties or pigmentation pattern has been show in a number of studies (Bruce & Langton, 1994; Vuong, Peissig, Harrison, & Tarr, 2005; Russell, Sinha, Biederman, & Nederhouser, 2006). The importance of surface properties for representations of facial identity is reflected in the disruptive effect of contrast negation on facial identification. This effect forms an interesting counter-example to the typical resistance of familiar face recognition to effects of many other types of image degradation (Bruce & Young, 2012) and evidence suggests that contrast negation disrupts recognition of faces to a greater degree than other types of visual stimulus (Vuong et al, 2005; Nederhouser, Yue, Mangini, & Biederman, 2007). Indeed the contrast negation effect is important in the face processing literature as it is a characteristic that distinguishes face recognition from basic level object recognition (Liu, Collin, Burton, & Chaudhuri, 1999). The contrast negation effect extends beyond identity recognition alone, gaze perception is also shown to be disrupted when faces are contrast negated (Sinha, 2000; Yoshizaki & Kato, 2011).

Early studies using contrast negation have shown that pigmentation and surface brightness properties are vital to the accurate percept of individual identity (e.g. Bruce & Langton, 1994; Kemp, Pike, White, & Musselman, 1996). Galper (1970) found that contrast negating a face image markedly reduces identity recognition rate to around 50%. It should be noted that although successful facial identification is reduced to around 50% (a marked impairment) the fact that these faces can be recognised to some extent also shows some potential role for facial shape in representations of facial identity (Russell, Sinha, Biederman, & Nederhouser, 2006; Troje & Bülthoff, 1996). It should be noted, however, that surface cues become more salient as faces become more familiar to the point where facial surface becomes the primary visual cue (Burton, Jenkins, Hancock & White, 2005 ; Russell & Sinha, 2007). Despite this though ,the contrast negation effect is highly disruptive of identity recognition both for familiar and unfamiliar faces (Galper, 1970; Galper & Hochberg, 1971; Phillips, 1972).

Luria & Strauss (1978) provided a further test of the contrast negation effect by investigating whether contrast negating face images led them to be perceived differently by observers. Using eye tracking Luria & Strauss's findings indicated that observers gaze patterns were different for contrast negated than contrast positive faces. Specifically, in the contrast negated condition participants spent longer viewing external facial features such as the outer cheeks, ears and hair rather than internal facial features. From this it is evident that contrast negation greatly disrupts properties crucial to the perception of a facial identity and forces a shift to the use of less informative external feature shape cues. From this early work it was apparent that when viewing contrast negated faces, the visual system does not perceive them accurately or efficiently.

The perception of contrast negated faces is interesting as contrast negation does not violate all properties of a normal face stimulus (e.g. feature spacing and 2D feature shape remain the same). An influential study by Kemp et al. (1996) investigated which surface based cues were most disrupted when faces were negated in different ways. A series of experiments were carried out testing recognition accuracy for faces which had hue, brightness or both hue and brightness reversed. Results showed that face images with negated luminance but normal hue saw a reduction in recognition accuracy, whereas faces with negated hue and normal luminance showed no recognition accuracy decrease. Kemp suggested that this demonstrated that of two known components of image negation (hue and brightness), it is reversal of brightness cues that disrupts cues to identity. Kemp and colleagues suggested that the difficulty in recognising contrast negated faces is due to the reversal of 3D shape from shading cues. Kemp argued that although both hue negation and brightness negation manipulations disrupt facial pigmentation patterns, it is only brightness negation that disrupts shading cues. This is because patterns of shading in a brightness reversed face would be uncommon and unnatural, making the extraction of correct 3D shape from shading identity information difficult. The importance of shape from shading as a cue in perception of all visual objects and their features has been argued before (e.g Ramachandran, 1988) and according to this explanation any image transformation (such as negation) that successfully disrupts shape from shading could be expected to produce a recognition deficit. Kemp and colleagues explanation that 3D shape from shading information is vital to identity recognition will now be expanded.

1.3.2 Is shape from shading a useful cue for facial identity perception?

It is clear that facial identity information is contained in natural facial surface characteristics. Although the findings of Kemp et al. (1996) led them to posit that 3D shape from shading cues are key to facial identity, earlier work by Bruce & Langton (1994) investigated whether the contrast negation effect could be replicated in faces devoid of pigmentation. Face stimuli were created using a 3D laser scanner that produced faces with only veridical shape information and no surface detail. Results indicated that when face stimuli were created without surface properties, contrast negation didn't produce any significant reduction recognition in accuracy performance. Bruce & Langton (1994) suggested that contrast negation may therefore exert its disruptive influence because it reverses the brightness of important pigmented regions of the face, so that light regions of skin become dark, dark hair becomes light etc. They suggest that these key textured patches are manipulated to a greater extent under contrast negation conditions as they are completely reversed whereas important shape information such as edges remains largely the same. This view opposes the explanation given by Kemp and colleagues. (1996) and provides a credible alternative explanation of which specific facial surface properties are important in conveying facial identity.

More recently facial shape from shading cues have been investigated from the perspective of naturalistic image lighting properties. Liu et al. (1999) used similar stimuli to Bruce and Langton (1994) and tested the contrast negation effect when manipulating lighting direction. They did this because in naturalistic everyday settings most objects including faces are lit from above, whereas contrast negating makes face stimuli appear to be bottom lit. Liu and colleagues suggested that it may be this unnatural lighting that disrupts our ability to extract identity information from contrast negated faces. Results showed that the source of lighting does affect perception of negated and non-negated faces. Specifically, contrast negated faces are better recognised when bottom lit (appearing normally lit), whereas normal contrast faces are better recognised when lit from above.

Liu and colleagues' findings suggest that at least part of the contrast negation effect seems to arise from the violation of natural lighting conditions. They also show that by maintaining a perception of natural lighting conditions, even though the photometric qualities of an image have been massively affected, high face identification rates can be restored. Furthermore, by removing facial pigmentation but maintaining a good level of facial identity recognition Liu and colleagues show that 3D shape from shading cues do to some extent convey useful facial identity information. Consequently the finding that contrast negation effects can be reversed by simply changing light source direction shows that facial shape from shading cues are sensitive to contrast negation.

1.3.3 The relative roles of shape and surface information for facial identity perception

Bruce and Langton’s (1994) hypothesis that surface pigmentation plays an important role in face identification has been elaborated on in a series of experiments carried out by Sinha, Russell and their colleagues. In one experiment Russell et al. (2006) investigated whether impairing facial shape or surface properties caused greater identity recognition deficits. Three main facial cues (shape, pigmentation or both) were manipulated to produce six experimental conditions. Stimuli consisted of faces with either averaged pigmentation and unique shape , averaged shape and unique pigmentation or unique shape and pigmentation which were viewed in either positive or negative contrast. Results showed that identification accuracy was significantly worse in contrast negated conditions when pigmentation was the uniquely identifying feature, supporting previous assertions that surface properties are sensitive to contrast negation. Results also demonstrated that shape could convey a significant amount of identity information, as evidenced by the relatively high recognition rate when shape was the unique identifier. However, surface properties appear more salient to the observer as evidenced by high recognition accuracy in conditions where surface information was the unique identifier. These findings suggest that disrupting facial surface properties has a greater disruptive effect on identification.

Russell et al's (2006) findings might only apply to unfamiliar faces as familiar face stimuli weren't tested. To extend their approach to familiar faces Russell & Sinha (2007) carried out a conceptually similar experiment using familiar faces. Stimuli were created in the same manner as those used by Russell et al. (2006) however the faces were personally recognisable to each participant and there were no contrast negation conditions. Results showed that recognition accuracy was marginally higher in conditions where facial pigmentation (surface properties) was the unique identifier rather than facial shape. Taken together these studies neatly demonstrate that surface properties appear more crucial and salient for representations of both familiar and unfamiliar facial identity.

Considering these findings, a theme begins to emerge in the literature that the importance of the role of surface properties in facial identification was originally underestimated. Despite this it remains difficult to speculate as to which cue is more important in face recognition as a growing number of studies suggest that shape and surface properties are approximately equally useful for face recognition (O'Toole, Vetter, & Blanz, 1999; Jiang, Dricot, Blanz, Goebel, & Rossion, 2009), although the aforementioned studies by Russell & Sinha provide convincing evidence that in the case of familiar face recognition, facial surface cues are of greater importance than previously thought. Based on these findings Chapter 5 tests whether there is a difference in neural sensitivity to facial identity conveyed by shape or surface information. To test this, a multivariate patterns analysis was performed on neural responses in face responsive regions to faces where identity was conveyed by either unique surface or unique shape properties.

1.4 Interactions of shape and surface: the interesting case of contrast chimeric faces

When considering the relative roles of facial surface properties in facial identity perception a recent study testing contrast chimeras has added a new dimension to the debate. Gilad, Meng, & Sinha (2009) demonstrated that neither of the theoretical explanations given for why contrast negation disrupts facial identity perception previously can be entirely correct. Gilad and colleagues created 'contrast chimeras' of famous individuals consisting of contrast negated faces in which the eye regions were left in positive contrast. Strikingly, recognition accuracy for these contrast chimeras was around 90% of the recognition rate for normal face photographs. To ensure this effect wasn't entirely driven by the positive contrast eye region Gilad and colleagues also included a condition where they superimposed a positive contrast eye region on silhouettes of heads. This condition displayed very poor recognition rates (13%), showing that the intrinsic recognisability of the eyes alone cannot drive the contrast chimera effect.

Gilad et al.'s findings are remarkable as they show that merely restoring the correct contrast polarity to a small region around the eyes is sufficient to eliminate the classic contrast negation effect on face identification. Yet the major proportion of a contrast chimera image still has reversed luminance values and therefore gives incorrect shape from shading and pigmentation cues for these parts of the face. This is at odds with traditional explanations of contrast negation that suggested either that the recognition deficit arises from disrupted 3D shape from shading information (Kemp et al., 1996) or from the reversal of large patches of pigmentation information (Bruce & Langton, 1994).

Although the traditional views of contrast negation may not be completely correct, they do not rule out the possibility that patterns of pigmentation or 3D shape from shading in the eye region are important for identity recognition. Gilad and colleagues suggested that their 'contrast chimera effect' arises because the eye region plays a critical role in creating a visual representation of a face and a contrast chimera face image maintains the natural ordinal contrast relationships around the eye region. More specifically Gilad et al point out that the ordinal contrast relationships between relatively lower luminance levels in the eye region and higher luminance values of the surrounding cheeks, nose and forehead remain constant across nearly all natural lighting conditions and viewpoints in which we encounter faces.

Gilad and colleagues suggest that contrast chimeras support high recognition because they maintain these ordinal luminance relationships between face regions that are stable and found in normal face images (Sinha, 2002). In their view this is because the eyes provide a stable and generally invariant feature around which to create face representations. Thus when negating the contrast of the whole face, the natural ordinal contrast relationships between the eyes and surrounding face are reversed, and key information in this region cannot be used. Contrast chimeras however maintain the normal lower luminance levels of the eye region and higher luminance values of the surrounding face regions, therefore maintaining the crucial ordinal contrast relationships from which identity can be accessed. This ordinal contrast relationship explanation has received support from Ohayon, Freiwald, & Tsao (2012) who found face patches in macaques produce similar responses to real faces and schematics of faces with correct contrast polarity relationships, whereas they do not to schematics of faces with random contrast polarity distribution.

The contrast chimera effect poses an interesting question for prior studies that used a contrast negation paradigm to probe the relative importance of facial surface information in representations of identity. The main question that arises is why does restoring a very small area of surface information (the eye region) almost restore identity recognition perfectly? This thesis also aims to addresses this question in a series of four behavioural experiments in Chapter 6. Experiment 1 tests whether the contrast chimera effect only arises when the eyes are returned to positive contrast or whether returning other non eye face regions can support facial identification. Experiment 2 tests whether the eye region alone can support good rates of recognition or whether the shape information contained in contrast negated non eye regions also aids the perception of facial identity in a contrast chimera. Experiment 3 tests whether the contrast chimera effect depends on a holistic representation of facial identity. Finally Experiment 4 aims to test Sinha's (2002) and Gilad and colleagues (2009) claim that the contrast chimera effect maintains the perception of identity because it keeps constant contrast polarity relationships in the eye regions.

1.5 Cortical networks involved in face perception

1.5.1 Neural models of face processing

The most well-known cognitive model of face perception was proposed by Bruce & Young (1986). Over the past 30 years attempts have been made to extrapolate the functional components identified by Bruce and Young (fig 1.1a.) and identify each one within the brain. A variety of techniques have been used to identify face selective neural populations including single cell recordings (Baylis, Rolls, & Leonard, 1985; Leonard, Rolls, Wilson, & Baylis, 1985) and event related potentials using methods such as EEG and MEG (Bentin, Allison, Puce, Perez, & McCarthy, 1996; Bötzel, Schulze, & Stodieck, 1995; Jeffreys, 1996). However, the most widely used neuroimaging method is now functional magnetic resonance imaging (fMRI) which has allowed the measurement of blood flow assumed to reflect neural responses in face selective neural populations in vivo in healthy participants. The advantage of fMRI is that it is a non-invasive method of testing neural response to a stimulus with a millimetre spatial resolution, allowing us to make inferences about specific regions of cortex involved in the perception of a visual stimulus.

Perhaps the most notable of the attempts to create a formalised neural model of face perception has been primarily based on past studies using fMRI and is proposed by Haxby, Hoffman, & Gobbini (2000). Haxby and colleagues identified a number of brain regions (shown in figure 1.2b) that were consistently identified in fMRI studies and neuropsychological case studies linking neural response to the perception of face stimuli. The following sections of the introduction will discuss in further detail how Haxby et al.'s 'core brain regions' are implicated in different aspects of face processing .

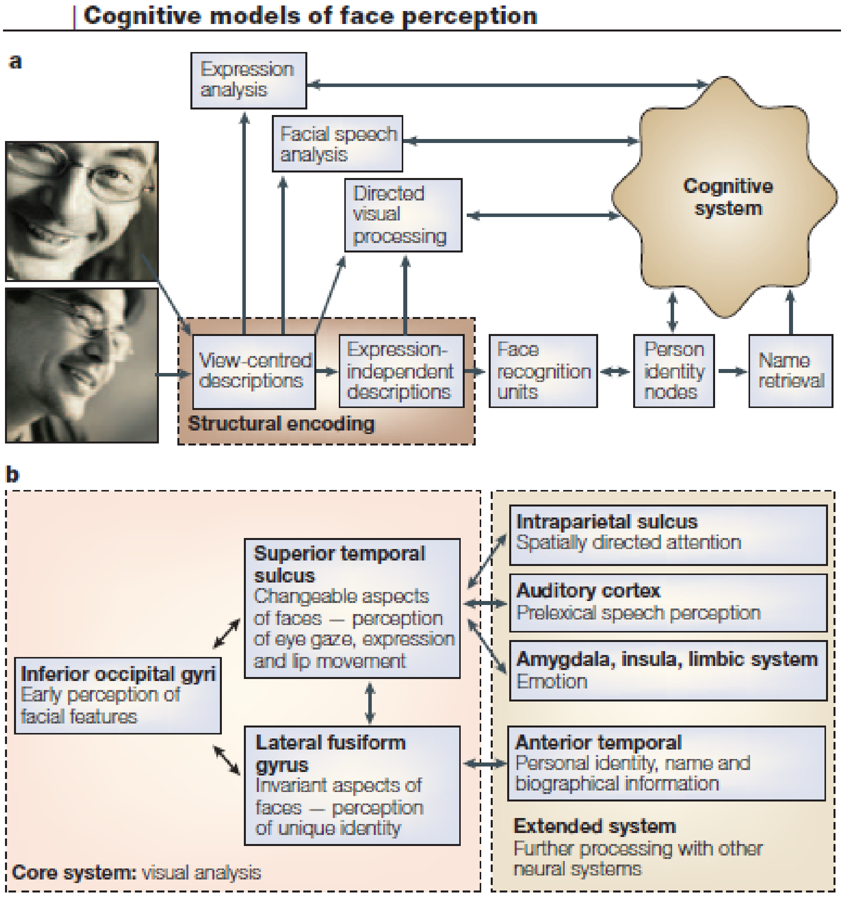

Figure 1.2: taken from Calder & Young (2005). A schematic figure (1.2b) showing how Haxby et al. (2000) identified brain regions involved in aspects of face processing analogous to the nodes originally proposed in figure (1.2a) by Bruce & Young (1986).

1.5.2 The Role of the Fusiform Face Area in face perception

One of the first successful attempts to identify a neural population that showed a preferential response to faces compared to non faces using fMRI was carried out by Kanwisher, McDermott, & Chun (1997). They identified an area that consistently showed a preferential selectivity for faces over other types of visual stimulus and named this region the Fusiform Face Area (FFA). This face selectivity in fusiform gyrus has been replicated in a number of studies (O'Craven & Kanwisher, 2000; Andrews & Schluppeck, 2004) and does not appear to be influenced by any task demands (e.g. passive viewing or one back) involved with viewing the faces (Berman et al., 2010). Since this first identification of the FFA a number of studies have sought to demonstrate that it is an area that primarily plays a role in perception of facial identity (George et al., 1999; Rotshtein, Henson, Treves, Driver, & Dolan, 2005; Haxby et al., 2000) or the judgment of face-like semblance of an object (Andrews, Schluppeck, Homfray, Matthews, & Blakemore, 2002; Andrews & Ewbank, 2004; Meng, Cherian, Singal, & Sinha, 2012).

The proposed role of the FFA as an area primarily concerned with facial identity perception is further supported by evidence that right side lateralised FFA lesions have been found in prosopagnosic patients (Barton, Press, Keenan, & O'Connor, 2002). Additionally, fMRI BOLD response abnormalities have been found in the FFA in prosopagnosic individuals (Schiltz & Rossion, 2006; Dricot, Sorger, Schiltz, Goebel, & Rossion, 2008; Steeves et al., 2006). Further evidence supporting the role of the FFA in judgements of facial identity comes from Grill-Spector, Knouf, & Kanwisher (2004), who found that activation in the FFA was correlated to the successful identification of and discrimination between individual faces. Moreover it was found that this effect was specific to faces and did not generalise to other forms of within category discrimination (e.g. between types of car, etc.).

An important aspect of understanding how facial identity is represented in the FFA is whether the FFA is, as some have suggested (e.g. Kanwisher et al., 1997; Reddy & Kanwisher, 2006), a self-contained module consisting of face selective neurons or whether FFA activation reflects a larger, more distributed pattern of activation including neighbouring regions (e.g. Haxby et al., 2001; Hanson, Matsuka, & Haxby, 2004). The latter argument represents a potentially important advance because it opens up the possibility that lower level aspects of a face image are coded in 'informative' neurons across a wide range of face responsive cortex including but not limited to the FFA. Recent developments in multivariate techniques have allowed researchers to test wide ranging patterns of neural response simultaneously. These multivariate techniques are powerful and have already shown differential responses to categories of stimuli from non-preferred categories as well as ones from the 'preferred category' in some brain modules (e.g. O'Toole, Jiang, Abdi, & Haxby, 2005; Hanson & Halchenko, 2008).

Interestingly, although the overwhelming body of evidence suggests that the FFA contains neural populations primarily selective to facial identity, a magnetoencephalography (MEG) study carried out by Lewis et al. (2003) found that the FFA is more sensitive to affective faces when compared to neutral faces. However, it is worth noting that with the MEG method it is difficult to localise this response with confidence to either the FFA or OFA, which are often located closely in fusiform gyrus. Despite this, a number of fMRI experiments have also shown that there is a greater response in the FFA to expressive faces than simply neutral faces (Ganel, Valyear, Goshen-Gottstein, & Goodale, 2005; Ishai, Schmidt, & Boesiger, 2005; Fox, Moon, Iaria, & Barton, 2009; Xu & Biederman, 2010). In fact recently Harry, Williams, Davis, & Kim (2013) used a multi voxel pattern analysis (MVPA) to show that it is possible to find distinct neural patterns of response to facial expressions within the FFA. For this reason, although the FFA is still primarily considered to be an area involved in detecting and processing the identity of faces, it seems that it may play a role in some aspects of processing variant aspects of faces such as expression. For this reason caution should be exercised when interpreting Haxby et al.'s (2000) strict definition of the role of the FFA as a region confined to purely the processing of invariant facial cues such as identity.

Despite this suggestion that the FFA may play some role in the perception of facial expressions, the majority of current face perception research demonstrates that the FFA is a brain region primarily concerned with the identification of the face stimulus and the perception of facial identity. Accordingly it occupies an important position in the Haxby et al. (2000) model as an area primarily involved in the processing of invariant facial cues such as identity. Despite this wealth of research it is important to temper current assumptions about the modularity and specialisation of the FFA as the only brain region concerned with facial identity perception. The recent advent of multivariate pattern analyses has shown that each facial judgement may draw on informative neural populations across a distributed area including regions adjacent to face selective regions.

1.5.3 The Role of the Occipital Face Area in face perception

The Occipital Face Area (OFA) is a face responsive brain region proximal to the FFA although nearer to areas involved in earlier downstream visual processing. The Haxby et al (2000) model places OFA as a part of the network where visual input arrives and early feature analysis is carried out. The model suggests that at this point this information is passed on to regions such as the FFA and STS where higher level facial analyses are made. There is debate about whether OFA input is necessary for further facial processing to be carried out as there is some evidence that FFA activation can occur in those with OFA lesions that prevent OFA activation (e.g. Rossion et al., 2003; Schiltz & Rossion, 2006; Rossion, 2008; Steeves et al., 2006). Despite this, recent evidence suggests that in healthy individuals the OFA carries out the first stage of visual face processing before passing on information to the FFA and STS (Nguyen, Breakspear, & Cunnington, 2013).

There exists clear evidence that there is a strong connectivity and co-activation between the OFA and FFA (Davies-Thompson & Andrews, 2012) suggesting that whatever the role of the OFA is, it carries out some processing of the face stimulus prior to its arrival in higher level regions such as the FFA and STS. This viewpoint has been supported by the findings of Rhodes, Michie, Hughes, & Byatt (2009) who found that the OFA shows sensitivity to changes in lower level facial features such as spatial relations between facial features. This perhaps explains the finding that the OFA shows sensitivity to faces that vary in either expression or identity (Rotshtein et al., 2005; Pitcher, Walsh, Yovel, &

Duchaine, 2007; Pitcher, 2008; Fox, Moon, Iaria, & Barton, 2009), both situations in which spatial relations between facial features can change markedly.

Current evidence thus converges with Haxby's (2000) view of the OFA as the earliest visual face responsive region that carries out local, part based analyses of facial characteristics before passing them on to more anterior fusiform and superior temporal regions (Yovel & Kanwisher, 2004; Pitcher et al., 2007; Schwarzlose, Swisher, Dang, & Kanwisher, 2008; Liu, Harris, & Kanwisher, 2010). These claims are strengthened in a recent study by Arcurio, Gold, & James (2012) showing that the OFA displays a preferential neural response to individual face parts and features that to intact feature configurations whereas the FFA shows no such preference. This mirrors prior findings that suggested that the OFA plays a greater role in the processing of face parts than the FFA (Betts & Wilson, 2010; Nichols, Betts, & Wilson, 2010). As such it is probable that the OFA carries out computations more tied to the facial image properties than the FFA, specifically diagnostic face parts before this information reaches higher level face responsive brain involved in more complex social judgements.

Currently, though, the role of the OFA is the least well understood of the three core regions of the Haxby face perception model. It appears that it is involved in aspects of both facial identity and expression processing but it is unclear to what extent. Although there have been suggestions that the OFA plays some role in facial identity processing, until now role of the OFA in expression perception may have been underestimated. This is because its preference for face parts may lead it to carry out significant processing of facial expressions where individual face parts can still carry significant expression information (McKelvie, 1995; White, 2000) despite generally being represented holistically (Calder et al., 2000; White, 2000) when compared to facial identity which is more purely dependent on holistic face representations as evidenced by the composite effect (Young et al., 1987) and the part-whole effect (Tanaka & Farah, 1993). For this reason the experiments in both Chapter 4 and 5 both seek to shed light on the role of the OFA in face perception. Firstly, Chapter 4 attempts to clarify whether there are consistent patterns of neural response to facial expressions and whether these are linked to perceptual patterns of response to facial expression, and in turn whether these perceptual patterns are based on image features of each expression. Secondly, Chapter 5 seeks to clarify whether the OFA plays a role identity perception and whether it responds to shape or surface based cues to facial identity.

1.5.4 The Role of the Superior Temporal Sulcus in face perception

The Superior Temporal Sulcus (STS) is a brain region that has been shown to play some role in the perception of social cues such as gaze (Haxby et al., 2000; Pourtois et al., 2004; Akiyama et al., 2006) and dynamic facial movements related to gaze and facial expressions (Calvert, Hansen, Iversen, & Brammer, 2001; Beauchamp, Nath, & Pasalar, 2010; Pitcher, 2014).

However, of all the components of facial analyses it is involved in, the anterior STS appears to be primarily implicated in facial expression perception, evidenced by the fact that there is a greater neural response in the anterior STS to facial expression compared to facial identity (Puce, Allison, Bentin, Gore, & McCarthy, 1998; Hoffman & Haxby, 2000; Allison, Puce, & McCarthy, 2000; Puce & Perrett, 2004). This finding is robust and has been supported by a number of studies all showing that the anterior STS shows a marked increase in neural response to facial expressions compared to other facial cues (Winston, Henson, Fine-Goulden, & Dolan, 2004;Calder & Young, 2005; Engell & Haxby, 2007; Ishai, 2008). Up until now however there has not been any clear evidence that links the neural response in the STS to lower level image features of the facial expression image. The previously mentioned Rhodes et al. (2009) study did not find any sensitivity in the STS to facial feature distance.

There also exists some evidence that the posterior STS also responds to facial identity, suggesting that the segregation between expression and identity processing proposed by Haxby et al (2000) is not as clear as first thought. Indeed Winston et al. (2004) found that when they sequentially presented face pairs where facial identity was the same there was a reduced fMRI signal in the pSTS reflecting an adaptation effect in neurons sensitive to facial identity within the pSTS. This reflected earlier behavioural studies that showed interactions between facial identity and expressions in facial aftereffects (Fox & Barton, 2007; Fox, Oruç, & Barton, 2008), tests of perception (Ganel, Goshen-Gottstein, & Ganel, 2004) and recognition (Kaufmann & Schweinberger, 2004). Much like the FFA may play some small role in facial expression, there appears to be some potential role for the posterior section of the STS in the perception of facial identity.

Despite its suggested role in facial identity perception, the main role of the STS seems to be in expression perception. The nature of expression representations in the STS has been elucidated recently by Harris et al. (2012). They aimed to test how facial expressions were represented in different brain regions, specifically, were there both continuous and categorical representations of expressions of emotion in the brain? They found that pSTS regions contained a more continuous representation of facial expressions than the amygdala which held categorical representations of facial expression. Related to this, earlier work by Said, Moore, Engell, & Haxby (2010) found that there was a significant correlation between neural response to 6 expressions (fear, anger, disgust, sadness, happiness and surprise) in the STS and the ratings of how perceptually similar these expressions were. The convergence of these two studies suggests that there is a graded representation of facial expressions of emotion within the STS. However neither Harris and colleagues or Said et al. (2010) could clarify whether these continuous representations of emotional expressions were based on their visual similarity, or on other non visual underlying dimensions of expression similarity such as those suggested by Russell (1980). This thesis aims to build on the findings of Said et al (2010) by testing whether these graded perceptual similarity ratings of expression are based on image properties of the expressive face. Chapter 4 also tests whether the neural response in other core face regions such as the OFA and FFA is correlated to perceptual similarity ratings of expression.

Much like the OFA, the STS has been shown to be sensitive to individual face parts but not the correct configuration of facial features (Liu, Harris, & Kanwisher, 2010). This sensitivity of the STS to specific facial features is thought to be due to the role of the STS in eye gaze perception (Materna, Dicke, & Thier, 2008). It was, until recently, unclear whether the sensitivity of the STS to individual face parts was based on their physical features or the expression information contained within those features. This question was addressed recently by Flack and colleagues (2015) who compared neural response patterns in the right pSTS in a composite face expression task in which participants were required to judge whether there had been any change in sequentially presented images. They found that neural responses in the right pSTS did not reflect holistic expression perception; rather they reflected a pattern where expressions were perceived on the basis of individual features. This adds detail to the Haxby et al (2000) model's view of the role of the STS, and suggests that the pSTS is involved at an early stage in the processing of facial expression in which facial features are represented independently. As such it is possible that there is a significant population of neurons within the STS region that codes feature-based properties of facial expressions of emotion.

To further elucidate this, in Chapter 4 one of the aims is to link the neural response in STS to both perceptual similarity of expressions and an image based properties of facial expression images. Chapter 5 also tests whether there is any shape or surface based neural representation of facial identity in the STS.

1.6 Representations of facial image properties in the cortical face perception network

In early visual brain regions, there is evidence that shape and surface image properties are processed in partially segregated visual areas and pathways (e.g., Tovée, 1996; Grill-Spector & Malach, 2004). Given this evidence and the previously discussed behavioural evidence that facial identity and expression are differentially sensitive to facial shape and surface cues, it is reasonable to attempt to identify whether there are differential neural responses to facial shape and surface information for different judgements (e.g. expression and identity). One of the main aims of this body of work is to better understand how facial shape and surface cues are represented in face responsive brain regions and this question is addressed in Chapters 4 and 5.

The fMRI adaptation paradigm designed by Tootell and colleagues (1998) has been used by Jiang et al. (2009) to probe whether face selective regions sensitive to identity are differentially sensitive to either shape or surface properties. Their adaptation experiment included conditions where faces were shown in sequences that were unchanged, changed in shape, changed in surface or changed in both shape and surface. Jiang and colleagues found that changes in facial shape were the main cue to facial identity in right lateralised face responsive regions (OFA and FFA). In left lateralised face responsive regions, there was sensitivity to both shape and surface cues to identity. Overall this study was successful in showing that neural response regions such as the FFA and OFA respond to physical facial characteristics that are involved in conveying facial identity.

Despite the success of Jiang and colleagues (2009) who showed the sensitivity of these face selective regions to different facial image properties, it is currently unclear whether these neural populations are responding to facial cues that form part of the representation of facial identities or merely to physical facial feature differences. One way to probe this would be to create an experiment with conditions in which either shape is the consistent identity cue or surface is the consistent identity cue. For this reason Chapter 5 employed a multi voxel pattern analysis technique to measure whether changes in facial shape or surface cues to identity elicited a consistent neural response. This allows an advance on the work of Jiang and colleagues (2009) and probes whether a neural population codes specific image based information rather than whether it is merely sensitive to a change in the face image.

One of the questions this thesis aims to answer is the extent to which higher level visual face regions are arranged based on facial image properties. This question lies at the centre of the debate about whether ventral temporal cortex is organised in a modular (Spiridon & Kanwisher, 2002; Kanwisher et al., 1997) or distributed (Carlson, Schrater, &

He, 2003; Cox & Savoy, 2003; Haxby et al., 2001) manner. O'Toole et al. (2005) investigated the extent to which visual representations of categories of different objects (faces and non-face object categories) were mapped in higher level regions in ventral temporal cortex based on their image similarity. They first defined consistent patterns of neural response to each object category and then compared these patterns of neural response to visual based object confusabilities. They found that the confusability of patterns of neural response to objects was highly correlated to image confusability. This suggests that the organisation of brain areas termed 'higher level face regions' may be to some extent dependent on facial image properties.

These findings relate to a suggestion recently posited by Op de Beeck, Haushofer, & Kanwisher (2008) who reviewed a wealth of fMRI literature to better understand whether ventral temporal cortex is organised into discrete, category selective modules or whether it can better be explained by a series of feature selective maps. They concluded that what has up until now appeared to be modular category selective regions in ventral temporal cortex is actually a series of overlapping continuous feature maps. Op de Beeck and colleagues suggest that a nonlinear summation of underlying feature selective maps gives rise to what appear to be category selective modules. When Goesaert & Op de Beeck (2013) attempted to test whether neural responses to facial identity were linked to pixel based image similarity they found that this was only the case in early visual cortex and not higher face selective regions. Although Op de Beeck and colleagues stated that there was as yet no evidence of simple image based similarities underlying these category selective regions, this thesis seeks to test this idea with facial expressions and to understand whether the organisational principles of face responsive brain regions can be linked to the image properties inherent within a face image.

Chapter 4 attempted to link perceptual similarity of expressions of emotion to their neural response similarity to establish whether this is based on either shape or surface image properties. To do this a multi voxel pattern analysis (MVPA) approach was employed to first define whether there are consistent distributed patterns of neural response to facial expressions of emotion. The aim is to understand whether representations of expressions within the STS, OFA and FFA might reflect their image properties and if so whether they are based primarily on shape or surface based image properties. To probe this, perceptual similarity data is first regressed with surface and shape based facial information to understand whether human perception of expression relies on image properties of each expression. Then a regression is carried out to test whether the patterns of neural response matches the perceptual similarity ratings of each expression.

These attempts to understand how facial shape and surface cues are represented in face responsive brain regions may also help elucidate another debate. A review of the neuroimaging literature by Calder & Young (2005) suggested that the original Bruce & Young (1986) and Haxby et al (2001) models of face perception may have provided an oversimplification of the dissociation between different facial cues. For example one of the central tenets of these models is the dissociation between brain regions that are selectively involved in the perception of either facial identity or facial expression. As Calder & Young (2005) point out, in fMRI research this dissociation appears to come from different studies that only test one of the two cues without comparing them directly to each other. With this in mind, this thesis will also provide a test of whether each brain region is sensitive to expression and identity in Chapters 4 and 5 using a sensitive MVPA paradigm. The MVPA paradigm allows the measurement of the full patterns of response in face responsive regions to expression and identity rather than the traditional relative univariate contrast method, which largely identifies areas of differential peak responses.

1.7 Thesis Outline and Aims

The first section of this literature review has outlined how current work on visual expression perception posits that shape cues are important for expression perception. However as discussed, there has been a lack of consideration of the role of facial surface properties and they may therefore have been underestimated. A main aim of this thesis is to clarify the extent to which representations of facial expression are based on facial surface as well as shape properties. The next section outlined the literature on facial identity perception demonstrating that current evidence supports the role of both facial shape and surface properties as important to the perception of facial identity. However standard explanations about how facial shape and surface cues combine to create the perception of facial identity have been thrown into doubt by the recently discovered 'contrast chimera effect'. In this thesis I aim to understand how the contrast chimera effect can be reconciled with past views on the role of facial shape and surface features in face perception. The next section of the review outlined current theories about how different cortical regions are implicated in facial expression and identity perception. Following this, I outlined the extent to which facial image properties may be represented in different face responsive brain regions and the extent to which different brain regions may response to facial shape and surface properties. This thesis aims to clarify the extent to which the neural response in different brain regions is dependent more on facial shape or facial surface cues to expression and identity. A full outline of the contribution of each Chapter to thesis aims and research questions is outlined below:

Chapter 2 –This Chapter also gives the theoretical background to the different types of multi voxel pattern analysis (MVPA) used in Chapters 4 and 5. Finally this Chapter discusses the variation of the Representational similarity analysis used in Chapter 4, its benefits and uses.

Chapter 3 – This Chapter explores the relative contributions of shape and surface cues to facial expression identification. The experiments in this Chapter provide conflicting shape and surface cues to different expressions and test which cue is preferred more frequently. Using morphs that disrupt either mainly facial shape or surface properties across all expressions, they test whether selectively disrupting either shape or surface cues produces a larger deficit in expression identification. In a final experiment, these images were contrast negated to test whether the image manipulations used in the second experiment selectively disrupt shape or surface cues.

Chapter 4 – having explored the relative importance of facial shape and surface properties in representations of facial expression in the previous Chapter, this Chapter sought to identify whether these image properties underpin the neural coding of expression. Specifically, is the neural representation of facial expression based on a continuous perceptual similarity that reflects image-based information? First, perceptual similarity responses were correlated to neural patterns of response to facial expressions of emotion. Following this the perceptual responses were correlated to expression similarity models based on shape or surface property similarity to test whether the perceptual similarity of facial expression is based on image similarity.

Chapter 5 – This Chapter investigates the relative contributions of facial shape and surface information to the neural representation of familiar identity. The fMRI Experiment in this Chapter investigates whether there is sufficient information contained in either only the shape or only the surface representations of facial identities to produce reliable patterns of neural response in face responsive brain regions.

Chapter 6 – This Chapter investigates the recently discovered contrast chimera effect which has proved difficult to explain by traditional accounts of facial identity perception. A wealth of contrast negation studies have shown that reversing the contrast polarity of an image severely disrupt the recognition of facial identity. However the contrast chimera effect demonstrates that simply restoring the eye region to the correct contrast polarity substantially restores identity recognition. This Chapter carried out a series of experiments to test the contrast chimera effect and attempt to understand why having a correct contrast polarity in the eye region is so vital to maintaining a useful identity percept.

Chapter 2 - Advanced functional magnetic resonance imaging techniques: linking neural responses to behavioural and stimulus measures

2.1. What can multi voxel pattern analysis approaches to functional magnetic resonance imaging tell us about the functional organisation of the brain?

Traditional functional Magnetic Resonance Imaging (fMRI) techniques test whether a blood oxygen level dependent (BOLD) response is evoked in a population of neurons by a particular stimulus or task (Huettel, Song, & McCarthy, 2009). In this paradigm the neural response evoked by two separate stimuli or experimental conditions is sampled and the neural response for each condition is contrasted in each voxel using a general linear model (Friston et al., 1995) to give a measure of the univariate neural response (Huettel et al., 2009). This univariate method allows identification of voxels tuned to a stimulus or cognitive state, and has been used to map the functional role of each brain region. The advantage of this method is it provides a straightforward measure of the differential response in each brain voxel that may be linked to different cognitive mechanisms.

Although the univariate fMRI sampling method has been used since the advent of fMRI to successfully map brain function, there are two main drawbacks to this method. Firstly, univariate analyses limit the understanding of the dimensional structure of the neural response to a stimulus (Norman, Polyn, Detre, & Haxby, 2006). This is because the univariate method is limited to testing whether each voxel (and the neurons contained within) can pass a statistical threshold in response to a stimulus. For example there is univariate adaptation-fMR evidence that neurons in both pSTS and FFA respond to facial identity (Winston et al., 2004), so while it is clear that these distinct populations of neurons respond to facial identity, it is not known whether they respond to different aspects of facial identity or whether they form part of a larger scale neural response to the same cues to facial identity. The Multi Voxel Pattern Analysis (MVPA) method means that neural response patterns which are widely distributed across brain areas can be compared simultaneously to test whether they form part of patterns that can be used to discriminate stimuli. Furthermore recent developments mean that MVPA can be combined with Representational Similarity Analysis (Kriegeskorte, Mur, & Bandettini, 2008b) to allow the neural response pattern in the brain to a particular kind of stimulus to be compared with models that represent theoretical or physical aspects of the stimulus.

A second limitation of the univariate fMRI method is that does not allow an understanding of how distributed patterns of neural to a particular stimulus might be (Haynes & Rees, 2006). This is because univariate analyses test each voxel individually, meaning that only voxels that pass a statistical threshold are considered to be responsive to a stimulus. This means that although the univariate method can be employed to test whether neurons within a voxel are differentially tuned to a particular stimulus, it is not clear whether combinations of voxels that are distributed more widely across the brain are also potentially involved in the neural representation of that stimulus. Multi-voxel pattern analyses on the other hand can be used to test whether patterns of neural response across the whole brain may be involved in the representation of a stimulus or cognitive state. For this reason attempts have been made to move away from traditional univariate analyses alone and more towards multi-voxel pattern analysis methods that aim to decode mental states and neural response profiles based on consistent neural activity across a number of distributed spatial locations in the brain.

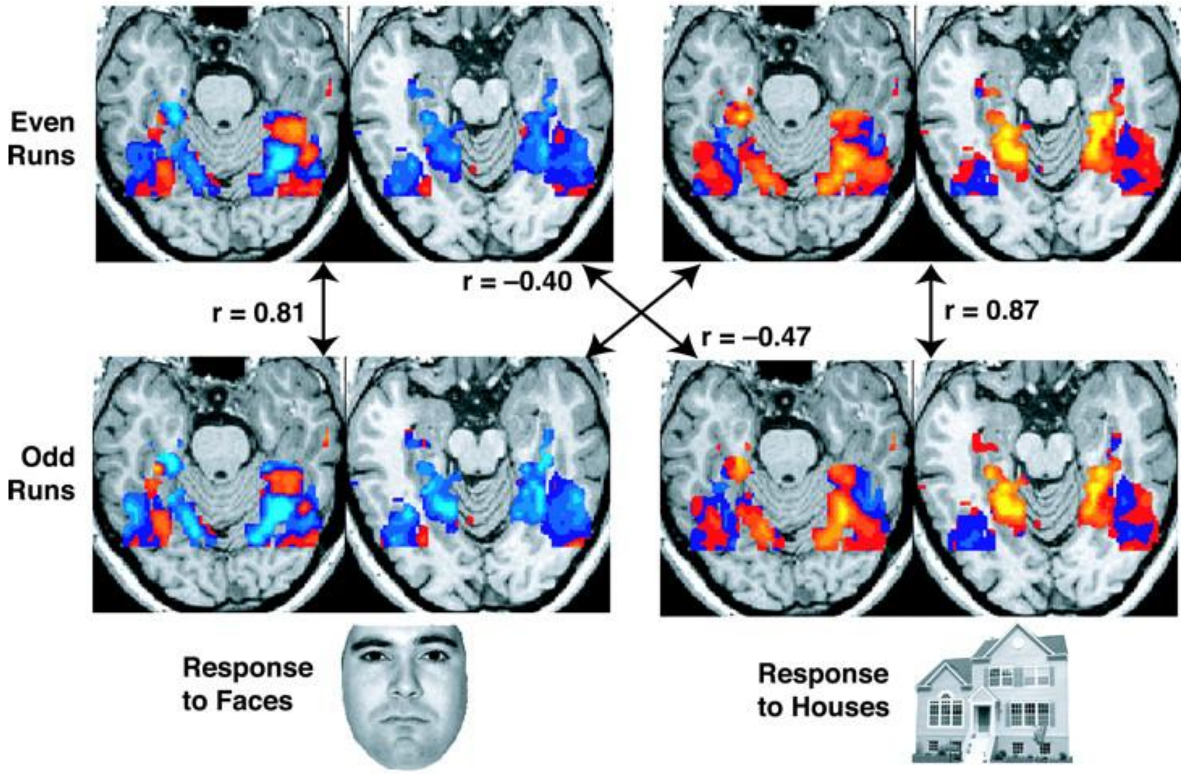

Figure 2.1: Image taken and altered from Haxby et al (2001) demonstrating how multi-voxel neural responses are measured and used for object category discrimination. The figure displays the cross validation correlation technique; experimental blocks are split into odd and even runs for cross validation. Odd and even runs where the presented stimulus was from the same category (e.g. faces) usually have similar response profiles and therefore higher correlation coefficients across voxel responses than odd and even runs from differing categories (e.g. one run of faces and one run of houses). All possible combinations of odd vs. even run correlations are carried out and yield Pearson's r correlations which are then significance tested against a chance level.

In contrast to traditional Univariate GLM based approaches, Multi Voxel Pattern Analysis (MVPA) aims to measure the reliability of neural response profiles evoked by a stimulus. In MVPA paradigms blocks of each stimulus are shown and evoked neural response is measured in the same way as in a univariate analysis. Next, a normalisation technique is typically used to remove any neural response that is general to all categories of stimulus shown. This is can be done by calculating the average neural response to all conditions and subtracting it from each condition in turn, thereby leaving a neural response in each voxel that is specific to that stimulus condition; this removes uninformative or 'noisy' voxels from the cross validation procedure (Haxby et al., 2001).

From here there are a number of ways to carry out cross validation, however Haxby and colleagues split their stimulus condition blocks into odd and even runs and correlations were carried out between patterns of neural response in these runs to yield Pearson's r values for each comparison (see figure 2.1). Within stimulus category correlation coefficients (e.g. even runs faces compared to odd runs faces) are then compared to between category conditions (e.g. odd runs of faces compared to even runs of non faces) and it is determined whether the within category condition correlation is greater than the between category condition correlations. If the within category condition correlation coefficient is significantly greater than the between category correlation coefficients then it is suggested that the population of neurons evoked in response to the stimulus can be reliably assumed to encode characteristics of that stimulus (Haxby et al., 2001). Using this cross validation procedure, experimenters can be said to' decode' neural responses to any of a number of stimulus categories. The term 'decode' refers to finding a consistent neural response to one stimulus category (e.g. faces) consistently in a cortical region (e.g. in the FFA) across a number of runs when compared to a stimulus from a different category (e.g. objects) in that same cortical region. By finding this consistent neural response to a stimulus category in a cortical region it is inferred that any future similar pattern of response in the same cortical region will have been evoked by a stimulus from the same category.

This MVPA paradigm designed by Haxby and colleagues (2001) and more advanced patterns classification analyses have been used to decode a number of mental states in the visual domain such as higher level object discrimination (e.g. Haxby et al., 2001), low level image features of objects and faces (e.g. Rice, Watson, Hartley, & Andrews, 2014), dominance in binocular rivalry (e.g. Haynes & Rees, 2005) and striped pattern discrimination (e.g. Kamitani & Tong, 2005). The prior success of the MVPA paradigm in discriminating facial expression (Said, Moore, Engell, & Haxby, 2010) and facial identity (Nestor, Plaut, & Behrmann, 2011) demonstrates this as an appropriate method to use to test patterns of neural response; therefore Chapters 4 and 5 use an MVPA paradigm to decode facial expression and identity as part of their design respectively.

Although the correlational MVPA method used by Haxby and colleagues (2001) is the most commonly used measure, other variants exist including the Leave-One Participant Out (LOPO) MVPA paradigm (Poldrack, Halchenko, & Hanson, 2009; Shinkareva et al., 2008; Kaplan & Meyer, 2012). This method differs in that instead of cross validating neural response profiles across runs within participants, it compares neural response to all stimulus conditions between an individual participant and the rest of the group. This is done by creating a group average (minus one participant) of the neural response to each condition and comparing it to the neural response profile in the individual that has been left out of the group analysis for those same conditions. This analysis is repeated once for each participant, each time removing that participant from the group analysis. This then yields a number of correlation coefficients for each condition comparison equivalent to the number of participants you have. Significance testing is then carried out in the same way as Haxby and colleagues (2001) by comparing correlation coefficients for neural responses to within category vs. between category conditions. In effect, the LOPO method measures the similarity between participants (i.e. between each participant and the rest of the group) whereas the odd and even runs correlation is a purely within-participants measure.

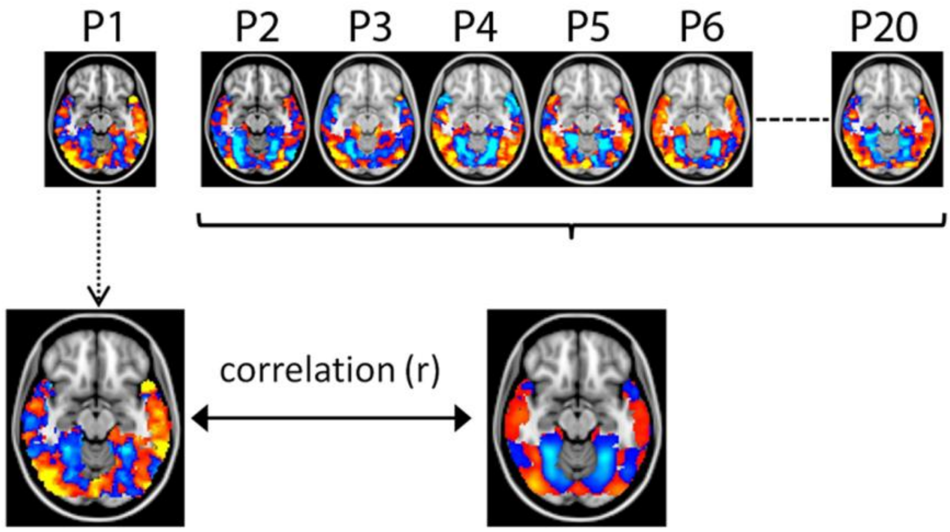

Figure 2.2: figure displaying LOPO MVPA method. This example represents the cross participant validation technique used in Chapter 5. Unlike the within participant odd vs. even run cross validation used by Haxby et al (2001), the LOPO method averages the neural responses in each experimental condition across all but one participant. These neural response patterns are then correlated between the group and the removed individual and the process is repeated for each experimental participant.

The LOPO MVPA approach used in Chapter 5 offers at least two benefits when compared to the traditional even vs. odd Haxby style cross validation used in Chapter 4. Firstly, it provides a more robust neural response comparison by averaging neural response across participants in each iteration; this increases the signal to noise ratio regarding the neural response to the stimulus category as the average comes from 23 samples each time (participant n - 1) rather than just 4 runs as in chapter 4. This allowed increased power in a relatively short functional run with few repetitions of each block (n=6). Secondly, as LOPO MVPA is carried out in a standardised brain space it provides a test of how reliable different functional neural response patterns are across different participants. Therefore it is provides a stricter test on whether neural response patterns to the same facial identities are consistent across different participants. Thus the LOPO MVPA also tested whether representations of each facial identity in this experiment were similarly distributed in face responsive brain regions across participants.

2.1.1 Limitations of MVPA

Although MVPA provides a paradigm that is much more sensitive than traditional univariate designs, there are limitations and situations in which it may not be able to successfully discriminate between neural response patterns. Firstly, although the spatial scale for sampling is improved in MVPA paradigms when compared to univariate GLM based methods is still nowhere near to the level of single cell recordings Logothetis (2008). Typically each fMRI voxel is a few millimetres in each spatial dimension and will contain and sample from millions of neurons. Therefore when the MVPA paradigm does not discriminate between patterns for a stimulus it cannot be assumed that the neural population samples do not contain information useful for encoding that stimulus. It could merely mean that the neural response profile occurs at a spatial resolution too fine to be sampled at the scale typically captured in each voxel (Haynes, 2015).

Another reason MVPA may not be able to discriminate between neural response patterns is because it possible that if neural tuning patterns were distributed in a random way within a small cluster of neurons then no coherent patterns would arise in the response profile at the voxel level (Haynes, 2015). Similarly if there are cases where a few neurons are involved heavily in encoding a stimulus but the surrounding neurons included in the MVPA contribute a significant amount of noise, it is possible that the informative voxels are in effect 'drowned-out' and the signal is lost (Etzel, Zacks, & Braver, 2013; Haynes, 2015).

Problems can also arise when interpreting pattern discrimination across different cortical areas. This is because sensitivity of the BOLD response varies across brain regions (Logothetis & Wandell, 2004), which in turn affects the signal to noise ratio in the fMRI sampling across these regions. Hypothetically it is possible that in a scenario two brain regions could be equally important in a task or function but one of these regions is less consistently sampled by fMRI for any of a number of reasons (e.g. signal drop out, artefacts affecting spatial resolution, noise from large nearby blood vessels). In this example the brain region with a better signal to noise ratio will appear to discriminate between stimulus categories or tasks more accurately. From this it may be incorrectly concluded that the region with higher discrimination accuracy is more important in encoding that specific stimulus/task when in reality this is an illusion. For this reason a presence or lack of discrimination between states or stimuli in a brain region could on occasion be a result of the signal to noise present in that regions, therefore the relative amount of signal between regions should be taken into account before drawing conclusions based on discrimination accuracy.

2.2 Regressing neural data to behavioural measures: linking behaviour to brain function

Representational Similarity Analysis (RSA; Kriegeskorte, Mur, & Bandettini, 2008; Kriegeskorte, 2009) provides a method for testing models based on theoretical predictions about the perceptual or neural representation of a stimulus or mental states. RSA can be used to compare information carried by a given representation in behavioural response patterns, neural activity patterns or a theory based representational model (seen in figure 2.3). RSA analysis usually begins by creating a dissimilarity matrix with each entry displaying the dissimilarity between the behavioural response and neural response patterns associated to two stimuli or states (Kriegeskorte, 2009). In Chapter 4, however, representational similarity matrices (RSM) rather than dissimilarity matrices were computed; this was done by reverse scaling neural and perceptual responses so that they display similarity rather than dissimilarity. This reverse scaling doesn't fundamentally change the analysis, rather it makes the output more intuitive as a good model fit corresponds clearly to a positive correlation and high similarity whilst a negative correlation corresponds to poor model fit and low similarity.

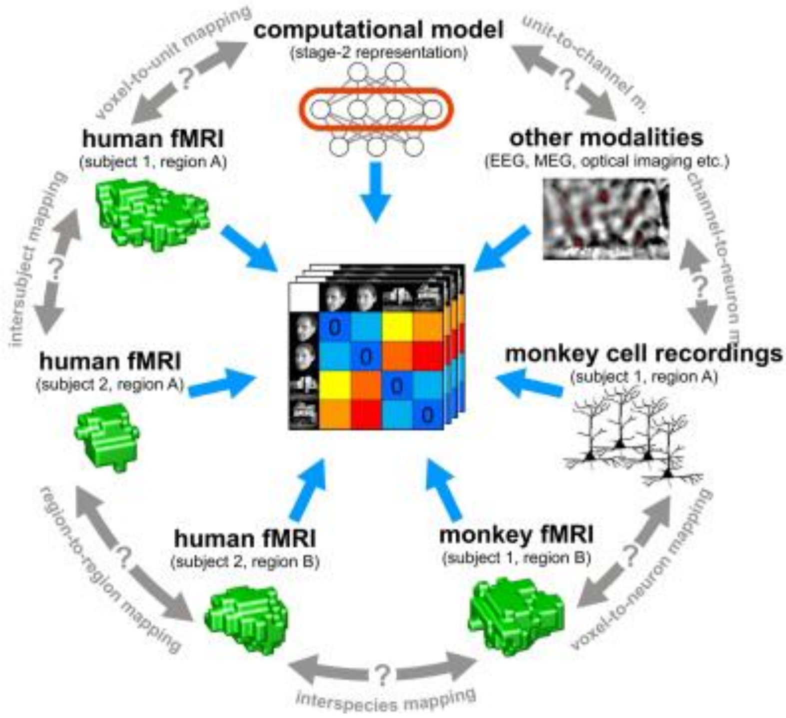

Figure 2.3: image taken from Kriegeskorte et al (2008). A visual representation of Representational Similarity Analysis and its ability to compare multidimensional models from a number of domains. An RSA analysis relies on the outcome variable having more than one observation per model dimension and normally distributed variance in the data.

Representational similarity analysis is a useful tool for probing the population coding of a given set of neurons by comparing activation patterns of neurons to perceptual or stimulus similarity models (Kriegeskorte & Kievit, 2013). With regards to testing models, for neural data it is recommended that that the RSMs for neural response are averaged across a participant group rather than using noisy single participant responses (Nili et al., 2014). Grouping data this way means that the RSM model that is used as a predictor has already been significance tested to determine whether the within category elements of the matrix (which are stimuli from the same category) can be successfully discriminated from between category elements (this cross validation procedure was described earlier in section 2.1 page 4) and whether the RSM therefore represents neural responses that can discriminate between different stimulus categories.

At this point the RSMs for neural and behavioural data can be tested for similarity through correlation or regression analysis. Firstly, the predictor variable RSM (e.g. neural response patterns to facial expressions in a brain region) is averaged across all participants and used to predict the outcome variable (e.g. participants' perceptual similarity responses to facial expressions) which contains RSMs for each participant. At this point because the predictor variable RSM (neural response patterns to expressions in a brain region) has the same number of dimensions as the outcome variable (participants' perceptual similarity responses to facial expressions), each point in the predictor RSM and the outcome RSM can be correlated and significance tested. An example of this method can be seen Chapter 4 where participants' perceptual similarity rating patterns to facial expressions are compared to shape and surface based visual models of facial expressions.

Significance testing must be carried out between RSMs to determine whether the predictor model can explain a significant amount of variance in the outcome variable. This can be done a number of ways but is carried out in Chapter 4 using simple linear regression and is similar to the regression analysis method used by Vass & Epstein (2013). Firstly the data in both predictor and outcome RSMs are normalised to produce the standard deviation of the mean. The competing image similarity models are also tested to check that they do not exceed common co-linearity thresholds, in this case a Variance Inflation Factor above 5 (Montgomery, Peck, & Vining, 2012). When this is the case, it can be assumed that each model can potentially explain a reasonable amount of unique variance in the outcome variable. Next, each image model is entered alone to predict the perceptual similarity responses for expression similarity. This regression coefficient is then significance tested against zero to test whether the predictor model explains a significant amount of variance in the outcome variable.

If the regression analysis indicates that the predictor model can significantly explain the outcome variance this allows inferences to be made about the representations held by the underlying neural population. This method is an advance on univariate measurements of neural populations. As discussed earlier, univariate paradigms can be used to estimate whether neurons are preferentially involved in a task or mental state over another. This gives precise spatial information about which brain voxels contain neurons that are differentially involved in encoding these stimuli or tasks. Multivariate methods on the other hand focus more on the whole patterns of activation rather than single response units and provide a measure of how similar the neural populations that encode different stimuli or mental states are.

This representational similarity analysis method has been successfully used recently to test how different representation of facial identity cluster in neurons in the macaque face patch (Freiwald & Tsao, 2010) and which neural populations are sensitive to gaze direction in human STS (J. Zhang, Kriegeskorte, Carlin, & Rowe, 2013) . As there have now been a number of successful attempts at decoding facial expressions from neural responses in face selective regions using MVPA paradigms, in Chapter 4 RSA is used in conjunction with MVPA to test the underlying information contained in these neural responses.

Chapter 3 - Contributions of feature shapes and surface cues to the recognition of facial expressions

This Chapter is adapted from: Sormaz, M., Young, A. W., & Andrews, T. J. (under review). Contributions of feature shapes and surface cues to the recognition of facial expressions.[1]

3.1 Abstract

Theoretical accounts of face processing often emphasise feature shapes as the primary visual cue to the recognition of facial expressions. However, changes in facial expression also affect the surface properties of the face. In this study, we investigated whether this surface information can also be used in the recognition of facial expression. First, participants identified facial expressions (fear, anger, disgust, sadness, happiness) from images that were manipulated such that they varied mainly in shape or mainly in surface properties. It was found that the categorization of facial expression is possible in either type of image, but that different expressions are relatively dependent on surface or shape properties. Next, we investigated the relative contributions of shape and surface information to the categorization of facial expressions. This employed a complementary method that involved combining the surface properties of one expression with the shape properties from a different expression. Our results showed that the categorization of facial expressions in these hybrid images was equally dependent on the surface and shape properties of the imchapter4

age. Together, these findings provide a direct demonstration that both feature shape and surface information make significant contributions to the recognition of facial expressions.

[1] The author, Mladen Sormaz, designed the experiment, analysed the results and wrote the article under the supervision of Prof. Andrew Young and Prof. Timothy Andrews.

3.2 Introduction

The human face has a complex musculature that allows it to create a remarkable variety of facial expressions (Du, Tao, & Martinez, 2014). Although there are individual differences between people in the precise anatomical arrangement of the facial muscles, those muscles involved in producing facial expressions of what are considered to be basic emotions (which include happiness, sadness fear, anger, and disgust) are highly consistent across individuals (Waller, Cray, & Burrows, 2008). These muscles allow a person to move critical expressive features such as the eyebrows, eyes, nose and mouth in ways that can change their shapes (e.g. raising or lowering the corners of the lips, widening or narrowing the eyes), their positions (raising or lowering the eyebrows), or often both (wrinkling the nose, or lowering the jaw to open the mouth).

Despite this well-known anatomical background, the nature of the visual information underlying recognition of facial expressions is poorly understood. While an obvious place to begin looking for critical visual cues might seem to be in the patterns of movement themselves, these are difficult to define and the good recognition of photographs of normal intensity basic emotions shows that the apex of a set of muscle contractions often creates an easily recognisable expressive configuration of the facial features. Moreover, notational systems such as the Facial Action Coding System (FACS: Ekman & Friesen, 1978) depend on the fact that the underlying pattern of muscle contractions that create an expression is evident even in a static image. Many studies therefore begin by exploiting the recognisability of well-validated photographs of facial expressions such as the Ekman and Friesen (1976) series, as was done here.

There are many ways of thinking about the visual information conveyed by a photograph of a face, but one that has proved very useful is in terms of its shape and surface properties. Any facial image consists of a set of edges created by abrupt changes in reflectance due to the shapes and positions of facial features and a broader pattern of reflectance based on the surface properties of the face – also known as texture or albedo (Bruce & Young, 1998, 2012). Shape properties can be operationally defined by the spatial locations of fiducial points that correspond to facial features; note that in this sense 'shape' properties will include both the feature shapes and their positions. In contrast, surface properties result from the pattern of reflectance of light due to the combination of ambient illumination, the face's pigmentation, and shape from shading cues.

The distinction of shape from surface properties is widely used in face perception research (Bruce & Young, 1998, 2012) and is implicit in standard approaches to computer image manipulation (B. P. Tiddeman, Burt, & Perrett, 2001). These image manipulation techniques allow quasi-independent changes to a face's shape or surface properties. Such changes cannot be fully independent, of course, because many of the shape and surface properties of images will necessarily covary. For example, the surface property of shading is clearly affected in part by the face's shape. None the less, such methods allow us to hold face shape fixed as closely as possible (by using the same fiducial positions for a set of images) or to hold the surface properties fixed as closely as possible (by using the same surface brightness patterns in a set of images). This then allows a direct test of the relative contributions of shape and surface information. Studies based on this approach have demonstrated independent contributions of shape and surface properties to the perception of a range of facial characteristics including gender, age, attractiveness and dominance (Burt & Perrett, 1995; Russell, 2003; Torrance, Wincenciak, Hahn, DeBruine, & Jones, 2014).

Thinking of facial images as broadly consisting of shape (feature positions) and surface (pigmentation, shading patterns) properties has also helped our understanding of facial identity recognition, where it is clear that both shape and surface cues can contribute (Russell, Sinha, Biederman, & Nederhouser, 2006; Troje & Bülthoff, 1996), but that the role of surface cues becomes more salient for familiar faces (Burton, Jenkins, Hancock & White, 2005 ; Russell & Sinha, 2007).

In contrast to the established role of surface cues in the perception of facial identity, judgements of expression are often thought to be based primarily on the shapes and positions of critical expressive features such as the eyebrows, eyes, nose and mouth. This makes sense because these shape changes are a direct consequence of facial muscle movements. Evidence for the primary importance of shape cues in facial expression recognition comes from contrast reversal (as in a photo negative). In a contrast-reversed image the edges that define feature shape properties remain in the same positions, despite the huge change in overall surface properties. Although contrast negation is well-known to be very disruptive of facial identity recognition (Bruce & Young, 1998, 2012), it turns out that judgements based on facial expression are still possible in contrast-reversed images (Bruce & Young, 1998; Magnussen, Sunde, & Dyrnes, 1994; White, 2001; Pallett & Meng, 2013; Harris, Young, & Andrews, 2014a). Similarly image manipulations that completely remove surface information, such as line drawings of faces, also show relatively preserved expression perception (McKelvie, 1973; Etcoff & Magee, 1992). Using such evidence, most current accounts posit shape information to be the most important cue in the perception and recognition of expression (Calder, Young, Perrett, Etcoff, & Rowland, 1996; Bruce & Young, 2012).

Although previous studies have suggested that feature shape is the dominant cue for the perception and recognition of facial expressions, there are grounds for thinking that surface information might also play a role (Calder, Burton, Miller, Young, & Akamatsu, 2001; Benton, 2009). For example, Benton (2009) found a decrease in the emotional expression aftereffect to facial expressions when images were negated, suggesting that the perception of facial expression can be affected by changes in surface information. Using Principal Component Analysis (PCA), Calder et al (2001) found that principal components (PCs) that convey variation in surface information could be used to categorize different facial expressions, albeit to a lesser extent than PCs that convey variation in shape. However, while these findings show a potential role for surface cues, they do not provide a direct test of whether surface properties are actually used for the recognition of facial expression. None the less, there are obvious ways in which surface properties might be useful to facial expression recognition. For example the feature shape change of opening the mouth will be accompanied by a bright region if the teeth are bared or a relatively dark region if the teeth are retracted; these different surface brightnesses are a direct reflection of muscle movements that clearly convey different expressions. Moreover, there are also indirect effects of underlying muscle movements such as the skin folding around the mouth and eyes resulting from smiling. These changes do not correspond to specific facial features, and are largely evident from their impact on surface shading patterns.

The aim of the current study was therefore to investigate the contribution of changes in the shapes of key expressive features (such as the eyebrows, eyes, nose and mouth) and changes in surface brightness patterns (such as those resulting from showing the teeth, or furrowing the brow) to the categorization of facial expression. In Experiment 1, we manipulated images to create facial expressions that varied primarily in shape or primarily in surface cues. This was achieved by reshaping images of different expressions to standardise the locations of the fiducial positions across the images, or by standardising the surface properties as far as possible by overlaying the same averaged surface onto the fiducials that characterise each expression. Because many of the shape and surface properties of images will necessarily covary, this method does not orthogonally manipulate shape and surface information, but it does allow us to hold the shape fixed as closely as possible (by using the same fiducial positions for all images) or to hold the surface properties fixed as closely as possible (by using the same surface brightness patterns in all images). This then allows a direct test of whether the information that remains free to vary across images can actually be used for the categorization of facial expression. In Experiment 2, we used contrast-reversed versions of the images used in Experiment 1 to further probe the role of shape and surface properties in the recognition of facial expressions. In Experiment 3, we then created hybrid images that combined the surface properties from one expression with the shape of a different expression. This approach offers a complementary method to that used in Experiment 1 and 2 for determining the relative contribution of surface and shape cues to the categorization of facial expressions.

3.3 Experiment 1

3.3.1 Methods

3.3.1.1 Participants

Participants (n = 20, female = 10, mean age = 24.8 years, SD = 3.8) were drawn from an opportunity sample of students and staff at the University of York. Participants gave informed consent and were paid or given course credit for their participation. All data were collected in accordance with the ethical guidelines determined by the Psychology Department of the University of York and were in accordance with the Code of Ethics of the World Medical Association (Declaration of Helsinki).

3.3.1.2 Stimuli

Figure 3.1 shows the stimuli for the three conditions used in Experiment 1 (original, shape varying, and surface varying; these are the 5x5 image matrices that form the leftmost columns in Figure 3.1). Static images of expressions were presented as these are well-recognised as long as they represent the apex of the pattern of muscle movements involved in producing the expression (see Bruce & Young, 2012). Five models (females F5, F6, F8, males M1, M6) were selected from the FEEST set (Young et al., 2002) of Ekman & Friesen (1976) photographs on the basis of high recognisability of their facial expressions and the similarity of the action units (muscle groups) used to pose each of the expressions from Ekman's coding of the action units given in the FEEST test manual (Andrew W. Young et al., 2002). For each model, images of expressions of fear, anger, disgust, sadness and happiness were used. These unmodified images from the Ekman & Friesen (1976) series formed the ‘original’ image condition (see Figure 3.1). Although also present in FEEST, expressions of surprise were omitted because the status of surprise as a basic emotion has been questioned (Oatley & Johnson-Laird, 1987); one can be pleasantly or unpleasantly surprised (Du, Tao, & Martinez, 2014).

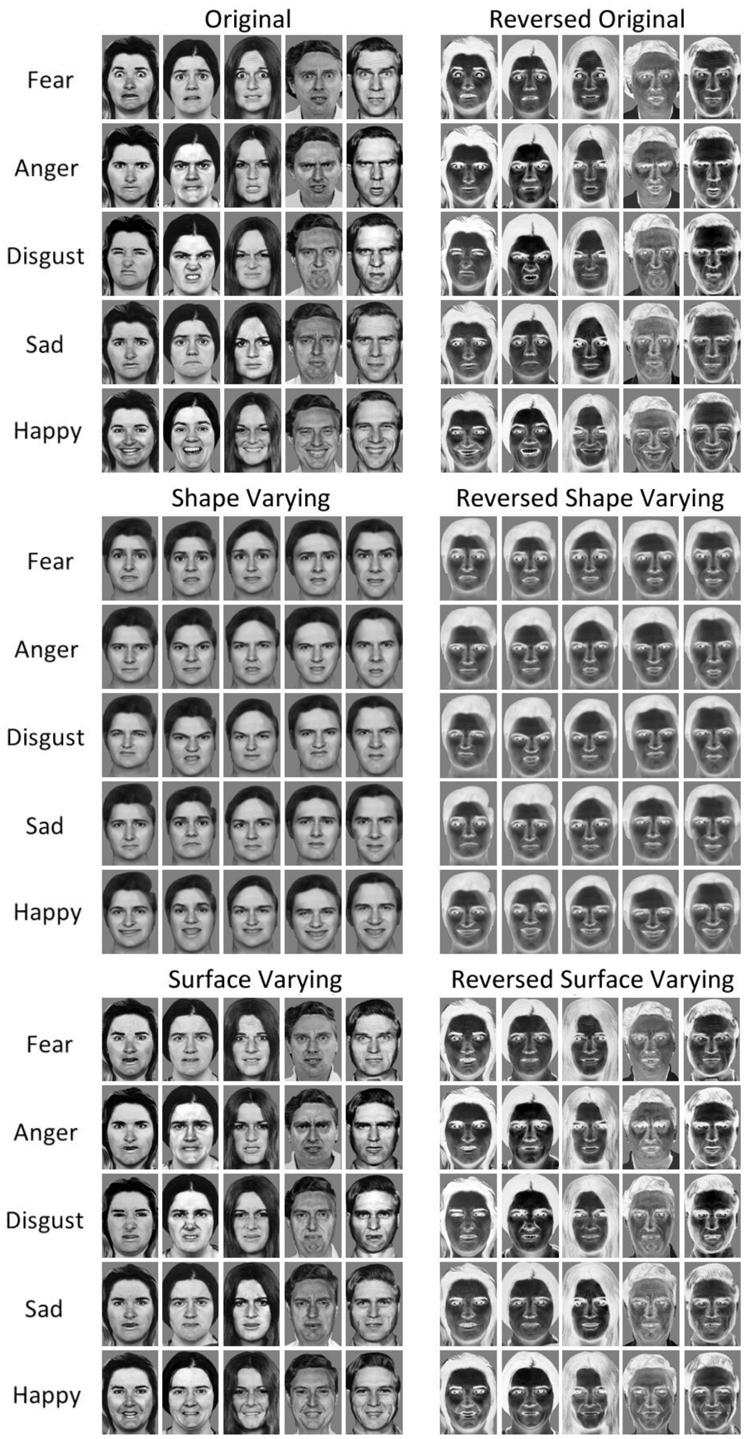

Figure 3.1 Images used in Experiment 1 and Experiment 2. Original images are taken from the FEEST image set (Young et al., 2002). These show 5 models posing 5 expressions. The shape-varying images were created by superimposing the average surface of all images onto the original shapes. The surface-varying images were created by superimposing the original surfaces onto the average shape of all images. Experiment 1 used the normal contrast versions of each of the images shown on the left of the display. Experiment 2 used the normal contrast versions of the original images

and the contrast-reversed versions of all images shown on the right of the display.

The aim of Experiment 1 was to measure the perception of images that vary primarily in shape or primarily in surface properties. To create the 'shape-varying' images, Psychomorph software was used to manually delineate 179 facial fiducial points on each of the 25 original images (Tiddeman et al., 2001). All 25 original images were then reshaped to the average shape (as defined by the fiducial positions) of all 25 images and averaged together to arrive at the average surface brightness across all 25 images. This averaged surface brightness was then reshaped back to the original shape (as defined by the original fiducial positions) of each of the 25 images. In this way, 25 new images were created, each of which had lost any expression-specific surface brightness information and only contained the unique shape cues associated with each individual expression.

A 'surface-varying' set of images was then created, in which primarily the surface information rather than shape information provided the cue to the posed expression. To do this the surface of each of the 25 original images was reshaped into the average shape across all 25 images. This removes most of the underlying shape cues to expression (as all images now shared exactly the same set of fiducial points), but leaves the surface information relatively unchanged (each image retains its surface brightness pattern).

3.3.1.3 Stimuli

Participants performed a 5-AFC (alternative forced choice) expression categorization task in which they indicated the perceived expression (Fear, Anger, Disgust, Sadness or Happiness) by a button press. Images were presented using PsychoPy (Peirce, 2008) at a viewing distance of 40 cm. Each trial began with a 500 ms fixation cross followed by a 1000 ms central presentation of one image (16^{o} x 11^{o}). Faces from the normal contrast original, shape-varying and surface-varying sets were presented in a randomised order and participants viewed each of the faces twice during the experiment, so each participant completed 150 trials. We recorded both accuracy and reaction time data for all participants. The experiment began with 15 practice trials using one variant of each expression from each experimental condition in a randomised order. The actors used in the practice trials were different from the ones in the main experimental trials.

3.3.2 Results

Our principal analyses used overall accuracies and reaction times to determine whether it was possible to identify the facial expression in images that primarily varied in surface or shape. Figure 3.2 shows that accuracy in the categorization task was above chance (20%) for the original (88.1 $\pm$ 2.2%), surface-varying (68.7 $\pm$ 3.5%) and shape-varying (70.2 $\pm$ 4.1%) images. An ANOVA showed a significant effect of condition ($F(1.42, 27) = 86.41$, $p < .001$, partial $n^2 = .82$). The main effect of condition was due to higher accuracy for the original images compared to both the surface-varying ($t(19) = 14.06$, $p < .001$) and shape-varying ($t(19) = 13.1$, $p < .001$) images. There was no significant difference between surface-varying and shape-varying conditions ($t(19) = 1.77$, $p = .1$).

We also found a significant effect of condition on reaction time ($F(2, 38) = 40.93$, $p < .001$, partial $n^2 = .68$). This main effect was due to faster reaction times to original images compared to either surface-varying ($t(19) = 7.3$, $p < .001$) or shape-varying ($t(19) = 7.23$, $p < .001$) images. There was no difference ($t(19) = .18$, $p = .86$) in reaction time between surface-varying (1557 $\pm$ 73 ms) or shape-varying (1553 $\pm$ 75 ms) images.

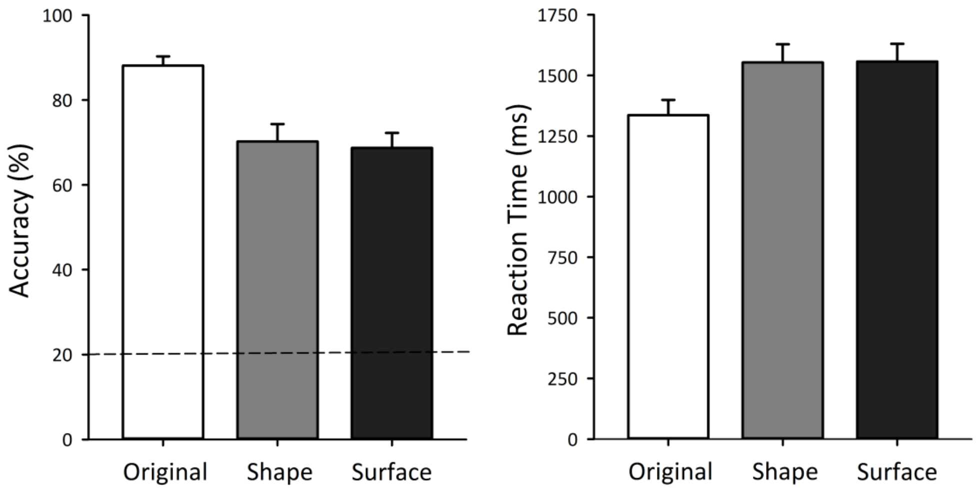

Figure 3.2 Experiment 1: Accuracy and reaction times for the categorization task averaged across all expressions. Accuracy was above chance levels (horizontal line) but lower than for the original images for both the shape-varying (shape) and surface-varying (surface) conditions. There was no significant difference in the accuracy or reaction time between the shape-varying and surface-varying conditions. Error bars represent standard error of the mean.

A subsidiary analysis was used to do determine whether different image cues are differentially important for different expressions, by including the five expressions as a factor in the accuracy analysis (Figure 3.3). A 2 way ANOVA showed a significant interaction between image type and expression ($F(8,152) = 6.12$, $p<.001$, partial $n^2 = .45$). All post hoc paired t tests carried out to compare different conditions and all stated p values were Bonferroni-Holm adjusted and tested at a critical alpha level of .05. Accuracy for anger ($t(19)= 3.38$, $p= .014$ was higher for surface-varying images compared to shape-varying images. In contrast, accuracy for shape-varying images compared to surface-varying images was higher for sad ($t(19) = 5.01$, $p<.001$) and happy ($t(19) = 10.22$, $p<.001$) expressions. Accuracy for the original images was higher compared to surface-varying images for fear ($t(19) = 7.21$, $p<.001$), anger ($t(19) = 3.5$, $p=.014$), sadness ($t(19) = 6.98$, $p<.001$) and happiness ($t(19) = 12.42$, $p<.001$). Accuracy for the original images was higher compared to shape-varying images for fear ($t(19) = 6.58$, $p<.001$), anger ($t(19) = 6.97$, $p<.001$) and disgust ($t(19) = 5.08$, $p= .001$). Accuracy for original images was not significantly greater to shape-varying images for happy (t (19) = 1, p= .79) and sad (t (19) =1.16, p= .79) expressions.

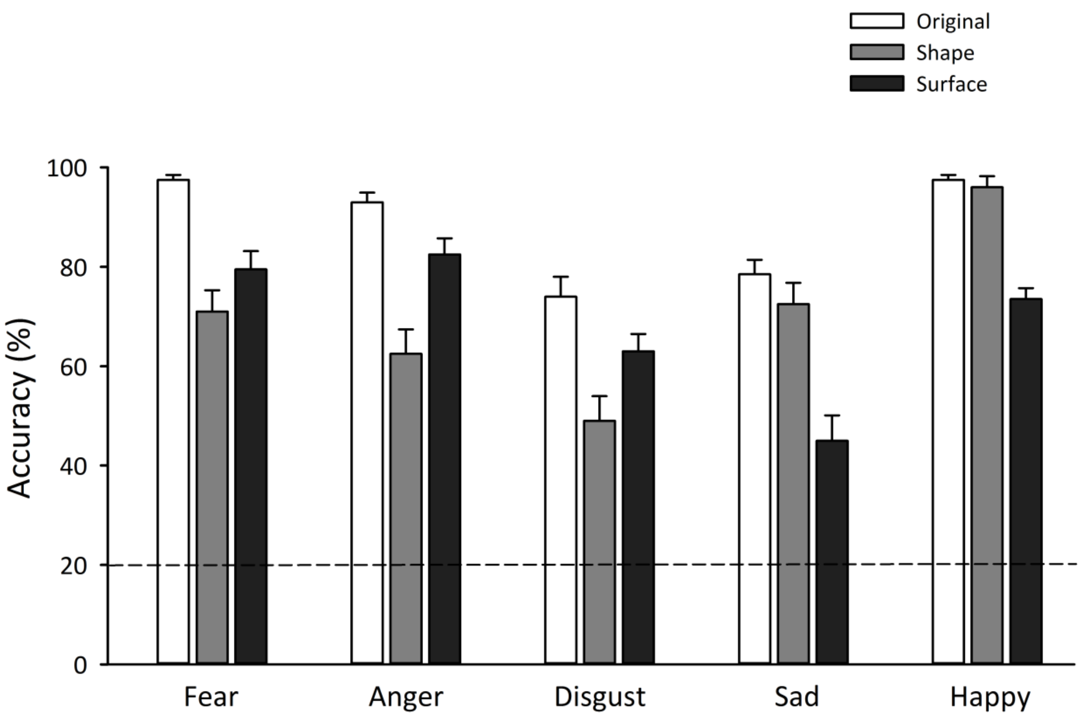

Figure 3.3 Experiment 1: Accuracy for the categorization task for each expression. Error bars represent standard error of the mean. Accuracy for shape-varying (shape) and surface-varying (surface) conditions was above chance level (horizontal line) for each expression. However, the relative importance of shape and surface properties differed across expressions, with shape cues being relatively important for happy and sad expressions, and surface cues to anger and disgust.

3.3.3 Discussion

Facial expression images that were manipulated to have fixed surface properties (the shape-varying images) or to have fixed shapes (the surface-varying images) were less well-recognised than original images from the Ekman & Friesen (1976) series, with longer reaction times and higher error rates. This shows that both shape and surface properties contribute to facial expression recognition.

Although there was no difference in the overall impact of holding fixed the shape or surface properties of the expressions, there were clear differences in how these contributed to recognising different emotions. For happiness and sadness, images with averaged surfaces (the shape-varying images) were recognised better than images with averaged shapes (the surface-varying images), showing the importance of shape cues for recognising these expressions. The opposite pattern held for anger and disgust, for which surface cues made a stronger contribution.

A potential limitation of our method for controlling shape is that only 179 fiducial positions were used to define the shape of each expressive feature, and human perceivers are known to be very sensitive to small differences in curvature that underlie the perception of feature shapes (Kosslyn, Hamilton, & Bernstein, 1995). We therefore ran an experiment using contrast-reversed images to confirm that the surface-varying images (i.e. those with the fixed fiducial positions) did not contain significant residual shape information. Contrast reversal has no effect on the positions of edges in the images, as the abrupt discontinuities in brightness values that create the perception of edges are still present, so it does not markedly affect what we here call shape information. Instead, contrast reversal has a substantial effect on surface properties because it inverts the relationships between all the relatively light and dark areas in the image. We therefore predicted that contrast reversal should result in a decrease in the recognition of expressions from the surface-varying images (as their fixed fiducial positions should largely have eliminated differences in the shape information that can survive contrast reversal), but should have no effect on shape-varying images (as these images do not contain any useful surface information that could be disrupted by contrast reversal).

3.4 Experiment 2

3.4.1 Method

3.4.1.1 Participants

Participants (n = 20, female = 10, mean age = 20.3 years, SD = 1.8) were recruited in the same way as for Experiment 1. Participants gave informed consent and data were again collected in accordance with the ethical guidelines determined by the Psychology Department of the University of York.

3.4.1.2 Stimuli

Stimuli are shown in Figure 3.1. They comprised contrast-reversed versions of the images used in Experiment 1 (the 5x5 matrices of images shown in the rightmost columns of Figure 3.1). Normal contrast versions of the original set of 25 images from the FEEST set (forming the 5x5 image matrix positioned in the upper left part of Figure 3.1) were also included as a point of comparison.

3.4.1.3 Procedure

The procedure followed that established for Experiment 1, with the exception that four sets of stimuli were use in Experiment 2 (normal contrast original images, contrast-reversed original images, contrast-reversed shape varying images, and contrast-reversed surface varying images).

3.4.2 Results

Mean accuracies and correct reaction times for recognition of emotion in each condition are shown in Figure 3.4.

Our principal analyses again compared overall accuracies and reaction times across conditions. Figure 3.4 shows that accuracy in the categorization task was above chance (20%) for original images (81.1 $\pm$ 3.4%), reversed original (70.9 +3.7%), reversed surface-varying (39 $\pm$ 4.9%) and reversed shape-varying (56.5 $\pm$ 5.3%) images. An ANOVA of the accuracy data with Greenhouse-Geisser corrected degrees of freedom (due to violation of sphericity) showed a significant effect of condition (F (2.17, 41.3) = 174.95, p<.001, partial n^2 =.90). Post hoc paired t tests p values were again Bonferroni-Holm adjusted. The main effect of condition was due to higher accuracy for the original images compared to the negated (t (19) = 8.2 , p<.001), the reversed surface-varying (t (19) = 22.42, p <.001), and the reversed shape-varying (t (19) = 10.07, p <.001) images. There was also a significantly higher recognition accuracy for the contrast-reversed original images than for contrast-reversed surface-varying (t (19) = 19.96, p<.001) and contrast-reversed shape-varying conditions (t (19) = 7.85, p <.001). Recognition accuracy was significantly higher in the reversed shape-varying condition than in the reversed surface-varying condition (t (19) = 7.38, p <.001)

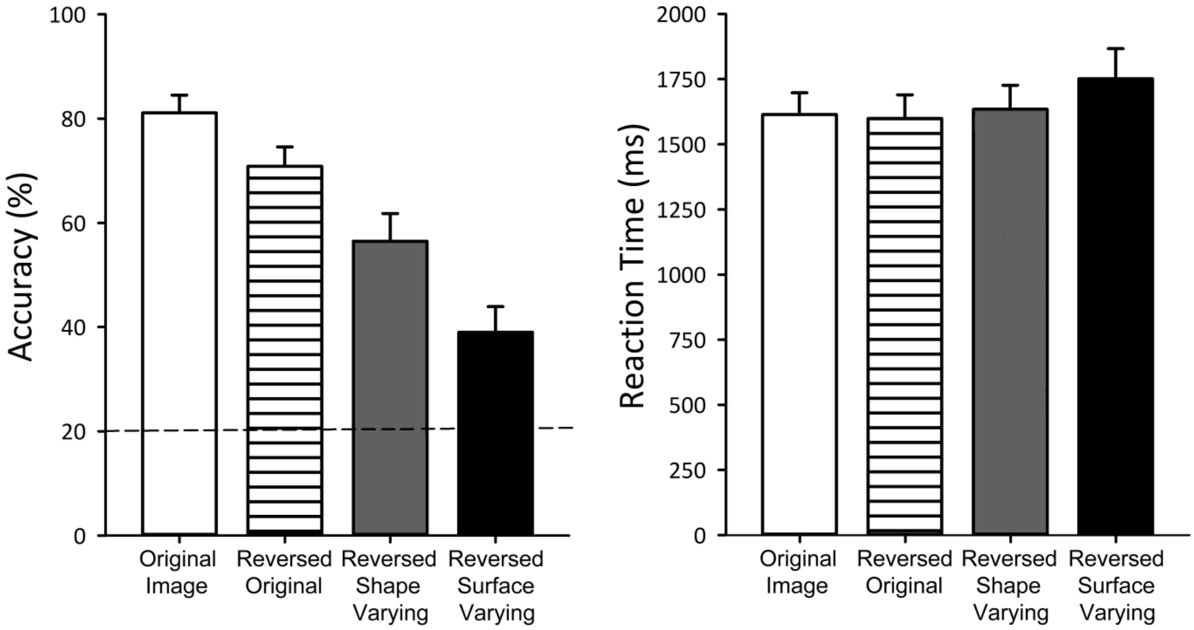

Figure 3.4 Experiment 2: Accuracy and reaction times for the categorization task averaged across all expressions. Accuracy was above chance levels (horizontal line) for all conditions, but the biggest drop in performance was evident for the contrast-reversed surface-varying images. The right panel shows reaction times in each of the experimental conditions for correct trials. Error bars represent standard error of the mean.

Figure 3.4 shows that the categorization task reaction times were similar in the original image (1614 $\pm$ 83 ms), contrast-reversed original image (1599 $\pm$ 91 ms), and contrast-reversed shape-varying (1635 $\pm$ 92 ms) conditions, and slowest in the contrast-reversed surface-varying (1751 $\pm$ 116 ms) condition. This pattern was confirmed with a second ANOVA (also with Greenhouse-Geisser corrected degrees of freedom) which demonstrated a significant effect of condition on reaction time (F (1.93, 36.6) = 6.03, p =.006, partial n^2 =.24). Post hoc paired t tests p values were again Bonferroni-Holm adjusted. This main effect reflected significantly faster reaction times to original images compared to contrast-reversed surface-varying (t (19) = 2.88, p =.04) images. There was no significant difference in reaction time between original images and contrast-reversed original images (t (19) =.22, p = 1) or contrast-reversed shape-varying (t (19) = .23, p = 1) conditions. There were significantly faster reaction times in the contrast-reversed original condition compared to contrast-reversed surface-varying (t (19) = 2.76, p = .04) but not contrast-reversed shape-varying conditions (t (19) = .38, p = 1). The reaction times in the contrast-reversed shape-varying condition were significantly faster than in the contrast-reversed surface-varying condition (t (19) = 3.54, p = .01).

In Experiment 1, we found no difference in accuracy for shape-varying and surface-varying original-contrast images, whereas in Experiment 2 there was greater accuracy for shape-varying compared to surface-varying contrast-reversed images. To confirm that this difference in the pattern of results was statistically reliable, we took the accuracy data from the three conditions of Experiment 1 (original, shape-varying, and surface-varying images) and the three corresponding contrast-reversed conditions in Experiment 2 (contrast-reversed original, contrast-reversed shape-varying, and contrast-reversed surface-varying images) and submitted these to a 2x3 ANOVA with image Format (normal or contrast-reversed) as a between-groups factor and experimental Condition (normal, shape-varying, surface-varying) as a within-group factor. This ANOVA revealed a significant main effect of image format (F (1, 38) = 75.73, p<.001, partial n^2 = .67), a significant main effect of condition (F (1.48, 56.06) = 202.54, p<.001, partial n^2 = .84), and a significant interaction between image format and condition (F (1.48, 56.06) = 10.12, p = .001, partial n^2 = .21). The image format x condition interaction confirms that the pattern of effects differed between the normal-contrast image format used in Experiment 1 and the reversed-contrast images used in Experiment 2.

3.4.3 Discussion

Experiment 2 found that contrast reversal had a greater impact on recognising the surface-varying than the shape-varying images created for Experiment 1. Contrast reversal has little effect on the edge cues that define features shapes but completely changes the brightness values that define surface patterns. On this basis, if the images created for Experiment 1 did indeed minimise the roles of surface or shape cues as intended, we would expect to find a substantial decrement of contrast reversal on the recognition of surface-varying images and less impact for shape-varying images. That this was precisely the pattern observed provides support to the view that the image manipulation techniques used to create the stimuli for Experiment 1 had achieved the intended effects.

Further support for the importance of surface properties in the recognition of facial expression is shown by the effect of contrast reversal on the original and shape-varying images. We found that contrast reversal of original images significantly lowered the recognition of facial expression. The only difference between these image sets is the surface properties (normal contrast compared to reversed contrast). This implies that surface properties are important for the recognition of facial expression. We also found that contrast reversed original images were recognized significantly better than contrast reversed shape-varying images. Again, the only difference between these images is the surface properties (reversed contrast of original images compared to reversed contrast of averaged images). This implies that the surface properties continue to contribute to the recognition of facial expression even in the contrast reversed original images. Thus, contrast reversal disrupts rather than eliminates surface cues.

To further explore the contributions of shape and surface properties we introduced a complementary method for Experiment 3, in which we tested the categorisation of all possible combinations of the average shape of one expression with the averaged surface of another expression. This method pits shape against surface properties, allowing a test of which dominates the expression seen in the hybrid image. For example, we can ask whether a hybrid of fear shape and happy surface will be seen as fear (shape dominance), as happiness (surface dominance) or as some other expression (because of the inconsistent cues).

3.5 Experiment 3

3.5.1 Methods

3.5.1.1 Participants

The participants used in Experiment 3 were those used in Experiment 1. They gave informed consent and were tested within the ethical guidelines of the University of York Psychology Department.

3.5.1.2 Stimuli

The aim of Experiment 3 was to compare the relative contribution of shape and surface properties to the perception of facial expression with a technique that would be complementary to that used in Experiment 1. Hybrid images were created that combined the average surface properties from one expression with the average shape from another expression. To generate the average surface properties from each expression, all 25 original images were reshaped to the average across all 25 images. The five images of each expression were then averaged to create an average surface for each of the five expressions. To generate the shape properties for each expression, we averaged the position of the fiducial points across the five images from each expression. This gave rise to one shape image for each expression. Finally, we combined the averaged surface for each expression with the averaged shape for each expression to create a matrix of 25 images (Figure 3.5) in which images on the top left to bottom right diagonal have the average surface and shape properties from the same expression. All other images in Figure 3.5 have the average surface properties from one expression and the average shape properties from a different expression, allowing us to estimate whether these hybrid expressions are recognised primarily from of their shape or their surface properties.

3.5.1.3 Procedure

The same 5AFC categorization task and presentation parameters were used as for Experiment 1, except that in Experiment 3 each of the 25 faces was presented five times, leading to 125 trials.

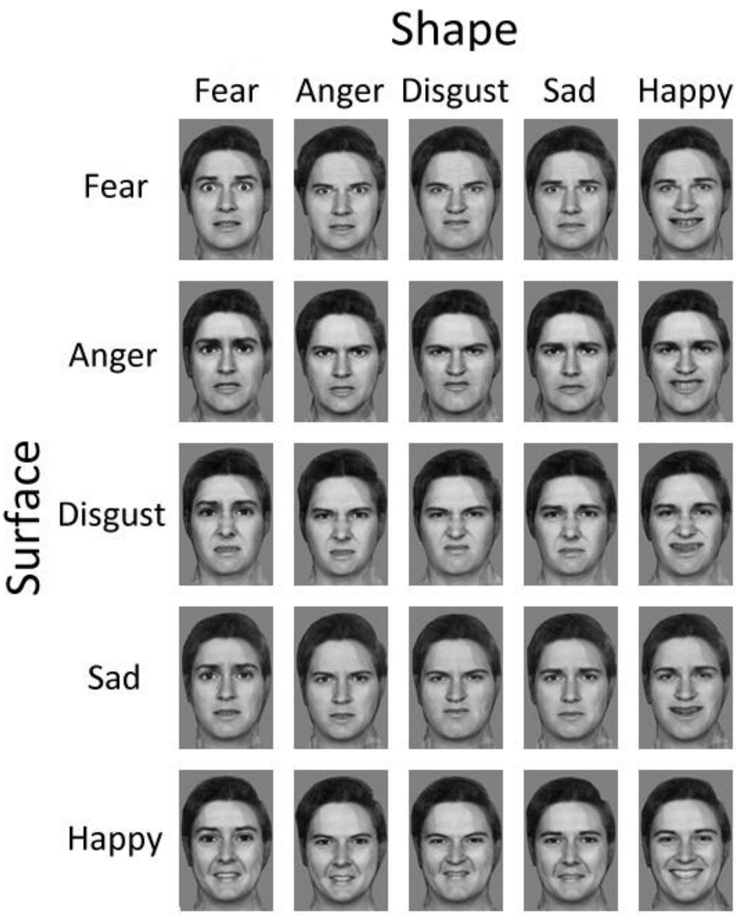

Figure 3.5 Hybrid images used in Experiment 3. Images were created by combining shape and surface properties from averaged facial expressions. Images on the top left to bottom right diagonal have the averaged shape and surface properties of the same facial expression. All other images have average shape and surface properties from different expressions. For example, the bottom left image combines an averaged fear shape with an averaged happy surface.

3.5.2 Results

In Experiment 3, we directly probed the relative contributions of shape and surface properties to the categorization of facial expression using hybrid images with the average surface properties from one expression and the average shape properties from another expression. Because these images were created from the averaged shapes and surfaces of each expression across the 5 models from the FEEST set, we began by checking that they were none the less seen as the intended emotion when the shape and surface of the same expression were combined (the images falling along the diagonal from top left to bottom right in Figure 3.5). Participants categorized these images in which the surface and shape cues conveyed the same expression with an accuracy of 92%, showing at least as good recognition as the original Ekman & Friesen (1976) images from Experiment 1.

Next, we determined the response for hybrid images in which the surface and shape cues conveyed different expressions. Figure 3.6 shows that participants reported the expression that corresponded to the surface properties of the image on 41.1 % of trials and the expression that corresponded to the shape of the images on 46.5 % of trials. Only a small proportion of responses were based on neither shape nor surface. An ANOVA showed a significant effect of condition (F (1.53, 29.08) = 95.52, p <.001, partial η^2 = .83). Correct responses based on shape and surface properties were significantly higher compared to responses that did not correspond to the expression from either the surface or shape properties (both p<.001). There was no overall difference in the proportion of trials in which participants chose surface compared to shape (t (19) = 1.73, p = .1).

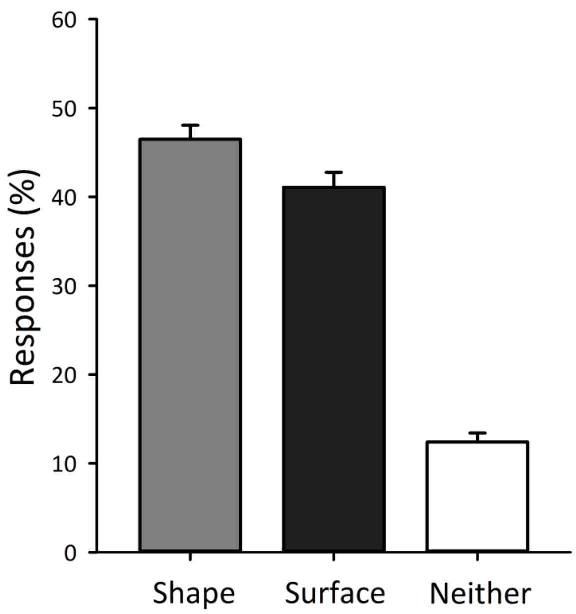

Figure 3.6 Experiment 3: Responses indicating whether the categorized expression corresponded to the Shape or Surface properties of the image, or when the response

did not correspond to the shape or surface information in the image (Neither). Responses based on Shape and Surface were significantly higher than the responses involving neither shape nor surface, but there was no significant difference between the responses based on Shape and Surface themselves.

To investigate the subsidiary issue of whether different image cues might be differentially important for different expressions, the data were separated by the expression identified in participants' responses, as shown in Figure 3.7. An ANOVA revealed a significant interaction between condition and expression (F (4.41, 83.86) =, p <.001, partial n^2 =.77). Post hoc paired t tests p values were again Bonferroni-Holm adjusted. In this analysis, the surface properties were more dominant than shape for disgust (t (19) = 3.11, p =.006) and sadness (t (19) = 4.95, p<.001). In contrast, shape cues were more dominant for happiness (t (19) =16.28, p<.001).

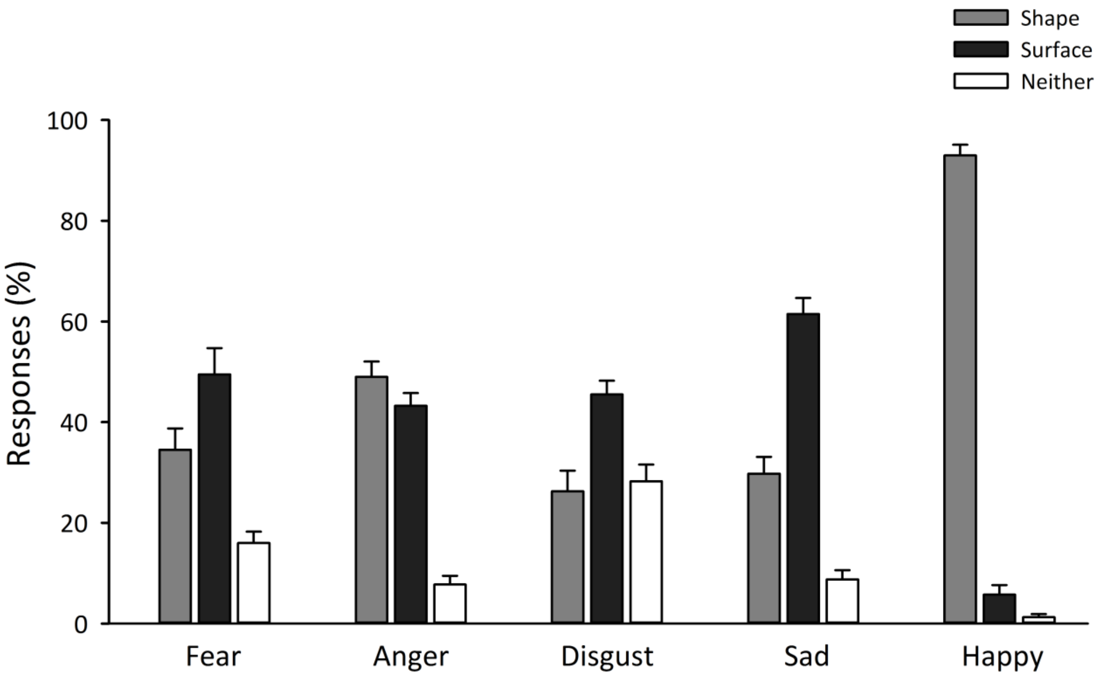

Figure 3.7 Experiment 3: Responses for each expression indicating whether the expression was categorized based on the Shape or Surface properties of the image, or when the response did not correspond to the shape or surface information in the image (Neither). Surface responses were significantly higher than Shape responses for disgust

and sadness. Shape responses were significantly higher than Surface responses for happiness.

3.5.3 Discussion

Like Experiment 1, the overall pattern of findings from Experiment 3 showed that both shape and surface properties are important to the recognition of facial expressions. As for Experiment 1, the analysis by expressions also showed that the relative importance of shape and surface cues varied across expressions. However, these detailed patterns differed somewhat across experiments, with the consistent findings being that perception of happiness is largely determined by feature shapes and disgust by surface properties. We can speculate that this reflects the salience of the distinctive mouth shape in happy expressions and the way that shading patterns enhance the nose wrinkling that characterises many expressions of disgust (Rozin, Lowery, & Ebert, 1994). The most important point, though, is again that both shape and surface properties contribute.

3.6 General Discussion

The aim of the present study was to investigate the roles of feature shapes and surface properties of the face in the recognition of facial expressions. To achieve this, we used complementary converging methods. In Experiment 1, we created images that held either shape or surface properties as constant as possible, to investigate the usefulness of each source of information in relative isolation. In Experiment 2, we validated the general method for varying shape and surface properties by testing recognition of contrast-reversed versions of the images used in Experiment 1. In Experiment 3, we created hybrid images that combined the averaged shapes and surfaces of different expressions, to investigate which type of information would dominate the perceived expression.

Our results show clearly that both shape and surface properties are useful for the recognition of facial expression. Indeed, despite the widely-shared opinion that the shapes of expressive features created by muscle movements are particularly important to expression, we instead found that both shape and surface information contributed more or less equally overall, albeit with some differences between different expressions. In Experiment 1, we found that identification accuracy for images that varied primarily in surface properties was well above chance (around 70% correct in 5AFC) and not significantly different from images that varied primarily in shape. In Experiment 2 we validated the properties of the images created for Experiment 1 by demonstrating a substantial decrement of contrast reversal on the recognition of expression for the surface-varying images and less impact for shape-varying images. This pattern is as expected because contrast reversal has little effect on the edge cues that define features shapes but completely changes the brightness values that define surface patterns. In Experiment 3, we directly compared the relative contributions of shape and surface cues to the categorization of facial expression. We found that participants were equally likely to use the surface or shape information to categorize facial expressions when viewing hybrid images that contain the surface properties from one facial expression and the shape properties from another facial expression.

So, while the present findings provide further support for the long held assertion that feature shape cues are important for the perception and recognition of expression, they also show that the importance of surface information in the representation of facial expressions has been underestimated. The novel finding from this study is that images that mainly contain variation in their surface properties can convey facial expression. Indeed, removing either shape or surface information impairs perception of expressions approximately equally. In both Experiments 1 and 3 categorisation of surface-only and shape-only images was significantly lower than for images containing appropriate shape *and* surface information. Taken together these findings show that the perceptual mechanisms that underpin the recognition of facial expression are tightly linked to both shape and surface information.

Impairment of shape or surface cues did not have an equivalent effect for each expression. This suggests that the informative cues in the face may vary for each expression, for example a smile signifying happiness produces a consistent shape change of an upturned mouth across all actors, making shape a salient and prominent cue. Conversely, a facial expression signifying disgust although also highly identifiable may produce less consistent or more subtle shape cues, therefore leaving the viewer to rely also on surface cues not present for other expressions. This reliance on both types of cue may reflect the natural covariance between shape and surface cues within many expressions. For example, fear expressions involve opening the mouth and widening the eyes (shape cues) and this creates salient contrast changes in the eye and mouth regions (surface cues). Moreover, some of the critical surface cues involve skin folding and other perturbations that are actually the result of movements of the facial features, allowing them to act as proxies for these even when the correlated feature shape changes themselves have been eliminated.

At first sight, these findings contrast with previous studies showing that the disruption of surface cues does not substantially impair perception of expressions (White, 2001; Pallett & Meng, 2013; Harris et al., 2014a). However, more careful, examination of such studies shows clear pointers to the conclusion we have reached. For example White (2001) showed in 2 of his 4 experiments a small but non-significant increase in error rate for contrast negated expressions when compared to normal images of facial expression. Similarly, although Pallett & Meng (2013) found preserved recognition accuracy for contrast-negated expressions they noted a reduced adaptation aftereffect for contrast-negated expressions. Indeed, as already noted Benton (2009) also found a reduced expression adaptation aftereffect to negated expressions, suggesting a role for surface information. However, our data go beyond these previous studies by showing that when the only differences between images are in the surface information, it can still be used to correctly identify some facial expressions. In other words, surface information alone can be sufficient.

Of course the photographs and computer-manipulated stimuli we used were two-dimensional static images, whereas faces we see in everyday life are three-dimensional and nearly always moving. However, the use of 2D static images brings advantages in terms of experimental control, and we do not think it has serious limitations. The role of movement in assisting facial expression recognition is most clearly evident with much more subtle and variable expressions than the type we have used here (e.g. Ambadar, Schooler, & Conn, 2005; Krumhuber, Kappas, & Manstead, 2013). The good recognition of photographs of normal intensity basic emotions shows, for these emotions at least, the apex of the set of muscle contractions forms a recognisable configuration and studies have not found much in the way of differences between neural responses to moving and static expressions of basic emotions (Johnston, Mayes, Hughes, & Young, 2013; Harris, Young, & Andrews, 2014b). So despite its importance, we need to be careful not to overstate the role of movement in facial expression recognition.

In a similar vein it should be noted that although our definition of surface information arguably contains both pixel intensity and 'shape from shading' information, we decided not to make this distinction. We concede that the two cues highly covary as a pixel intensity change is likely to occur in areas where facial expression change may induce the face appear to have wrinkles, lines or folds. For example in an image of a happy facial expression as well as shape from shading cues in the wrinkles around the eyes, there will be a great pixel intensity change around the mouth as the teeth become visible. We do not see this as a confound rather as a reflection of the diversity of facial expression information contained in the facial surface, for example were one to produce computer manipulated issues with no facial surface and only shape contours, there would be little variance in either pixel intensity or shape from shading across different facial expression images.

Much the same point applies to three-dimensionality. Although the face is a complex 3D structure, decades of research has established that humans are good at recognising age, sex, familiar identity and expression from 2D photographs (Bruce & Young, 2012). What may be important here, though, is that some of this excellent performance with photographs may reflect the presence of 3D cues in the form of shape from shading, but the standard image-manipulation methods we use here will assign shading information to surface properties and not to the feature shapes per se. This does not seem to us a serious limitation as there are no grounds at present for thinking that (say) a smile will be less obvious because a person has a protruding nose. Where it does have an influence, though, is that we have been careful to point out that the skin folds that result from raising the corners of the mouth, wrinkling the nose, or screwing up the eyes will also be treated as strongly covarying surface properties and not as changes in the feature shapes themselves.

It may be interesting for future studies to more accurately measure the exact degree of covariance between facial shape and surface change in images of facial expression and whether this is consistent across a large number of exemplars (i.e. only 5 actors and expressions were used in the current study). Future work may also benefit by separating the aforementioned underlying surface cues by generating face stimuli that preserve all 3D cues such as laser scanner expressive faces. Such stimuli could be designed for use in an expression labelling study where facial expression recognition accuracy is compared for faces with either pixel intensity difference or shape from shading as the unique varying expression cue. This could further elucidate whether pixel intensity or shape from shading cues are differentially important to recognition of different facial expressions.

In conclusion, we show that both shape and surface information can be used to identify facial expressions. We also show that the relative importance of shape and surface varies across different expressions, presumably reflecting the extent to which either cue can be diagnostic of a particular facial expression. In some ways, finding that both shape and surface properties play important roles fits what we noted in the Introduction concerning the perception and recognition of age, sex and identity. But because prominent theories (Haxby, Hoffman, & Gobbini, 2000; see also Calder & Young, 2005) draw a strong distinction between changeable (expression) and relatively invariant facial characteristics (age, sex, identity), it was important to investigate what happens in the case of facial expression. The key new finding is thus that both shape and surface information play important roles in the recognition of facial expressions.

Chapter 4 - Modelling the perceptual similarity of facial expressions from image statistics and neural responses

This Chapter is adapted from: Sormaz, M., Watson, D. M., Smith, W. A. P., Young, A. W., & Andrews, T. J. (2016). Modelling the perceptual similarity of facial expressions from image statistics and neural responses. *NeuroImage*, 129, 64–71.[2]

4.1 Abstract

The ability to perceive facial expressions of emotion is essential for effective social communication. The current study investigated how the perception of facial expression emerges from the image properties that convey this important social signal, and how neural responses in face-selective brain regions might track these properties. To do this, the current study measured the perceptual similarity between expressions of basic emotions, and investigated how this is reflected in image measures and in the neural response of different face-selective regions. The current study found that the perceptual similarity of different facial expressions (fear, anger, disgust, sadness, happiness) can be predicted by both surface and feature shape information in the image. Using block design fMRI, the current study found that the perceptual similarity of expressions could also be predicted from the patterns of neural response in the face-selective posterior superior temporal sulcus (STS), but not in the fusiform face area (FFA). These results show that the perception of facial expression is dependent on the shape and surface properties of the image and on the activity of specific face-selective regions.

4.2 Introduction

The ability to visually encode changes in facial musculature that reflect emotional state is essential for effective social communication (Ekman, 1972; Bruce & Young, 2012). A full

[2] The author, Mladen Sormaz, designed the experiment, analysed the results and wrote the article under the supervision of Prof. Andrew Young and Prof. Timothy Andrews. Dr. William Smith provided advice on the implementation of the procrustean analysis and David Watson provided advice for the implementation of the Multi Voxel Pattern Analysis.

understanding of the mechanisms that underpin the perception of facial expression requires understanding both the way in which these processes are driven by visual properties of the image and the way in which different brain regions are involved (Haxby, Hoffman, & Gobbini, 2000; Bruce & Young, 2012).

Any facial image consists of a set of edges created by abrupt changes in reflectance that define the shapes and positions of facial features and a broader pattern of reflectance based on the surface properties of the face, also known as the albedo or texture (Bruce & Young, 1998, 2012). Shape can be defined by the spatial location of fiducial points that correspond to key features of the face. In contrast, surface properties reflect the reflectance of light that is caused by pigmentation and shape from shading cues. Shape and surface properties have both been proposed to contribute to the perception of identity and expression (Bruce & Young, 1998; Calder, Young, Perrett, Etcoff, & Rowland, 1996), but with the perception of familiar identity being relatively dominated by surface cues (Burton et al., 2005; Russell & Sinha, 2007) and feature shapes being relatively dominant in perceiving facial expressions (McKelvie, 1973; Etcoff & Magee, 1992; Butler, Oruc, Fox, & Barton, 2008). This differential use of image properties in the perception of identity and expression is consistent with models of face perception which propose that they are processed independently (Bruce & Young, 1998, 2012; Haxby, Hoffman, & Gobbini, 2000).

Support for the critical role of shape information in the perception of facial expression is found in studies that show manipulations of the image that degrade surface information, but leave shape information intact, have little impact on perceptual and neural responses to facial expression (Bruce & Young, 1998; Magnussen, Sunde, & Dyrnes, 1994; White, 2001; Pallett & Meng, 2013; Harris, Young, & Andrews, 2014). Similarly, image manipulations that completely remove surface information, such as line drawings of faces, also show relatively preserved expression perception (McKelvie, 1973; Etcoff & Magee, 1992).

Although previous studies have suggested that feature shape is the dominant cue for the perception of facial expressions, there is some evidence to suggest that surface information may also play a role. Calder, Burton, Miller, Young & Amakatsu (2001) found that Principal Components (PCs) that convey variation in surface information could be used to categorize different facial expressions, albeit to a lesser extent than PCs that convey variation in shape. More recently, Benton (2009) found a decrease in the emotional expression aftereffect to facial expressions when images were negated, suggesting that the perception of facial expression can be affected by changes in surface information. So, it remains uncertain how different image properties contribute to the perception of facial expression.

The first aim of the current study was therefore to explore the relative importance of shape and surface properties to the perception of facial expression. Specifically, the current study asked whether the perceptual similarity of different facial expressions could be predicted by corresponding similarities in the shape or surface properties of the image. The perceptual similarity task involved rating the degree of similarity in expression between pairs of pictures of facial expressions. This task was used to generate a matrix of perceived (rated) similarities between exemplars of facial expressions of five basic emotions. This is equivalent to the procedure used to establish widely-adopted perceptual models such as Russell's circumplex (Russell, 1980), where expressions of emotion lie proximally or distally on a two-dimensional surface based on their perceived similarity, with the distance between expressions reflecting their similarity or confusability to human observers.

The second aim of this study was to determine if the perceptual similarity of facial expressions is reflected in the patterns of neural responses in face-selective regions of the brain. Neural models of face perception suggest that a network of face-selective brain regions underpins the perception of faces (Allison, Puce, & McCarthy, 2000; Haxby, Hoffman, & Gobbini, 2000; Ishai, 2008), with the posterior superior temporal sulcus (STS) playing a key role in processing facial expression (Winston, Henson, Fine-Goulden, &

Dolan, 2004; Engell & Haxby, 2007; Harris, Young, & Andrews, 2012; Baseler, Harris, Young, & Andrews, 2014; Psalta, Young, Thompson, & Andrews, 2014). Recent evidence has shown that it is possible to successfully decode some properties of facial expressions from face responsive brain regions (Wegrzyn et al., 2015, Said, Moore, Engell, & Haxby, 2010). Nevertheless, the extent to which the neural response can predict the fine-grained perception of facial expression remains unclear. Using multi-voxel pattern analysis (MVPA) techniques, the current study asked whether the perceptual similarity of expressions could be explained by the neural response in different face-selective regions. The current study predicted that patterns of response in regions associated with processing of facial expression should predict the perception of facial expression.

4.3 Methods

4.3.1 Participants

Twenty-four healthy volunteers took part in the fMRI experiment and the behavioural similarity ratings experiment (12 female, mean age = 25.2 years). All participants were right-handed and had normal or corrected to normal vision with no history of neurological illness. The fMRI work was approved and conducted following the guidelines of the York Neuroimaging Centre Research Ethics Committee, University of York, and the behavioural study by the Department of Psychology Ethics Committee. All participants gave written consent prior to their participation.

4.3.2 Stimuli

Figure 4.1 shows all the stimuli from the five expression conditions. Static images of expressions were presented as these are well-recognised as long as they represent the apex of the pattern of muscle movements involved in producing the expression (see Bruce & Young, 2012). By using well-validated images from the Radboud Face database (Langner et al., 2010) the current study ensured that this criterion was met. Images were selected on the basis of high recognisability of their facial expressions and the similarity of the action units (muscle groups) used to pose each of the expressions. Only male faces were used to avoid any confounds from characteristics introduced by gender differences in the images themselves. For each of five models, images of expressions of fear, anger, disgust, sadness and happiness were used.

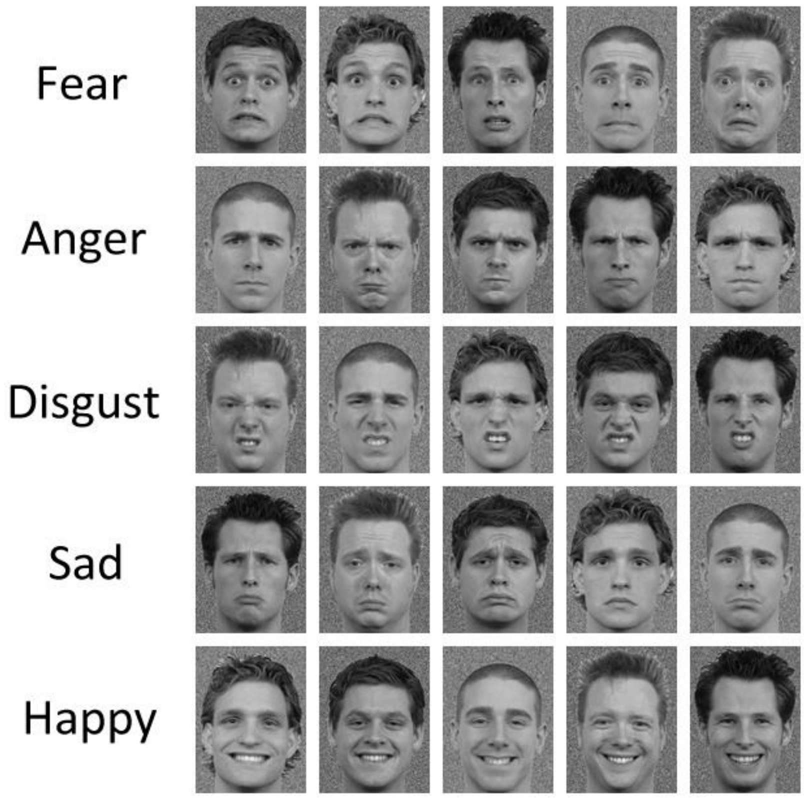

Figure 4.1 Images used in behavioural and fMRI experiments. Images were taken from 5 identities posing expressions of 5 basic emotions. Each row shows a typical sequence that might form a block in the fMRI experiment (presenting the images from left to right one at a time).

4.3.3 Perceptual Similarity Experiment

First, the current study determined the perceptual similarity of different facial expressions. Participants carried out a perceptual similarity rating task. Pairs of images were presented either side of a fixation cross and participants were asked to rate the images on the similarity of expression on a scale of 1-7 (1: not very similar expressions, 7: very similar expressions). Each possible combination of pairs of different images from the set of expressions was displayed once in the perceptual similarity rating experiment, excluding pairs of images from the same identity. This resulted in 200 trials in total. From these ratings the average rated similarity between examples of expressions of same or different basic emotions was derived. These similarity ratings were z-scored and then incorporated into a similarity matrix for each participant.

4.3.4 Image Properties

To determine whether the patterns of perceptual similarity found in the behavioural task could be explained by shape information in the face images, the locations of 140 fiducial points corresponding to expressive features in each of the face images were defined using PsychoMorph software (Tiddeman, Burt, & Perrett, 2001). This produced a 2 x 140 matrix for facial feature positions in 2D image space, with x and y co-ordinates for each fiducial point (Figure 4.2). These fiducial locations were then used to provide a measure of facial feature shape by entering the fiducial location matrices into a procrustean comparison (Schönemann, 1966) to measure the similarity in feature locations between every possible pair of images. The procrustean analysis rigidly aligns fiducial points allowing shape translation, rotation or scaling to correct for image position or size without morphing or non-linear image distortion. After alignment of a pair of images in this way, the procrustean metric computes the averaged squared distance between each pair of aligned points giving a value between 0-1. To create a similarity matrix, each value was subtracted from 1 and then z-scored.

A surface measure of image differences that controlled for the position of the facial features in the image was also calculated. To do this each of the 25 original images was reshaped (using a wavelet-based Markov random field sampling method) to the average shape across all 25 images (Tiddeman, Stirrat, & Perrett ,2005). This removed any underlying shape cues to expression (as all images now shared exactly the same set of fiducial points), but left the surface information relatively unchanged. The pixel values from each face were then correlated for the same image pair combinations as for the procrustean analysis. These pixel correlations were transformed using Fisher's Z-transform. The values were z-scored to create an average surface similarity measure between each expression pairing.

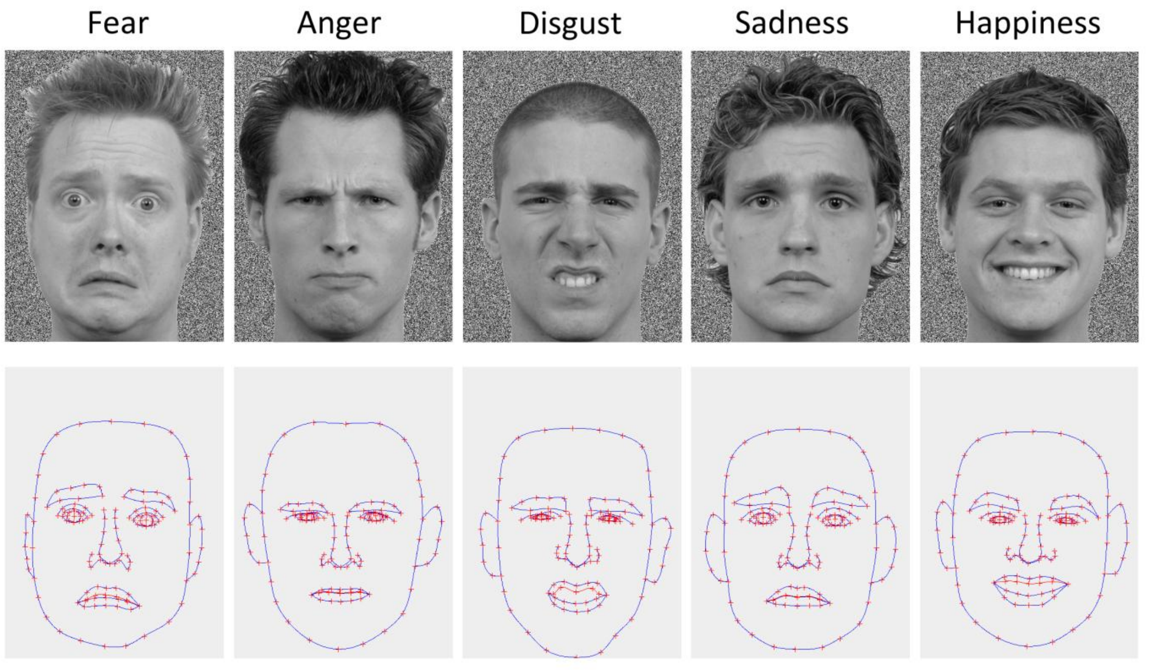

Figure 4.2 Exemplars of faces posing different expressions (top) and the location of the key fiducial points in each face (bottom).

4.3.5 fMRI experiment

To determine whether the patterns of perceptual similarity response in our behavioural task could be explained by patterns of response in face-selective regions, the neural response in face-selective regions to different facial expressions was measured. A block design was used with each block comprising a series of face images depicting one of the five expressions (fear, anger, disgust, sadness and happiness). Within each block, 5 images were each presented for 1 second followed by a 200 ms fixation cross, giving a block duration of 6s (Peirce, 2008). Stimulus blocks were separated by a fixation cross on a grey screen for 9s. Each condition was repeated eight times in a counterbalanced order, giving a total of 40 blocks. To minimise any influence of task effects on the patterns of neural response to expression, participants were not required to respond to the facial expressions during the fMRI scan. Instead, an irrelevant task of pressing a button when a red spot appeared was used to ensure that they paid attention to the stimuli without responding to their expressions per se. A small red spot appeared on 1 or 2 images in each block and participants were instructed to press a response button whenever they saw the red spot. Participants correctly detected the red spot on over 90% of trials (mean accuracy = 95.3 $\pm$ 2 %, SD = 2).

Scanning was performed at the York Neuroimaging Centre at the University of York with a 3 Tesla HD MRI system with an eight channel phased array head coil (GE Signa Excite 3.0 T, High resolution brain array, MRI Devices Corp., Gainesville, FL). Axial images were acquired for functional and structural MRI scans. For fMRI scanning, echo-planar images were acquired using a T2*-weighted gradient echo sequence with blood oxygen level-dependent (BOLD) contrast (TR = 3 s, TE = 32.7 ms, flip-angle = 90°, acquisition matrix 128 x 128, field of view = 288 mm x 288 mm). Whole head volumes were acquired with 38 contiguous axial slices, each with an in-plane resolution of 2.25 mm x 2.25 mm and a slice thickness of 3 mm. T1-weighted images were acquired for each participant to provide high-resolution structural images using an Inversion Recovery (IR = 450 ms) prepared 3D-FSPGR (Fast Spoiled Gradient Echo) pulse sequence (TR = 7.8 s, TE = 3 ms, flip-angle = 20°, acquisition matrix = 256 x 256, field of view = 290 mm x 290 mm, in-plane resolution = 1.1 mm x 1.1 mm, slice thickness = 1 mm). To improve co-registration between fMRI and the 3D-FSPGR structural image a high resolution T1 FLAIR was acquired in the same orientation planes as the fMRI protocol (TR = 2850 ms, TE = 10 ms, acquisition matrix 256 x 224 interpolated to 512 giving effective in-plane resolution of 0.56 mm). First-level analysis of the facial expression scan was performed with FEAT v 5.98. The initial 9s of data were removed to reduce the effects of magnetic stimulation saturation. Motion correction (MCFLIRT, FSL) was applied followed by temporal high-pass filtering (Gaussian-weighted least-squares straight line fitting, sigma = 120s). Spatial smoothing (Gaussian) was applied at 6 mm (FWHM). Individual participant data were entered into a higher-level group analysis using a mixed-effects design (FLAME, http://www.fmrib.ox.ac.uk/fsl). Parameter estimate maps were generated for each experimental condition; fear, anger, disgust, sadness and happiness. These maps were then registered to a high-resolution T1-anatomical image and then onto the standard MNI brain (ICBM152). Regions defined by the localiser scan were used to constrict MVPA analyses to face-responsive regions only.

To identify face-selective regions, data from a series of localizer scans with a different set of participants (n = 83) was used (Flack et al., 2014). The localizer scan included blocks of faces and scrambled faces. Images from each condition were presented in a blocked design with five images in each block. Each image was presented for 1 s followed by a 200-ms fixation cross. Individual participant data were entered into a higher-level group analysis using a mixed-effects design (FLAME, http://www.fmrib.ox.ac.uk/fsl). Face-responsive regions of interest were defined by the contrast of faces>scrambled faces at the group level and spatially normalised to an MNI152 standard brain template. The peak voxels for the OFA, FFA and STS in each hemisphere were determined from the resulting group statistical maps. Then the 500 voxels with the highest z-scores within each region were used to generate a mask. Masks were combined across hemispheres to generate 3 masks for the OFA, FFA and posterior STS, which form the core face-selective regions in Haxby et al's (2000) neural model .

Parameter estimates in the main experimental scan to each expression were normalised independently in each voxel by subtracting the mean parameter estimate across all expressions and then registered onto the standard MNI152 brain. Pattern analyses were then performed using the correlation-based MVPA method devised by Haxby and colleagues (Haxby, Gobbini, Furey, Ishai, Schouten & Pietrini, 2001). After separating the data across odd and even blocks for each participant (as was done by Haxby, et al., 2001), we determined the reliability of the patterns within participants by correlating patterns across odd and even runs for each condition. This procedure was performed 24 times (i.e. once for each participant) for each of the 15 possible combinations of basic emotions. The final correlation matrix provides a measure of the similarity in the pattern of response across different combinations of facial expressions. These neural correlations were transformed using Fisher's Z-transform and then converted into z scores.

4.3.6 Regression analyses

To then determine whether the pattern of perceptual similarity responses was best predicted by variance in facial shape or surface information, a linear regression analysis was performed using the similarity matrix for shape and surface analyses as independent regressors and the perceptual similarity rating correlation matrices from each individual as outcomes. The linear regression method used in the current study is similar to a Representational Similarity Analysis (RSA; Kriegeskorte, Mur, & Bandettini, 2008; Kriegeskorte, 2009) which can characterise the information carried by a given representation in behavioural response patterns, neural activity patterns or a representational model. By analyzing the correspondence between participant responses and neural response it is possible to test and compare different models. For example if either the shape or surface regressors are able to explain a significant amount of the variance in the corresponding perceptual similarity rating matrices, the model regression coefficient can be expected to be significantly greater than zero. All regressor and outcome variables were Z-scored prior to the regression analysis. However, it is important to note that the similarity responses are not fully independent. The same method was used to measure similarity between predictor models based on neural response patterns in OFA, FFS and STS regions and perceptual ratings of expression similarity as outcomes.

4.4 Results

4.4.1 Perception of facial expression is predicted by shape and surface properties of the image

Figure 4.3 shows the average perceptual similarity scores for each of 15 possible combinations of facial expression across all participants. The extent to which perceptual similarity of facial expressions could be predicted by the normalized shape and surface properties of the image was determined, by generating a corresponding similarity matrix for these image properties. The group averaged matrix for perception was significantly correlated with both shape (r (15) =.61, p=.016) and surface (r (15) =.77, p<.001) properties. In these analyses, images were normalized through rigid realignment of fiducial positions in the shape (procrustes) analysis and through a non-rigid transform to create fixed-shape images for the measure of surface similarity.

An important question concerns whether these transforms were necessary, or superfluous because the same characteristics were present in low-level properties of the untransformed images. A similar analysis with the raw images failed to show a significant relationship between perception and either shape ($r(15)=.27$, $p=.31$) and surface ($r(15)=.37$, $p=.16$) properties. This suggests that the mechanism underlying the perception of facial expression involves some form of equivalent normalization process.

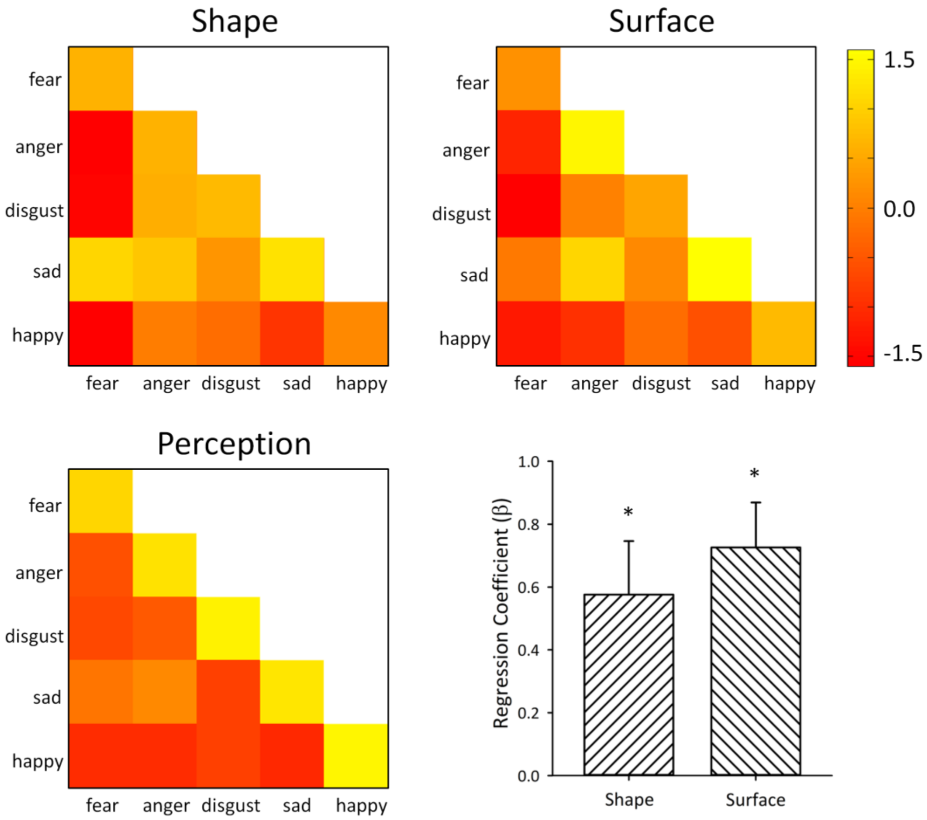

Figure 4.3 Regression analyses of the perceptual similarity data with shape and surface properties of the image. The analysis shows that the perceptual similarity of facial expressions can be predicted by both the shape and surface properties of the face. Error bars represent 95% confidence intervals. * denotes p<.001. Colour bars for each grid represent z score scale.

To measure the reliability across participants, a regression analysis was performed in which the models derived from the shape or surface analyses were independently used as predictor variables and the perceptual similarity ratings matrices from each individual as outcomes (Kriegeskorte et al. 2008). First, the image property models (shape model and surface model) were tested for colinearity and found not to be colinear. The variance inflation factor (VIF) value for the shape and surface models was 3.17, which does not exceed the recommended threshold of 5 (Montgomery, Peck, & Vining, 2012). The output of the regression analysis shows that the perceptual similarity of the facial expression could be explained by both the shape ($F(1,358) = 178$, $\beta=.58$, $p<.001$) and the surface ($F(1,358) = 399.5$, $\beta=.73$, $p<.001$) properties in the images.

In Figure 4.3 it is clear that the perceptual similarity between expressions can in part be driven by high similarity ratings along the diagonal (where one fear expression is seen as very similar to another fear expression, and so on). These will henceforth be referred to as within-category comparisons. To determine the extent to which these within-category comparisons were responsible for the result of the regression analysis, the analysis was repeated with just the between-category (off-diagonal) comparisons, looking to see whether the pattern of perceptual similarities between different expressions might still be tracked by the image properties. Again, it was found that the perceptual similarity of the expressions was significantly predicted by both the shape ($F(1,238) = 51.81$, $\beta=.42$, $p<.001$) and the surface ($F(1,238) = 61.14$, $\beta=.46$, $p<.001$) properties of the image, offering strong evidence of their importance.

4.4.2 Perception of facial expression is predicted by neural responses in face-selective regions

Figure 4.4 shows the average correlation matrix for expressions involving each of the 15 possible combinations of basic emotions in each of the core face-selective regions. To measure the reliability of the neural response to each facial expression, the data were analysed in each face responsive region with a 5 x 2 repeated measures ANOVA with Comparison (within-category, between-category) and Expression (Fear, Anger, Disgust, Sadness and Happiness) as factors. There was a significant main effect of Comparison in the STS ($F(1, 23) = 5.27$, $p=.03$) and OFA ($F(1, 23) = 6.45$, $p=.018$), but not in the FFA ($F(1, 23) = 0.067$, $p=.8$). This suggests that there are reliable patterns of response to facial expression in STS and OFA. We did not find any effect of Expression (STS: $F(4, 92) = .59$, $p=.67$, OFA: $F(4, 92) = 1.2$, $p=.31$, FFA: $F(4, 92) = 1.38$, $p=.248$) or any interaction between Comparison and Expression (STS: $F(4, 92) = .77$, $p=.55$, OFA: $F(4, 92) = .94$, $p=.45$; FFA: $F(4, 92) = 1.32$, $p=.27$) in any of the core face-selective regions. This suggests that the ability to discriminate expressions was not driven by any specific expressions, but rather by a generalised ability to discriminate all patterns of neural response to expressions.

Next, it was determined how the pattern of perceptual similarity might be linked to the patterns of response in different face-selective regions. The similarity of patterns of response to different facial expressions in each face-selective region were compared (see Fig. 4.4) with perceived similarity of the expressions (see Fig. 4.3). There was a significant correlation between perception and patterns of response in the STS ($r(15) = 0.62$, $p = .014$) and OFA ($r(15) = 0.67$, $p<.001$). However, there was no significant correlation between perception and patterns of neural response in the FFA ($r(15) = -0.08$, $p=.77$).

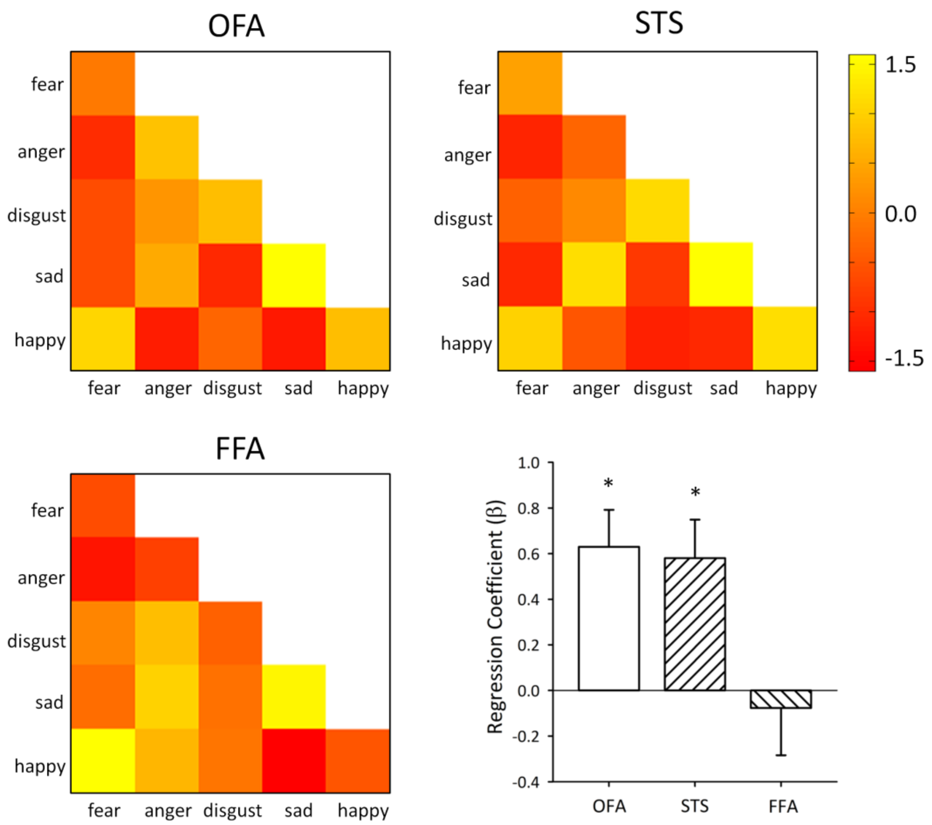

Figure 4.4 Regression analyses of the perceptual similarity data (shown in Figure 4.2) with the fMRI data from different face-selective regions. The analysis shows that the perceptual similarity of facial expressions can be predicted by the pattern of response in the OFA and STS, but not in the FFA. Error bars represent 95% confidence intervals. * denotes p<.001. Colour bars for each grid represent z score scale.

To measure the reliability across participants, a linear regression analysis was used with the neural responses in the different face responsive regions (OFA, FFA and posterior STS) responses as individual regressors and the perceptual similarity ratings matrices from each individual as the outcome. The perceptual similarity of the facial expressions could be predicted by neural response to facial expressions in STS ($F(1,358) = 181.2$, $\beta=.58$, $p<.001$) and OFA ($F(1,358) = 235.7$, $\beta=.63$, $p<.001$) regions but not in the FFA region ($F(1,358) = 2.11$, $\beta=-.08$, $p = .15$).

Again, one possible interpretation of these results is that they might be driven primarily by the higher within-condition compared to between-condition correlations. To determine if this was the case, the analysis was repeated only using the off-diagonal elements of the correlation matrices. As before, results showed that the perceptual similarity of the facial expressions could be predicted by neural response to facial expressions in the STS ($F(1,238) = 7.18$, $\beta=.17$, $p<.008$) and OFA ($F(1,238) = 9.96$, $\beta=.25$, $p<.002$), but not in the FFA region ($F(1,238) = 1.5$, $\beta=.08$, $p = .22$).

To determine whether the patterns of response in the face-selective regions could be explained by the magnitude of response to different expressions, a univariate analysis was performed on each region of interest. Table 1 shows the % MR signal to each expression. In contrast to the MVPA, Table 1 shows that similar levels of activation were evident to all expressions within each region. A repeated measures ANOVA showed that there was an effect of Region ($F=63.0$, $p<0.001$), which was due to lower responses in the STS. However, there was only a marginal effect of Expression ($F=2.38$, $p=0.073$) and a marginal interaction between Region and Expression ($F=2.11$, $p=0.066$). This marginal interaction likely reflects a relatively larger response to happiness compared to other expressions in the OFA and STS, but a relatively larger response to fear compared to other expressions in the FFA. It may also reflect the low response to sadness in the FFA but the high response to sadness in the OFA and STS.

	Fear	**Anger**	**Disgust**	**Sad**	**Happy**
OFA	0.87 + 0.08	0.79 + 0.07	0.79 + 0.08	0.79 + 0.09	0.88 + 0.09
STS	0.34 + 0.08	0.33 + 0.07	0.28 + 0.08	0.31 + 0.09	0.34 + 0.09
FFA	0.82 + 0.08	0.73 + 0.07	0.72 + 0.07	0.70 + 0.08	0.81 + 0.08

Table 4.1 % MR signal in face-selective regions to different facial expressions.

Finally, the current study determined how the pattern of perceptual similarity might be linked to the patterns of response in regions outside the core face-selective regions. The localiser scan was able to define other face-selective regions in the inferior frontal gyrus (IFG), amygdala and precuneus, which are part of the extended face processing network. The similarity of patterns of response to different facial expressions in each face-selective region were compared with perceived similarity of the expressions. There was a significant correlation between perception and patterns of response in the IFG ($r(15) = 0.63$, $p = 0.01$), but not in the amygdala ($r(15) = 0.28$, $p = 0.31$) or precuneus ($r(15) = -0.25$, $p = 0.37$). However, when only the between-category comparisons were measured there were no significant correlations in any of these face regions (IFG: $r(10) = -0.01$, $p = 0.97$; amygdala: $r(10) = -0.33$, $p = 0.23$; precuneus: $r(10) = 0.02$, $p = 0.94$).

To determine whether regions outside the face-selective ROIs could also predict patterns of response to facial expression, the analysis was repeated using the Harvard Oxford anatomical masks (http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/Atlases). First it was asked whether there were distinct patterns of response to different facial expressions. From the 48 anatomical regions, only the Inferior Temporal Gyrus posterior (ITGp, $F = 17.9$, $p < .001$) and the Middle Temporal Gyrus posterior division (MTGp, $F = 4.3$, $p = .048$) showed distinct patterns (Suppl. Table 1). Next, the similarity of patterns of response to different facial expressions in each region were compared with perceived similarity of the expressions. In contrast to the face-selective ROIs, neither the ITGp ($r(15) = 0.15$, $p = .59$), the MTGp ($r(15) = 0.48$, $p = .077$) nor any other anatomical region showed a significant correlation between patterns of response and perceptual similarity.

4.5 Discussion

Facial expressions are signalled by complex patterns of muscle movements that create changes in the appearance of the face. The aims of the present study were to determine how our perception of expression is linked to (1) the image properties of the face and (2) the neural responses in face-selective regions. Together, these findings show that the mechanisms that underpin the perception of facial expression are tightly linked to both shape and surface properties of the image and to the pattern of neural response in specific face-selective regions.

The use of a measure of the perceptual similarity between expressions allows a more fine-grained analysis than the more standard method of categorizing each expression as one of the basic emotions (e.g. Mattavelli et al., 2013). Instead, the current study was able to track the magnitude of perceived differences between emotions, and to demonstrate that this pattern of between-category differences could still be modelled both from normalized image properties and from neural responses in STS. The fact that the link between image properties and perception was still evident when the within-category correlations (fear with fear, etc.) were removed from the analysis shows that the findings are not driven solely by the relatively high within-category relationships. Rather, it suggests a more continuous representation of facial expression involving a distinct between-category structure.

Different facial expressions can be defined by edge-based shape cues that result from changes in the shape of the internal features (Ekman, 1972; Bruce & Young, 1998, 2012). Previous studies have suggested that these shape cues are important for the perception of facial expressions (Bruce & Young, 1998; Magnussen et al., 1994; White, 2001; Harris et al., 2014). Although changes in facial expression also affect the surface properties of the face (Andrew J. Calder et al., 2001), this information has not been thought to be particularly diagnostic for discriminating facial expression (Bruce & Young, 1998). In this study, results showed that both the shape and surface properties correlated highly with perceptual judgements. So, while the present findings provide further support for the long held assertion that shape cues are important for the perception of expression, the novel finding from this study is that surface properties are as important. This usefulness of both types of cue may reflect the natural intercorrelation between shape and surface cues within many expressions. For example, fear expressions involve opening the mouth and widening the eyes (shape cues) and this creates salient contrast changes in the eye and mouth regions (surface cues).

Neuroimaging studies have previously revealed a number of regions that respond selectively to facial expression (Haxby et al., 2000; Allison et al., 2000). The current study found that the perceptual similarity of different facial expressions could be predicted by the similarity in the pattern of neural response in the OFA and STS. That is, facial expressions that were perceived as being similar had more similar neural patterns of response in these regions, which is of course consistent with Haxby et al.'s (2000) idea that they are important to the analysis of changeable aspects of faces such as expression. Our findings are also consistent with a recent study showing that patterns of neural response correlated with the perceptual similarity of dynamic facial expressions in the posterior superior temporal sulcus (Said, et al 2010). Indeed, the correspondence between perception and neural response in the superior temporal region is consistent with the role of this region in the perception of facial expression (Haxby et al., 2000; Winston et al., 2004; Engell & Haxby, 2007; Harris, Young & Andrews, 2012; Harris et al., 2014; Baseler, Harris, Young & Andrews , 2014; Pitcher, 2014; Psalta et al., 2014; Wegrzyn et al., 2015).

The OFA is thought to be the primary input area in the face processing network and has projections to both the STS and FFA (Haxby et al., 2000). However, more recently there is evidence that face processing can occur in the absence of input through the OFA (Rossion et al., 2003). The current study's finding that the OFA can decode expression and contains representations of perceived similarity of these images suggests that it is involved in representing facial expression. This fits with other studies showing that the OFA adapts to facial expression (Fox et al., 2009) and that applying TMS to the OFA disrupts the perception of facial expression (Pitcher, 2014).

In contrast to the STS and OFA, patterns of response in the FFA did not predict the perception of facial expression. Although the current study's findings are consistent with neural models that suggest that this region is important for the representation of relatively invariant facial characteristics associated with recognition of identity (Allison et al., 2000; Haxby et al., 2000), they contrast with more recent studies that have shown responses in the FFA can be linked to the perception of facial expression (Harry, Williams, Davis & Kim, 2013; Wegrzyn et al., 2015). One potentially crucial difference between the current study and these previous studies is that they asked only whether patterns of response to different facial expressions were distinct. The current study addressed the more fine-grained question of whether the perceptual similarity of different facial expressions can be explained by the similarity in the patterns of neural response.

A final point to consider is that we do not draw any direct inference about whether neural responses in the current study could be predicted by image measures of surface and shape. This is because the block design employed did not allow a large number of exemplars of each expressions to be used to measure neural response to each individual example of each facial expression. This is not a problem for the regression between the image models and behavioural ratings of similarity as these averages were drawn from an extremely large number of rating trials and were therefore more stable. The block design employed allowed a more coarse scale measurement of the mean neural response across all actors posing each expression, this reduced the number of sample points in the regression and therefore reduced the predictive power of any neural level regression. To increase predictive power and the ability to draw these inferences at the neural level it would be ideal to use an event related design which allows for comparison of 'stimulus rich' model matrices (as suggested by Kriegeskorte and colleagues , 2008) which contain the neural similarity for every pairwise combination of experimental stimuli. We did not employ this method due to the fact we prioritised the increased power allowed by a block design to discriminate neural patterns of response to facial expressions , however future experiments may benefit from employing a stimulus rich comparison design.

In conclusion, the current study has shown that perceptual patterns of response to facial expression are correlated with statistical properties of face images and with neural responses. The current study found that changes in both the shape and surface properties of the face predict perceptual responses to facial expression and that difference in the neural patterns of response in the STS, but not the FFA can also predict perceptual responses to facial expressions. Together, these results show the importance of image properties in understanding higher level perceptual judgements and suggest that these factors may be an important organizing principle for the neural representations underlying the perception of facial expression.

Chapter 5 - Contributions of feature shapes and surface cues to the neural representations of facial identity

5.1 Abstract

A range of behavioural evidence has shown the pre-eminence of surface compared to shape information for the recognition of facial identity. The aim of the current study was to compare the relative importance of shape and surface information in the neural representation of familiar faces. Image morphing techniques were used to generate hybrid faces that had the shape properties from one identity and the surface properties from a different identity. A block design fMRI experiment was used to test the sensitivity of face-selective regions in the human brain to the shape or surface properties of the face. Faces were presented in blocks in which either (1) the face shape was the same, but the surface properties varied or (2) the surface properties were the same, but the shape varied. Based on previous behavioural evidence, it was predicted that brain regions responsible for familiar face recognition should show distinct patterns of response to face blocks that had the same surface properties, but not show distinct patterns to face blocks that had the same shape properties. In contrast to this prediction, the FFA and OFA showed distinct patterns to face blocks that had the same shape properties, but failed to show distinct patterns to face blocks with the same surface properties. The dissociation between the neural results and previous behavioural findings suggests that representations of facial identity in these core face regions are not solely based on image properties.

5.2 Introduction

Neuroimaging studies have revealed a network of regions in the occipital and temporal lobe which form a core system for the visual analysis of faces (Kanwisher, McDermott, & Chun, 1997; Haxby, Hoffman, & Gobbini, 2000). These studies have consistently found regions that show stronger responses to faces than other visual stimuli in the fusiform gyrus (the fusiform face area, or FFA), occipital cortex (the occipital face area, or OFA) and the posterior superior temporal sulcus (pSTS). These three functionally localisable regions form a core system for the visual analysis of faces in the widely used neural model of Haxby, Hoffman and Gobbini (2000), with the FFA being thought to be particularly closely linked to the processing of relatively invariant facial characteristics such as identity. Here, we are interested in developing a more detailed analysis of the information that is represented in the FFA, and in particular whether it corresponds to the information that is critical to familiar face recognition.

Any facial image consists of a set of edges created by abrupt changes in reflectance due to the shapes and positions of facial features and a broader pattern of reflectance based on the surface properties of the face – also known as texture or albedo (Bruce & Young, 1998, 2012). Shape properties can be defined by the spatial location of fiducial points that correspond to key features of the face. In contrast, surface properties result from the pattern of reflectance of light due to the combination of ambient illumination, the face's pigmentation, and shape from shading cues. The distinction of shape from surface properties is widely used in face perception research (Bruce & Young, 1998, 2012) and is implicit in standard approaches to computer image manipulation (B. Tiddeman et al., 2001). These image manipulation techniques allow quasi-independent changes to a face's shape or surface properties. Such changes cannot be fully independent, of course, because many of the shape and surface properties of images will necessarily covary. For example, the surface property of shading is clearly affected in part by the face's shape. None the less, such methods allow us to hold face shape fixed as closely as possible (by using the same fiducial positions for a set of images) or to hold the surface properties fixed as closely as possible (by using the same surface brightness patterns in a set of images), as shown in Figure 5.1. This then allows a direct test of the relative contributions of shape and surface information.

The majority of experiments that have investigated the roles of shape and surface cues in face recognition have used unfamiliar faces. These studies have reported that both shape and surface cues can contribute to judgements of facial identity (O'Toole, Vetter, & Blanz, 1999; Jiang, Blanz, & O'Toole, 2006; Russell & Sinha, 2007). However, because participants performing unfamiliar face recognition tasks have no other experience with the faces, they have no way of separating aspects of appearance that are specific to those images from aspects of appearance that are stable across different viewpoints, lighting, and facial expressions. Therefore, performance on unfamiliar face recognition tasks could be driven by properties that are specific to particular images; it can involve picture recognition rather than true face recognition (Bruce, 1982; Hay & Young, 1982; Longmore, Liu, & Young, 2008).

Familiar face perception differs markedly from unfamiliar face perception because the participant has previous experience of seeing familiar faces across many different viewing conditions (Hancock, Bruce, & Mike Burton, 2000). This familiarity with a face allows recognition to proceed using invariant representations that are not affected by changes in viewpoint, lighting, and facial expression (Bruce & Young, 1986; Bruce, 1994; Burton, Bruce, & Hancock, 1999; Burton, 2013). A number of studies have shown that the surface properties of faces play a critical role in the invariant representation that is used for the recognition of familiar faces (Hole, George, Eaves, & Rasek, 2002; Burton, Jenkins, Hancock & White, 2005; Russell & Sinha, 2007). For example, familiar face recognition is not substantially affected if the surface properties are presented on a standardized shape (Burton et al , 2005), or when a face's shape is distorted by stretching the image (Hole et al., 2002). In contrast, line drawings of faces, which lack any surface properties, are not usually sufficient for recognition (Davies, Ellis, & Shepherd, 1978; Leder, 1999). The reason shape information may not be a reliable cue for the recognition of familiar face identity is thought to be that shape cues (particularly from the internal features of the face) are less invariant across different images of the same face (Burton, 2013). Despite this wealth of behavioural evidence, very few neuroimaging studies have yet directly tested the extent to which facial shape or surface (independently of each other) are employed in neural representations of facial identity.

Within the core system of face-selective regions, the fusiform face area (FFA) and occipital face area (OFA) are thought to be particularly important for the representation of invariant facial characteristics that convey facial identity (Haxby et al., 2000; Grill-Spector, Knouf, & Kanwisher, 2004). Consistent with the view that the FFA is crucial to the perception of facial identity a number of recent Multi Voxel Pattern Analysis (MVPA) fMRI studies (e.g. Liu, Harris, & Kanwisher, 2010; Natu et al., 2010; Nestor, Plaut, & Behrmann ,2011) have succeeded in finding distinct patterns of neural response to individual facial identities within the FFA. However during these experiments, facial shape and surface changes between facial identities were not held constant or manipulated in any consistent manner, therefore it is still unknown to what extent the neural response patterns to facial identity can be discriminated on the basis of facial shape or surface features.

The aim of the current study was to directly test the relative importance of facial shape and surface image properties in the neural representation of facial identity (Burton et al., 2005). The current study used a Leave One Participant Out Multi Voxel Pattern Analysis (Haxby et al., 2001) to attempt to discriminate between patterns of neural response to different facial identities. This different MVPA was used for the increased statistical power it allows. The LOPO MVPA approach also provides validation across participants that they all showed similar patterns of neural response to shape or surface cues to facial identity rather only within participant validation that is offered by the odd vs. even run approach used un chapter 4. It was hypothesised that in keeping with previous behavioural evidence, that brain regions responsible for familiar face recognition should show distinct patterns of response to face blocks that had the same surface properties, but not show distinct patterns to face blocks that had the same shape properties.

5.3 Methods

5.3.1 Participants

Twenty healthy volunteers took part in the fMRI experiment (10 female, mean age = 25 years 6 months SD = 3 years, 10 months. All participants were right-handed and had normal or corrected to normal vision with no history of neurological illness. The fMRI work was approved and conducted following the guidelines of the York Neuroimaging Centre Research Ethics Committee, University of York. All participants gave written consent prior to their participation.

5.3.2 Stimuli and design

Figure 5.1 shows the original images from the eight conditions used in the fMRI experiment. Hybrid images of the unique shape of one actor and the unique surface of another were combined (stimuli never had shape and surface cues from the same actor). Stimuli were organised into eight blocks of 6 images that had either a consistent shape or consistent surface of an actor whilst the other cue varied. This gave 4 shape invariant blocks (Kyle Shape, Lineker Shape, Paxman Shape and Sugar Shape) and 4 surface invariant blocks (Kyle Surface, Lineker Surface, Paxman Surface and Sugar Surface) these images were created using well-validated images from the Burton et al. (2005) face set. Images were selected on the basis of high recognisability of the individual pictured; participants had to successfully verbally identity a picture of each celebrity used in the experiment with 100% accuracy to be allowed to participate in the fMRI experiment.

A block fMRI design was used with eight stimulus conditions. Within each block, 6 images of representing an invariant facial shape or surface cue (depending on the experimental condition for that block) were each presented for 800 ms followed by a 200 ms fixation cross, giving a block duration of 6s. Stimulus blocks were separated by a fixation cross on a grey screen for 9s. Each condition was repeated six times in a counterbalanced order, giving a total of 48 blocks. To minimise any influence of task effects on the patterns of neural response to expression, participants were not required to respond to the facial identities during the fMRI scan. Instead, an irrelevant task of pressing a button when a red spot appeared was used to ensure that they paid attention to the stimuli without responding to their identities per se. A small red spot appeared on 1 or 2 images in each block and participants were instructed to press a response button whenever they saw the red spot. Participants correctly detected the red spot on over 90% of trials (mean accuracy = 95.6 %, SD = 3.2).

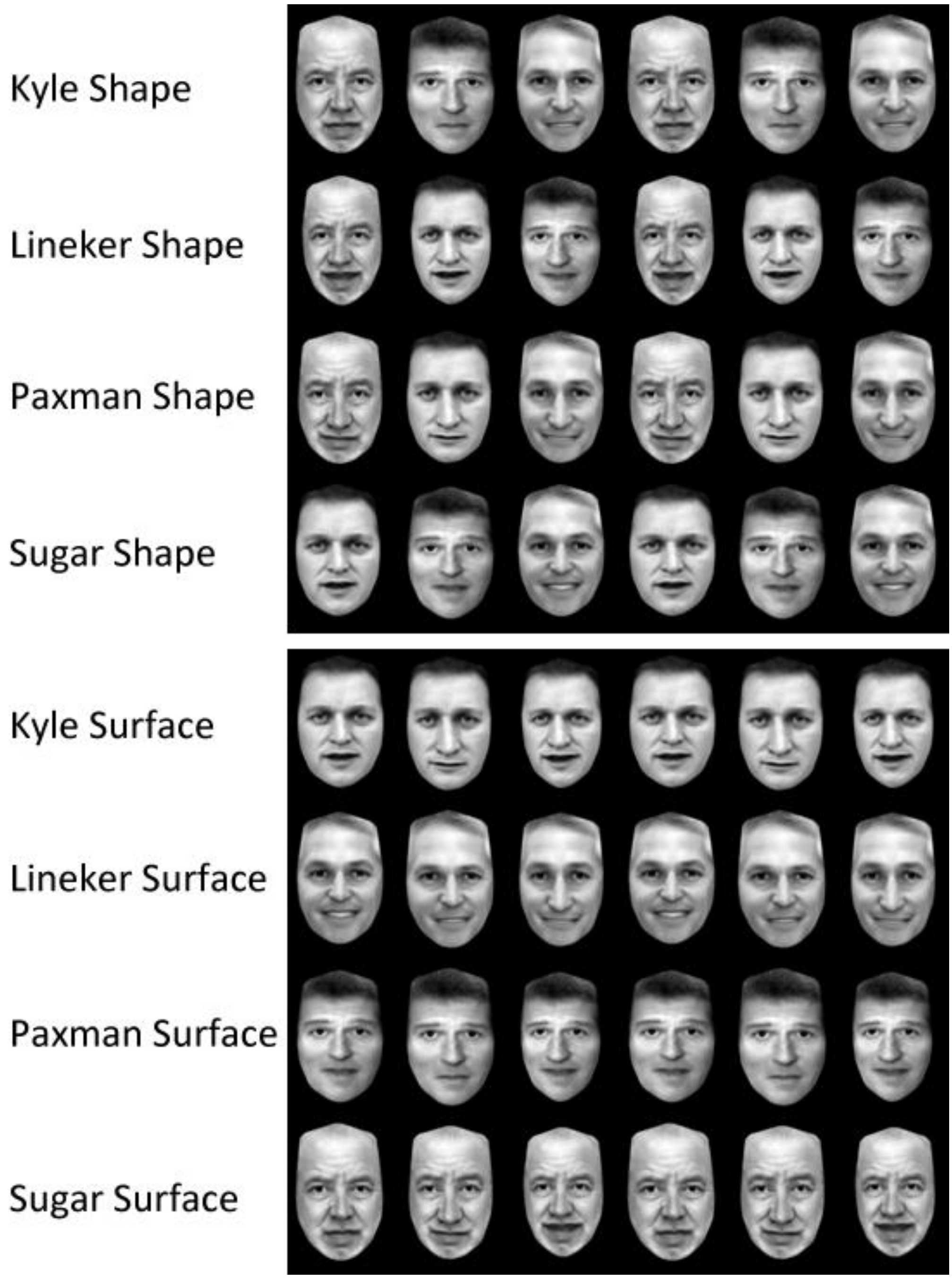

Figure 5.1: Stimulus conditions used in the fMRI experiment. Images were generated from the facial identities of British celebrities Jeremy Kyle, Gary Lineker, Jeremy Paxman and Sir Alan Sugar. Each row represents a stimulus condition. Images in each row were presented in a block. The top four rows represent stimulus conditions in which the shape was specific to one of the 4 identities, the bottom four rows represent stimulus conditions in which the surface was specific to one of the four identities.

5.3.3 fMRI Acquisition Parameters

Scanning was performed at the York Neuroimaging Centre at the University of York with a 3 Tesla HD MRI system with an eight channel phased array head coil (GE Signa Excite 3.0 T, High resolution brain array, MRI Devices Corp., Gainesville, FL). Axial images were acquired for functional and structural MRI scans. For fMRI scanning, echo-planar images were acquired using a T2*-weighted gradient echo sequence with blood oxygen level-dependent (BOLD) contrast (TR = 3 s, TE = 32.7 ms, flip-angle = 90°, acquisition matrix 128 x 128, field of view = 288 mm x 288 mm). Whole head volumes were acquired with 38 contiguous axial slices, each with an in-plane resolution of 2.25 mm x 2.25 mm and a slice thickness of 3 mm. T1-weighted images were acquired for each participant to provide high-resolution structural images using an Inversion Recovery (IR = 450 ms) prepared 3D-FSPGR (Fast Spoiled Gradient Echo) pulse sequence (TR = 7.8 s, TE = 3 ms, flip-angle = 20°, acquisition matrix = 256 x 256, field of view = 290 mm x 290 mm, in-plane resolution = 1.1 mm x 1.1 mm, slice thickness = 1 mm). To improve co-registration between fMRI and the 3D-FSPGR structural image a high resolution T1 FLAIR was acquired in the same orientation planes as the fMRI protocol (TR = 2850 ms, TE = 10 ms, acquisition matrix 256 x 224 interpolated to 512 giving effective in-plane resolution of 0.56 mm).

5.3.4 Localiser Scan

To identify face-selective regions, data from an independent series of localizer scans with a different set of participants (n = 83) were analyzed by comparing the response to intact faces compared to phase-scrambled faces for further details see Sormaz et al., 2016). This use of an independent localiser was intended to allow us meaningfully to compare patterns of activation across individual participants with the LOPO-based MVPA method described below. This procedure produced masks in the following regions: Occipital Face Area (OFA), Fusiform Face Area (FFA) and Superior Temporal Sulcus (STS).

5.3.5 Facial Identity Scan

First-level analysis of the facial identity scan was performed with FEAT v 5.98. The initial 9s of data were removed to reduce the effects of magnetic stimulation saturation. Motion correction (MCFLIRT, FSL) was applied followed by temporal high-pass filtering (Gaussian-weighted least-squares straight line fitting, sigma = 120s). Spatial smoothing (Gaussian) was applied at 6 mm (FWHM). Individual participant data were entered into a higher-level group analysis using a mixed-effects design (FLAME, http://www.fmrib.ox.ac.uk/fsl). Parameter estimate maps were generated for shape invariant conditions (Kyle Shape, Lineker Shape, Paxman Shape, Sugar Shape) and surface invariant conditions (Kyle Surface, Lineker Surface, Paxman Surface and Sugar surface). These maps were then registered to a high-resolution T1-anatomical image and then onto the standard MNI brain (ICBM152). Regions defined by the independent localiser scan were used to constrict MVPA analyses to face-responsive regions only.

Parameter estimates to each condition were normalised independently in each voxel based on whether the neural response came from shape invariant or surface invariant block. For shape invariant blocks, normalisation was carried out by subtracting the mean parameter estimate across all shape invariant conditions for that voxel. For surface invariant conditions normalisation was carried out in the same way, by subtracting average response to all surface invariant conditions from each voxel. This was done to account for any differential response in the brain to shape or surface image cues. Pattern analyses were then performed using an adapted variant of the correlation-based MVPA method devised by Haxby and colleagues (Haxby et al., 2001; Hanke et al., 2009); http://www.pymvpa.org/). However, rather than separating the data across odd and even runs for each participant (as was done by Haxby et al., 2001), we determined the reliability of the patterns across participants using a leave-one-participant-out (LOPO) method (Rice, Watson, Hartley, & Andrews, 2014; Watson, Hartley, & Andrews, 2014). This method involves taking a parameter estimate map from each individual and correlating it with the parameter estimate map from a group analysis of the remaining 19 participants. This procedure was performed 20 times (i.e. once for each participant) for each of the 10 possible combinations of experimental conditions. The final correlation matrix provides a measure of the similarity in the pattern of response across different combinations of experimental conditions.

5.4 Results

Figure 5.2 shows the average correlation matrix for the pattern of response in face-selective regions to all combinations of facial identity in which shape was invariant within a block.

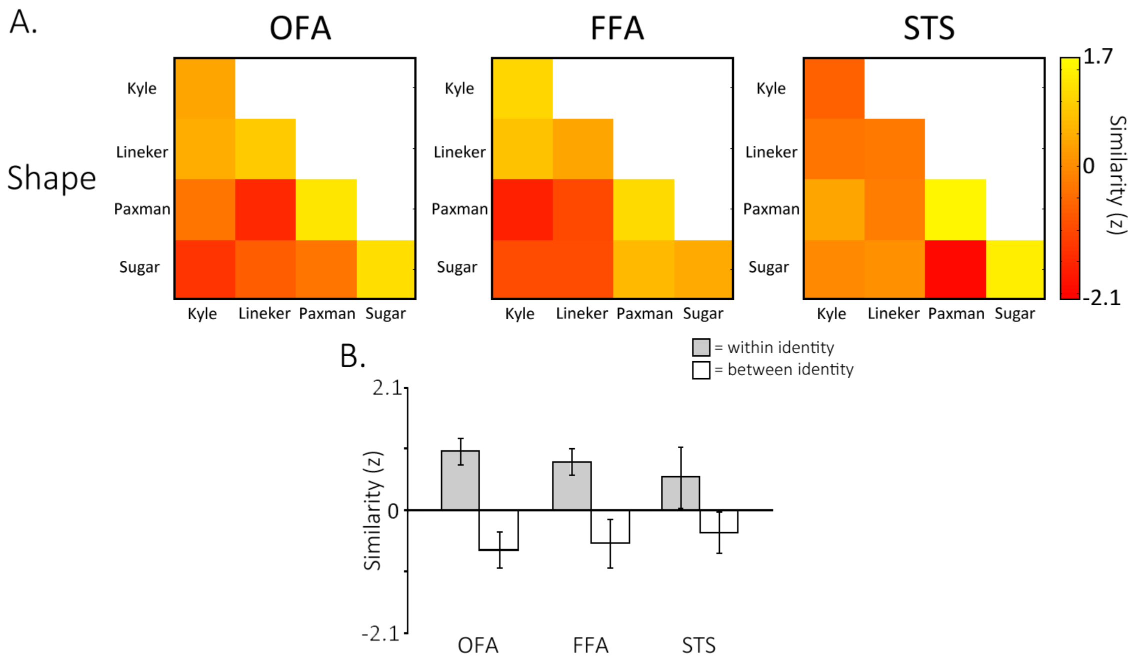

Figure 5.2: A) Similarity matrices for the pattern of neural response across different face-selective regions for shape invariant blocks. Correlation coefficients are expressed in the colour bars as z-scores representing standard deviations from the mean for each correlation coefficient in each face responsive region. B) The z-scored mean within vs. between level correlation coefficients for shape invariant identity blocks. Greater within condition correlation coefficients than greater than between level correlation coefficients reflected the ability of the OFA and FFA to significantly discriminate between identities on the basis of invariant shape information (both p<.05).

To measure the reliability of the neural response to each face, the data in each face responsive region for each condition were separately analysed using a 4 x 2 repeated measures ANOVA with Identity (Jeremy Kyle, Gary Lineker, Jeremy Paxman and Alan Sugar), and Comparison (within-category, between-category) as main factors. For shape invariant conditions in the OFA there was no significant main effect of identity ($F(3, 57) = 2.19$), $p=.1$), but there was a significant main effect of comparison ($F(1, 19) = 20.71$, $p<0.001$) with higher within-category compared to between-category correlations. There was no significant interaction of identity x comparison ($F(3, 57) = 0.84$, $p = 0.48$) showing that no particular facial identity shape disproportionately contributed to the discriminability of identities.

For shape invariant conditions in the FFA there was a no significant main effect of identity ($F(3, 57) = 1.06$), $p=0.13$), but there was a significant main effect of comparison ($F(1, 19) = 25.09$, $p<0.001$) with higher within-category compared to between-category correlations demonstrating a generalised ability for the FFA to discriminate between shape based identity information for all conditions. There was a marginally significant interaction of identity x comparison ($F(3, 57) = 2.51$, $p = 0.07$, Greenhouse-Geisser corrected due to violation of sphericity). The lack of significant interaction means that all conditions were contributing approximately equally to the discrimination of within vs. between category correlations.

For shape invariant conditions in the STS there was a significant main effect of identity ($F(3, 57) = 2.84$), $p=0.046$). There was no significant main effect of comparison ($F(1, 19) = 3.12$, $p=0.094$) with higher within-category compared to between-category correlations. There was a significant main interaction of identity x cue x comparison ($F(3, 57) = 2.84$, $p = 0.046$). We probed these differences with post hoc t tests and found the interaction was caused by Alan Sugar face shape information conditions having a significantly higher within vs. between category condition correlation (p=.017) and therefore being more discriminable when compared to other identities.

Figure 5.3 shows the average correlation matrix for the pattern of response in face-selective regions to all combinations of facial identity in which surface properties were invariant identity cues within a block.

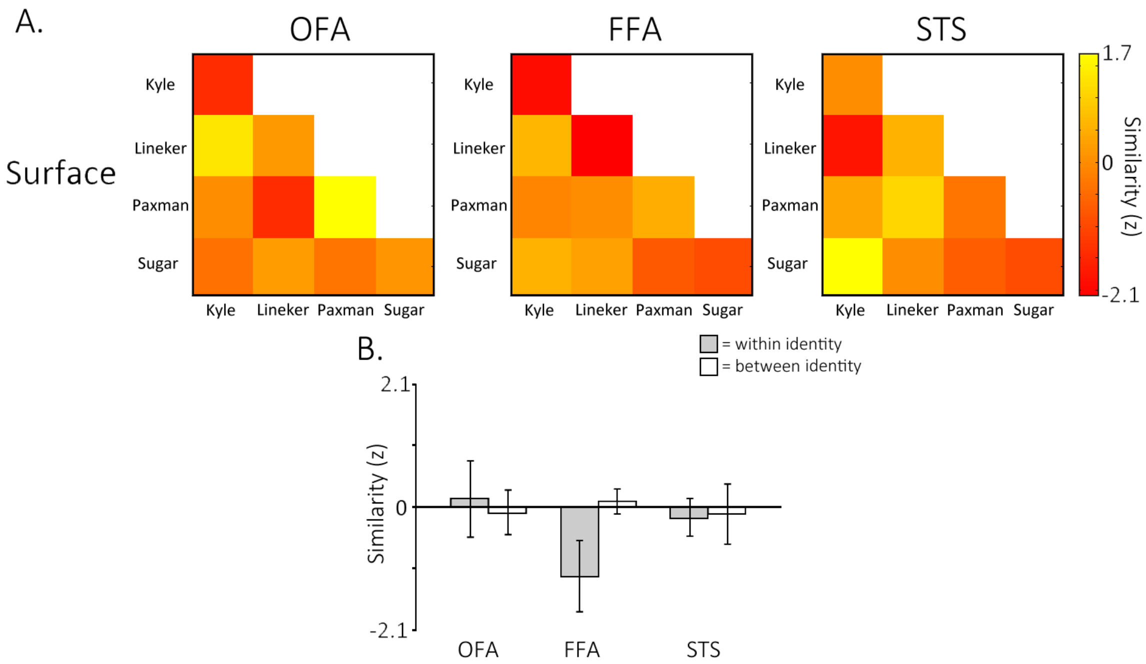

Figure 5.3: A) Similarity matrices for the pattern of neural response across different face-selective regions for surface invariant blocks. Correlation coefficients are expressed in the colour bars as z-scores representing standard deviations from the mean for each correlation coefficient in each face responsive region. B) The z-scored mean within vs. between level correlation coefficients for shape invariant identity blocks. There was not within condition correlation coefficient significantly greater than a between level correlation coefficient. This signified that no face responsive brain regions were able to discriminate between identity on the basis of invariant surface information.

For surface invariant conditions in the OFA there was a significant main effect of identity ($F(3, 57) = 3.6$), $p=.02$) reflecting differences in neural response to different identities.

There was no significant main effect of comparison ($F(1, 19) = .106$, $p=0.75$), but here was a significant interaction of identity x comparison ($F(3, 57) = 3.93$, $p = 0.013$). The significant interaction means there was some difference in the level of within vs. between category correlations across identities. We probed these differences with post hoc t tests and found the interaction was caused by Jeremy Kyle face surface information having a significantly lower within compared to between category condition correlations when compared to other identities ($p<.001$).

For surface invariant conditions in the FFA there was a significant main effect of identity ($F(3, 57) = 2.84$), $p=0.046$). There was no significant main effect of comparison ($F(1, 19) = 3.06$, $p=0.1$) There was a marginally significant interaction of identity x comparison ($F(3, 57) = 2.98$, $p = 0.039$. The significant interaction means there was some difference in the level of within vs. between category correlations across identities. We probed these differences with post hoc t tests and found the interaction was caused by Jeremy Kyle and Gary Lineker face surface information having a significantly lower within vs. between category condition correlations (both $p<.001$) when compared to other identities.

For surface invariant conditions in the STS there was no significant main effect of identity ($F(3, 57) = .616$), $p=0.61$). There was no significant main effect of comparison ($F(1, 19) = 0.002$, $p=0.96$) with higher within-category compared to between-category correlations. There was also no significant main interaction of identity x comparison ($F(3, 57) = 0.153$, $p = 0.93$).

5.5 Discussion

The current study aimed to find consistent patterns of neural response to invariant shape and surface cues to identity. Results showed that consistent patterns of neural response that discriminated between identities could only be found in conditions where identity information was based on invariant facial shape information. Conditions where facial shape was the invariant identity cue identified consistent neural patterns in the Fusiform Face Area and Occipital Face Area but not the Superior Temporal Sulcus. In conditions where facial surface was the invariant identity cue there were no consistent patterns of neural response in core face responsive regions to discriminate between identities. These findings show that it is possible to find consistent patterns of neural response in core face responsive regions to facial shape information.

The findings of the current study are striking because recent evidence from Anzellotti & Caramazza (2014) and (2015) suggests that it is easier to discriminate invariant features of facial identity (e.g. facial identity across viewpoint) in the anterior temporal lobes than the FFA. This mirrors earlier findings by Kriegeskorte, Formisano, Sorger, & Goebel (2007) showing that patterns of neural response to facial identity were easier to decode in anterior temporal brain regions rather than in the FFA suggesting that the anterior temporal regions hold representations of identity based in invariant features. In contrast to these studies, the current study demonstrates both that consistent neural response patterns to facial identity can be seen FFA and that these neural responses are evoked by the physical characteristics of the face (namely shape characteristics).

With this in mind it should be noted that the current study provided the first MVPA based, systematic test of the extent to which neural responses that encode facial identity are based on facial shape and surface features. The MVPA analysis was able to firstly, discriminate facial identities in the FFA in accordance with prior MVPA studies that have also been able to discriminate identities in the FFA (e.g. Natu et al., 2010; Liu, Harris, & Kanwisher, 2010; Nestor, Plaut, & Behrmann, 2011). However, importantly the MVPA analysis in the current study provided an advance on previous studies by demonstrating that both the OFA and FFA are able to decode facial identities based on shape based facial information. This finding provides further evidence that representations of facial identity are not divorced from facial image properties (specifically shape based image properties) in core face responsive regions as had been previously suggested (Ramon, Dricot, & Rossion, 2010; Kriegeskorte, Formisano, Sorger, & Goebel, 2007; Anzellotti & Caramazza, 2015). Interestingly the findings of the current study reflect a similar pattern of results found by Liu et al., (2010), who employed an MVPA paradigm to discriminate between neural responses faces in the FFA on the basis of facial features being in either correct or scrambled feature configuration. Liu and colleagues showed that the FFA is sensitive to facial feature configurations; this is interesting as one of the main invariant shape cues to identity in the current study was the shape and positioning of facial features for each face. Taken together the findings of Liu et al (2010) and the current study suggest that the FFA may contain representations of facial identity based on facial feature configuration cues.

Interestingly, the current study demonstrated that the OFA has the ability to discriminate between facial identities in the on the basis of invariant facial shape cues. This reflects the possibility that the OFA, an area implicated in lower level visual processing of the face stimulus encodes enough identity specific shape information for effective identity discrimination. This result is interesting as the aforementioned study by Liu et al (2010) found that the OFA was not sensitive to facial feature configuration, rather only the presence of facial features. Despite this the similarity of the ability of the ability of the OFA and FFA to decode facial identity is not surprising as they are both anatomically proximal have central roles in facial identity processing (Haxby, Hoffman, & Gobbini, 2000; Bruce & Young, 2012). Both regions have also been shown to be sensitive to physical relationships between facial features that are key to facial identity (Rhodes et al., 2009) and have been shown to display similarity in their level of sensitivity of tuning to specific facial identities (Gratton, Sreenivasan, Silver, & D'Esposito, 2013). In accordance with previous studies that have shown that the OFA is more sensitive to facial shape than surface properties (e.g. Yue et al., 2013) the current study showed that the OFA could discriminate identities based on facial shape information only. This supports previous assertions that neural populations facial representations of identity in the OFA may be represented by facial shape cues rather than facial surface cues.

The findings that both the OFA and FFA can discriminate identities based on physical shape properties of the image are consistent with suggestions that areas appearing to respond to stimulus categories (e.g. selective FFA response to faces) may emerge from neural activation in nearby visually mapped regions containing lower level image based representations of the face stimulus (e.g. Op de Beeck, Haushofer, & Kanwisher, 2008; Andrews et al., 2015). This explanation has been supported by recent findings from Rice et al. (2014) who found that category selectivity in higher visual cortex arose from the low level image properties of each category, this included neural response to face stimuli in face responsive brain regions. In the case of the current study this may be because the FFA and OFA masks were in close proximity to each other and therefore contained overlapping topographic maps that contributed to the image based representation of the face image. Nevertheless, the results in the current study demonstrating that the FFA does contain image based representations of facial identity are novel.

The relatively higher sensitivity of face responsive regions to facial identity based on face shape rather than surface information was unexpected as previous behavioural research shows that both of facial shape and surface information are sufficient to produce approximately equal recognition rates of facial identities (e.g. Troje & Bülthoff, 1996; Russell et al., 2006) and even some advantage for surface based information (Russell & Sinha, 2007). Furthermore our results are contrary to previous findings from fMRI adaptation, for example Jiang et al (2009) showed that there was a release from adaptation both to shape and surface information in both FFA and OFA. These previous studies suggest that neural populations exist in OFA and FFA that code for facial identity based on shape and surface whereas our results show that neural patterns based on shape only can be discriminated. An explanation of the difference in results may be provided by a recent ERP study Caharel et al (2009) who showed that coding of facial shape specific identity information occurs earlier (reflected in the N170) than combined shape and surface information (reflected in the N250), similarly recent work by Itz, Schweinberger, Schulz, & Kaufmann (2014) has shown that ERP responses to facial shape occur earlier (N170 and N200) than to facial surface reflectance information (N250) , however, they posit that the later N250 response to surface information is crucial to the percept of familiar identities . In addition to this temporal difference between shape and surface information Lai, Oruç, & Barton (2011) found that although both facial shape and surface both elicited significant identity specific perceptual encoding after effects, the identity after effect was significantly greater for shape features rather than surface properties. The difference between the findings of the current study and previous studies may represent the possibility that the patterns of neural response to shape based identity cues we measured were a better measure of early neural responses on a time scale where shape information is more salient and being encoded more strongly than surface reflectance information which become more salient later. For this reason it may prove useful for future studies attempting to discriminate between facial identities on the basis of shape or surface information to use neuroimaging methods such as Magnetoencephalography (MEG) which can discriminate neural response patterns that happen at different time points.

One question that remains is to what extent the shape information used to discriminate identities in the current study was image or identity dependent. The current study used shape information generated from an average of 100 ambient front facing images of each individual. The strength of such an approach is that it allows the creation of highly recognisable images of each facial identity that contains a wider range of image based variance (e.g. lighting cues) contained across a large number of samples of the same individual (Burton et al, 2005) . This approach, however, limited us to using exactly the same shape or surface information to express each identity. This leaves open the possibility that our ability to discriminate facial identities on the basis of neural response patterns to facial shape was image specific rather than identity specific. The extent to which FFA representations of facial identity are image dependent is controversial as a number of fMR-adaptation studies have shown dependence of identity adaptation in the FFA on use of the same image of an individual (e.g. Andrews & Ewbank, 2004; Davies-Thompson, Gouws, & Andrews, 2009 ; Xu, Yue, Lescroart, Biederman, & Kim, 2009) whilst others in have shown FFA adaptation even when different images of the same individual are used (e.g. Rotshtein, Henson, Treves, Driver, & Dolan, 2005; Davies-Thompson & Andrews, 2012). Based on recent evidence (e.g. Yue, Tjan, & Biederman, 2006; Xu et al.,

2009) it is likely that the level of release from adaptation in OFA in FFA is proportional to the degree of change in the face image. Therefore whilst the ability to discriminate identity on the basis of facial shape in the current study is unlikely to be completely image dependent, future studies may benefit from using different image of the individuals to capture a greater degree of shape or surface variance within the same identity exemplars.

The lack of discriminability for patterns of neural response to surface information could be due to the spatial scale of the fMRI response. The fMRI response is limited by the size of the voxel. Each fMRI voxel samples thousands of neurons (N K Logothetis, 2008). It is possible therefore that MVPA paradigms can fail to discriminate significantly distinct neural patterns if they are distributed at a finer spatial scale (Haynes, 2015). For this reason it may be that unlike neurons tuned to facial shape identity information, neurons tuned to facial surface information may not be as widely distributed making their response pattern more difficult to discriminate.

In conclusion, we demonstrate that it is possible to discriminate facial identities on the basis of the neural response to their shape information in the OFA and FFA. In contrast patterns of neural response to invariant surface based identity information in the OFA, FFA and STS could not be discriminated. These findings show that although there are image based neural representation of facial identity, it is facial shape cues that are more easily discriminable using MVPA. This finding does not fit with previous behavioural experiment findings that suggest the pre-eminence of surface information for the recognition of identity.

Chapter 6 - Contrast negation and the importance of the eye region for holistic representations of facial identity

This Chapter is adapted from: Sormaz, M., Andrews, T. J., & Young, A. W. (2013). Contrast negation and the importance of the eye region for holistic representations of facial identity. *Journal of Experimental Psychology: Human Perception and Performance*, 39(6), 1667–77.[3]

6.1 Abstract

Reversing the luminance values of a face (contrast negation) is known to disrupt recognition. However, the effects of contrast negation are attenuated in chimeric images, in which the eye region is returned to positive contrast (Gilad, Meng, & Sinha, 2009). The present study probes further the importance of the eye region for the representation of facial identity. In the first experiment, it was asked to what extent the chimeric benefit is specific to the eye region. The results showed a benefit for including a positive eye region in a contrast negated face, whereas chimeric faces in which only the forehead, nose or mouth regions were returned to positive contrast did not significantly improve recognition. In Experiment 2, it was confirmed that the presence of positive contrast eyes alone does not account for the improved recognition of chimeric face images. Rather, it is the integration of information from the positive contrast eye region and the surrounding negative contrast face that is essential for the chimeric benefit. In Experiment 3, it was demonstrated that the chimeric benefit is dependent on a holistic representation of the face. Finally, in Experiment 4, it was shown that the positive contrast eye region needs to match the identity of the contrast negated part of the image for the chimera benefit to occur. Together, these results show the importance of the eye region for holistic representations of facial identity.

[3] The author, Mladen Sormaz, designed the experiments, analysed the results and wrote the article under the supervision of Prof. Andrew Young and Prof. Timothy Andrews.

6.2 Introduction

Reversing the luminance values of an image (contrast negation) disrupts the recognition of familiar faces (Galper, 1970; Galper & Hochberg, 1971; Phillips, 1972; Johnston, Hill, & Carman, 1992). The effect of contrast negation provides an interesting counter-example to the typical resistance of familiar face recognition to other types of image degradation (Bruce & Young, 2012), and contrast negation appears to disrupt face recognition to a greater degree than recognition of other types of visual stimuli (Vuong, Peissig, Harrison, & Tarr, 2005; Nederhouser, Yue, Mangini, & Biederman, 2007). The effect of contrast negation is easily seen (by those old enough to remember chemical photography) in the difficulty of recognising faces from photographic negatives. However, the negative of a normal colour image reverses both its luminance and its hue. A seminal study by Kemp, Pike, White, & Musselman (1996) demonstrated that face images with negated luminance but normal hue saw a reduction in recognition accuracy, whereas faces with negated hue and normal luminance showed no recognition accuracy decrease. It is therefore clear that information critical to face recognition is somehow carried through luminance values.

The explanation for why contrast reversal has such a profound effect on face recognition has been dominated by two different ideas. The first is Kemp et al's (1996) suggestion that the difficulty in recognising contrast negated faces is due to the reversal of 3D shape from shading cues. Their underlying premise is that disrupting depth cues through contrast reversal interferes with the use of information about 3D shape. However, more recent studies have tended to downplay the potential role of 3D information in face recognition and place more emphasis on the importance of surface pigmentation (Bruce & Young, 2012). For example Liu, Collin, & Chaudhuri (2000) found that recognition was poor for faces with intact 3D information but missing surface pigmentation. In line with such findings, Bruce & Langton (1994) developed the second main approach to the impact of contrast negation by suggesting that it exerts its disruptive influence because it reverses the brightness of important pigmented regions of the face, so that light regions of skin become dark, dark hair becomes light etc. Recent work expanding on this hypothesis has been carried out by Russell, Sinha, Biederman, & Nederhouser (2006), who propose that much of the contrast negation effect results from negating pigmentation cues that carry substantial information about facial identity. This has been elaborated by Sinha, Russell and their colleagues (Sinha, Balas, Ostrovsky, & Russell, 2006; Russell & Sinha, 2007) into a more general claim that because pigmentation carries much information about facial identity, any manipulation (e.g. contrast negation) that disrupts pigmentation rather than shape will have deleterious effects on recognition. Consistent with this hypothesis, further work by Santos & Young (2008) shows that contrast negation effects are not only confined to identity judgements but also impact on a range of social inferences made to faces in which pigmentation cues are likely to play a substantial role. Importantly, too, a study by White (2001) showed a degree of independence of contrast negation effects on facial identity and expression. White found that the perception of facial expression was much less affected by contrast negation than perception of identity. Since contrast negation does not affect the locations of regions of rapid change in luminance that define the shapes of facial features critical to the perception of expression, this dissociation is in line with the suggestion that contrast negation effects disrupt facial characteristics heavily dependent on pigmentation cues (i.e. identity) rather than shape (i.e. expression).

Recently, however, a dramatic finding by Gilad et al. (2009) has added to the debate about how and why contrast reversal disrupts recognition of identity by demonstrating that neither the pigmentation nor the shape from shading accounts given previously can be entirely correct. Gilad and colleagues (2009) created ‘contrast chimeras’ of famous individuals consisting of contrast negated faces in which the eye regions were left in positive contrast. Strikingly, recognition accuracy for these contrast chimeras was around 90% of the recognition rate for normal face photographs. This result is remarkable as it shows that merely restoring the correct contrast polarity to a small region around the eyes is sufficient to eliminate much of the classic contrast negation effect on face identification. Yet the major proportion of a contrast chimera image still has reversed luminance values and therefore gives incorrect shape from shading and incorrect pigmentation cues. Traditional approaches leave it unclear why restoring correct luminance values to only a small part of the image should have such a profound effect.

Gilad et al. (2009) suggested that this 'contrast chimera effect' arises because the eye region plays a critical role in creating a visual representation of a face and a contrast chimera face image maintains the natural ordinal contrast relationships around the eye region. More specifically, they point out that the ordinal relations between relatively lower luminance levels in the eye region and higher luminance values of the surrounding cheeks, nose and forehead remain constant across nearly all natural lighting conditions and viewpoints in which we encounter faces. Because these relationships are generally so stable, Gilad and colleagues suggested that the eyes provide a reliable feature around which to create face representations. Negating the contrast of the whole face reverses the natural ordinal contrast relationships between the eyes and surrounding face, disrupting the critical role of this region. Contrast chimeras however maintain the normal lower luminance levels of the eye region and higher luminance values of the surrounding face regions, therefore maintaining the correct ordinal contrast relationships from which an overall representation of the face can be created. This hypothesis offers a potentially important new perspective that regards ordinal contrast polarity relationships in the eye region as central to facial identity representations.

The aim of the present study is to use the chimeric effect to probe further how the eye region is involved in the perception of facial identity. The first experiment asked whether the chimeric benefit is specific to the eye region or whether it can be found with chimeras in which the forehead, nose and mouth region are returned to contrast positive. The second experiment investigated whether the eye region alone is sufficient or whether integration of information from the positive contrast eye region and the surrounding negative contrast face is essential for the chimeric benefit. In the third experiment, it was tested whether the chimeric benefit is dependent on a holistic representation of faces. In the final experiment, it was asked how precisely the information from the positive contrast eye region needs to match that from the contrast negated part of the face for the chimera benefit to occur.

6.3 Experiment 1

A first step towards understanding the contrast chimera effect is to identify the key facial region involved. Although Gilad et al. (2009) account focussed on the eyes, they noted that any beneficial effect of creating contrast chimeras based on the mouth region did not reach statistical significance, but did not report testing the effect of putting any other facial regions into photo positive. Because recent work by Ohayon, Freiwald, & Tsao (2012) has shown increased neural responses when macaques were exposed to non-eye regions from the face that had consistent ordinal contrast relationships (e.g. forehead-nose), Experiment 1 investigated whether contrast chimeras involving other regions of the face might also improve recognition compared to full contrast negated stimuli. Specifically, this experiment compared recognition rates for contrast chimeras involving the forehead, eyes, nose, or mouth regions. These regions were chosen in part based on the work of Dakin & Watt (2009) which suggested the nose and mouth regions as well as the eyes contain potentially informative horizontally oriented spatial frequencies. The forehead was chosen as a control region unlikely to contain critical information, so that it could be used as a baseline from which to measure whether simply revealing any region of the face in positive contrast might improve recognition accuracy.

In Gilad et al. (2009) study they created each of their chimeric stimuli by carefully blending the positive contrast eye region with the rest of the contrast negated image. This way of creating contrast chimeras makes it less obvious how the image has been changed and allows the region of change to be tightly circumscribed. However, because this experiment wanted to arrive at a manipulation that would be equivalent across different face regions, it was decided instead to use a simpler technique of using a constant-shaped rectangular frame to delineate the contrast positive region, with no blending of the borders of this rectangular frame or other artistic effects. This allowed experimenters to move the frame up or down within an image to bring into positive contrast equivalently-shaped regions highlighting the forehead, eyes, nose or mouth.

6.3.1 Method

6.3.1.1 Stimuli

Face images were produced using Adobe Photoshop version 5. Photographs of 32 (16 male, 16 female) famous and highly recognisable individuals (actors, politicians and other famous people) were used. Care was taken to ensure that each image was of a similar size, pose and image resolution, and that the faces were of people well-known to most undergraduate students in the UK (established by pilot screening). All images used were as close as possible to full face views. Images were greyscaled and cropped to only show the face (including hair), and superimposed on a uniform grey background. Faces were presented in a pseudo-random order in a 32 page booklet, with one face per page at the centre of an A4 sheet of paper. When viewed from around 30 cm away, each image subtended on average 13 degrees horizontally and 16 degrees vertically.

Each of the 32 greyscale face images was converted into a contrast negated (greyscale photo negative) version, and for each of these contrast negated images chimeric variants were produced involving by putting a constantly sized and shaped rectangular region around the forehead, eyes, nose or mouth into photo positive. Examples of these manipulations are shown in Figure 6.1.

6.3.1.2 Participants

Forty participants (mean age: 22 years, 20 females), took part in Experiment 1. All participants were from a western cultural background and had normal or corrected to normal vision. The study was approved and conducted following the guidelines of the Ethics committee at the University of York Psychology Department. Participants were paid or took part for course credit.

6.3.1.3 Design and Procedure

An independent measures design was used, with each participant allocated to one of five image manipulation conditions (8 participants per group). Each participant saw a pseudo-randomised sequence of faces in a booklet corresponding to their allotted image manipulation condition. They were asked to identify each face in an answer booklet. Correct recognition occurred when the full name of the face, an alias or an identifier unique to an individual (e.g. David Cameron could be identified as ‘Current prime minister of the UK’) were provided.

The five experimental conditions were as follows:

(i) A full contrast negative face.

(ii) Forehead chimera: negative contrast face except for a rectangular positive contrast forehead region. The positive contrast forehead region did not overlap the space covered in any other chimera conditions and was the same size as that used to create the eye chimera.

(iii) Eye chimera: each image seen by participants in this group was contrast negated except for a rectangular positive contrast eye region. The eye regions were demarcated to contain at least the whole eye and eyebrows of each individual.

(iv) Nose chimera: each image seen by participants in this group was contrast negated except for a positive contrast nose region. Again, the positive contrast nose region did not overlap the space covered in any other chimera conditions and was the same size as that used to create the eye chimera.

(v) Mouth chimera: each image seen by participants in this group was contrast negated except for a positive contrast mouth region. The positive mouth region did not overlap the space covered in any other chimera conditions and was the same size as the eye chimera.

After completing their designated image manipulation condition, all participants were given a booklet of the same 32 faces in normal greyscale image form presented in a new pseudo-random order, and were again asked to identify each face. This was used as an 'original image' control condition to determine which faces were known to each participant.

6.3.2 Results

Accuracy for each condition was calculated as a proportion of the faces recognised in the original image control condition (the normal greyscale images). In this way, faces that were not recognised in the control condition (i.e. faces that were unknown to participants) were eliminated from the analysis. A similar procedure was used by Gilad et al. (2009) and gives rise to recognition accuracy in the original image control condition of 100%.

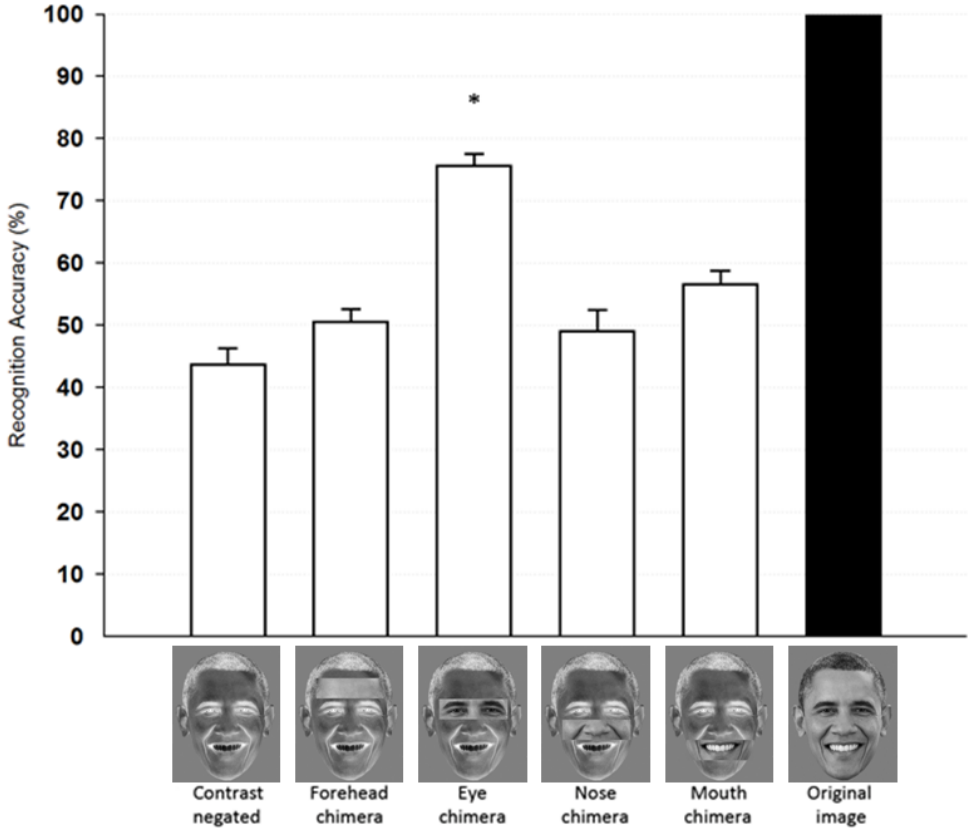

Figure 6.1: Histogram displaying recognition accuracy in each condition of Experiment 1, with standard error. Performance of the experimental conditions is conditionalised on recognition of the original image photographs. Asterisk marks condition with superior recognition rate in comparison to contrast negated images and all other chimera manipulations.

Percentage correct recognition accuracies in each condition are shown in Figure 6.1. Statistical analyses were conducted both by participants and by items.

For the analysis by participants, the percentage correct scores for the five image manipulation conditions were arcsine converted and subjected to a one way between groups ANOVA. Because of the use of an arcsine transform, mean accuracy results are given below in radians as well as the more easily interpreted percentages. The ANOVA indicated a significant effect of the image manipulation condition on recognition accuracy with a large effect size ($F_{4, 35}$ = 6.93, p<.001, n^2 =.44). A Tukey's HSD post hoc test was carried out with a corrected alpha value (a= .05, critical arcsine value = .25 radians) to compare performance between conditions. This showed that performance in the eye chimera condition (M= 75.7%, SD= 10.7, mean radians = .87) was significantly higher than for the contrast negated (M = 43.8%, SD= 14.2, radians = .46), forehead chimera (M = 50.6%, SD= 11.2, radians = .53), nose chimera (accuracy = 49%, SD= 19.9, radians =.53) and mouth chimera (accuracy = 56.6%, SD = 12.7, radians = .61) conditions. There were no significant differences between the four remaining conditions (contrast negated, forehead chimera, nose chimera, and mouth chimera).

To test whether the effects revealed from the analysis by participants generalised across stimulus items (i.e. across familiar faces), and to complement the limited power of the between participants design used, we also carried out an items analysis of our data. Percentage correct scores for each face in the five image manipulation conditions were arcsine converted and subjected to a one way between groups ANOVA. This items analysis revealed a pattern of results that matched the main analysis by participants. There was a significant effect of the image manipulation condition on recognition accuracy with a medium effect size ($F_{4, 155}$ =5.66, p<.001, n^2 = .14). A Tukey's HSD post hoc test (a= .05) showed that performance in the eye chimera condition was significantly higher than for the contrast negated, forehead chimera, nose chimera and mouth chimera conditions. There were no significant differences between the four remaining conditions (contrast negated, forehead chimera, nose chimera, and mouth chimera).

6.3.3 Discussion

The results of experiment 1 show that the chimeric benefit is maximal for the eye region. No significant improvements in face recognition were apparent for chimeras in which the forehead, nose or mouth region were in contrast positive compared to full negative contrast faces. That said, the overall recognition accuracy for eye region contrast chimeras (76%) did not seem as high as that in Gilad and colleagues (2009) study (91%). This may be due to differences in recognisability across the 32 famous faces we used compared to the 8 faces used by Gilad et al. (2009), or it might result from the less 'natural' rendering of the contrast positive region by simply 'letterboxing' it. This was done to make the area of the image returned to photo positive consistent across the chimeric conditions, but it introduces irrelevant lines into each image that define the border of the rectangular area, and adding irrelevant lines to an image of a face can affect its recognisability (Ellis, Davies, & Shepherd, 1978). The key point, though, is that in spite of this slightly reduced accuracy for our eye region contrast chimeras compared to that observed by Gilad et al. (2009), there was none the less a substantial benefit to recognition from returning the eye region of a contrast negated face back into contrast positive.

The finding that this contrast chimera benefit only existed in the eye chimera condition suggests that Gilad et al. (2009) correctly pinpointed the region of the face where ordinal contrast relationships appear to be most important in identity recognition. This is interesting as recent work by Dakin & Watt (2009) indicates that faces carry potentially useful high spatial frequency information in the horizontal plane especially around the eyes, nose and mouth. Consequently, creating chimeras with positive contrast in the nose and mouth regions, which contain the aforementioned horizontal high spatial frequency information, might also have improved recognition. Whilst inspection of Figure 6.1 does suggest that a study with greater statistical power might reveal some benefit for the mouth chimeras, any such effect is much less pronounced than for the eyes.

6.4 Experiment 2

Having established from Experiment 1 that the eye region offers the greatest recognition benefit for contrast chimeric images, Experiment 2 was carried out to test how the different parts of the eye chimeric image might be used by the observer in creating this contrast chimera benefit. In Gilad et al.'s (2009) paradigm, the contrast chimeric benefit is primarily measured by the difference in performance between a contrast chimera (the 'normal eye chimera' in Figure 6.2) and a standard greyscale negative (shown as 'contrast negated' in Figure 6.2). Logically, the only difference between these images is that the eye region of the normal eye chimera is contrast positive. Gilad et al.'s (2009) interpretation of the benefit of adding this contrast positive eye region was that it allowed information from the contrast negated region to be used to assist in identifying the individual. However, there is a variety of other possible reasons, of which the current experiment sought to test three:

- First, there might be enough information in the positive eye region to allow recognition based on using that region alone, and ignoring the rest of the chimeric stimulus. For this reason, an 'eyes only' condition was included.

- Second, the combination of a correct eye region might be used together with a fairly general cue from the negated region, such as overall shape. For this reason, an 'eyes on silhouette' condition in which the contrast positive eye region was placed on a correct face outline for the person. Gilad et al. (2009) had also used this condition, but as it is an important condition to replicate it was included here because the participants' overall performance in experiment 1 with the contrast chimeras seemed lower than that found by Gilad et al. (see discussion of Experiment 1).

- Third, the benefit might be less to do with the presence of a contrast positive eye region than with the removal of the misleading contrast negative eye region. For this reason, a 'blank eye negated' condition was included to see how well the rest of a contrast negated face could be recognised if the misleading eye region was removed.

6.4.1 Method

6.4.1.1 Stimuli

The same greyscale images of 32 famous individuals (16 male, 16 female) used in Experiment 1 were used to create stimuli for Experiment 2.

6.4.1.2 Participants

Forty new participants (mean age: 21 years, 20 females) participated in Experiment 2. All participants met the same demographic requirements as those in Experiment 1.

6.4.1.3 Design and procedure

An independent measures design was used with each participant allocated to one of five image manipulation conditions (8 participants per group):

(i) Contrast negated: each image seen by participants in this condition was a full contrast negated image was also included.

(ii) Positive eyes only: each image seen by participants in this condition was a positive contrast eye region, the same size and shape as that in the normal eye chimera condition, superimposed on a homogeneous grey background.

(iii) Positive eyes on face silhouette: each image seen by participants in this condition involved a positive contrast eye region of the same size and shape as that in the normal eye chimera condition, superimposed in the correct position on a head shape silhouette.

(iv) Blank eye negated face: each image seen by participants in this condition was contrast negated but with a homogeneous grey eye region that was the same size and the same average luminance as the eye region in the normal eye chimera condition.

(v) Normal eye chimera: each image seen by participants in this group was manipulated in the same way as those from the eye chimera condition in Experiment 1.

All other aspects of design and procedure were as for Experiment 1. As in Experiment 1 an 'original image' control condition was seen by all participants after their allocated image manipulation condition, to determine which faces each participant could recognise from a normal greyscale image.

6.4.2 Results

Percentage correct recognition accuracy scores were collated following the same method as in Experiment 1 and can be seen in Figure 6.2. For statistical analysis, the results were again analysed by participants and by items.

For the analysis by participants, percentage correct scores for the five image manipulation conditions were arcsine converted and subjected to a one way between groups ANOVA. This indicated a significant effect of the image manipulation condition on recognition accuracy with a large effect size ($F_{4,35}$ = 10.09, p<.001, n^2= .52). A Tukey's HSD post hoc test (corrected alpha value a= .05, critical Tukey's value = .26) was used to compare performance between conditions. Results indicated that recognition accuracy in the normal eye chimera condition (M= 72.4%, SD= 13.2, radians =.83) was significantly higher than contrast negated (M = 47.8%, SD= 9.7, radians = .50), positive eyes on silhouette (M = 36.4%, SD= 18.6, radians = .38), blank eye negated (M = 34.9%, SD =11.3, radians= .36) and eyes only (M = 41.9%, SD =18.1, radians = .44) conditions. Comparisons between the remaining conditions (contrast negated, positive eyes on silhouette, blank eye negated, and eyes only) showed no significant differences.

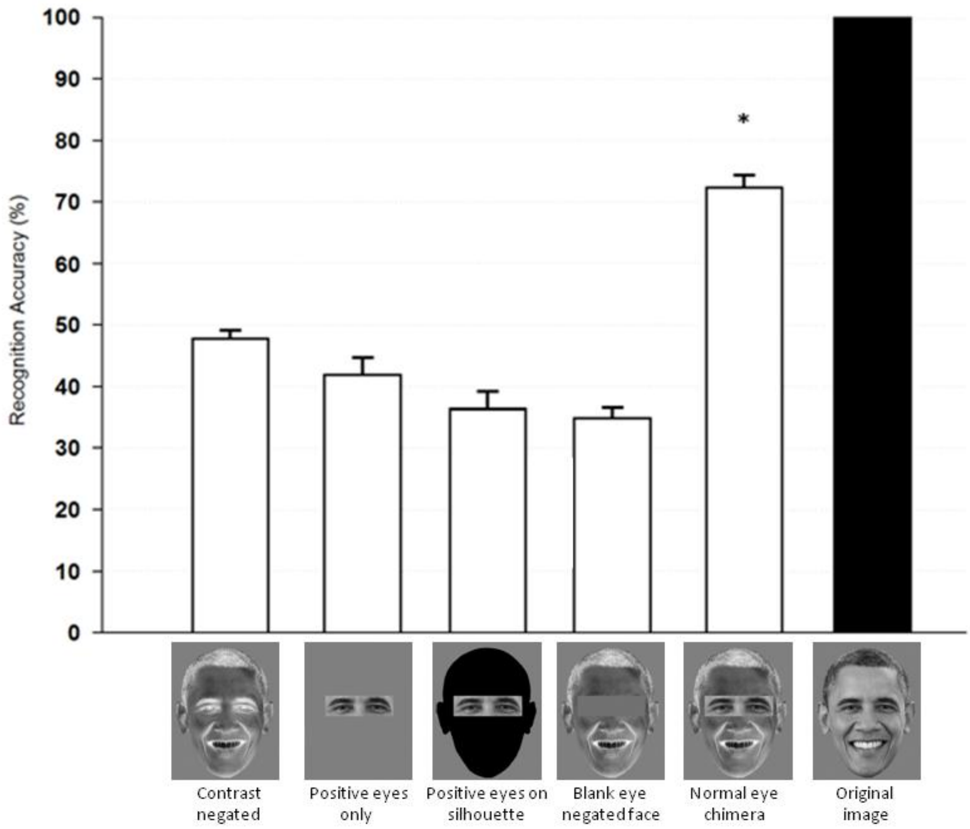

Figure 6.2: Histogram displaying recognition accuracy in each condition of Experiment 2, with standard error. Asterisk marks the condition with a higher recognition rate than all other experimental conditions.

As for Experiment 1, in Experiment 2 an items analysis was carried out on the data. The percentage correct scores for the five image manipulation conditions were arcsine converted and subjected to a between groups ANOVA. However, as the assumption of homogeneity of variance was violated, a Welch's F was used to adjust the degrees of freedom and the F ratio. This items analysis revealed a pattern of results that matched the main analysis by participants. Welch’s F indicated a significant effect of the image manipulation condition on recognition accuracy with a large effect size ($F_{4,\ 76.2} = 8.05$, $p<.001$, $w^2= .41$). A Games-Howell post hoc test was used to compare performance between conditions. Results indicated that recognition accuracy in the normal eye chimera condition was significantly higher than for contrast negated (p=.01), positive eyes on silhouette (p=.001), blank eye negated (p = .001) and eyes only (p=.001) conditions.

Comparisons between the remaining conditions (contrast negated, positive eyes on silhouette, blank eye negated, and eyes only) showed no significant differences (all p values greater than .05).

6.4.3 Discussion

The recognition accuracy data indicate that only the normal eye chimera condition led to a boost in performance, with all other image manipulations showing no difference from the baseline contrast negated condition. These findings are important as they show that specific components of the contrast chimeras cannot support good recognition rates alone. For example, the positive eye region alone is not particularly well recognised, and this is the case even when it is combined with the face's silhouette. Neither does the contrast negated region seem to contain useful cues in the absence of positive eyes. The benefit created by the positive eye region of the contrast chimeras therefore arises from some form of interaction between information from the positive eye region and the contrast negated remainder of the image. Experiment 3 investigated the nature of this interaction.

A minor point is that the recognition of 'positive eyes on face silhouette' did not differ significantly from the contrast negated condition, whereas Gilad et al. (2009) found that their equivalent condition was actually more detrimental to recognition than their contrast negated condition. This may reflect the fact that the technique used by the current experiment to create the contrast chimeras results in somewhat larger positive eye regions, therefore revealing more of the face and possibly giving a slightly stronger cue to identity from the eye region itself than the method used by Gilad and colleagues.

6.5 Experiment 3

Experiment 3 investigated how the positive eye region interacts with the contrast negated part of the image to boost recognition of contrast chimeras. Specifically, this experiment looked at whether the chimera benefit results from holistic perception of the face-like chimeric image. To achieve this, the current experiment tested whether it would be possible to create the contrast chimera effect by presenting the components of a contrast chimera in a configuration that was not face-like.

The concepts of 'configuration' or 'configural processing' are widely used in face perception research, but unfortunately can carry different meanings (Bruce & Young, 2012). Maurer, Grand, & Mondloch (2002) distinguish three importantly different senses in which such terms are often used. These are the first-order configuration of eyes above nose above mouth shared by all faces, the second-order configuration involving the precise spacings of these features unique to an individual, and thirdly the concept that the face is processed 'holistically' rather than as a collection of parts. For Experiment 3, a separate eyes and head chimera condition was created in which the eye region was placed alongside the rest of the face. This arrangement violates the first order configural properties as defined by Maurer et al. (2002), and violations of the first-order configuration (such as isolating or misaligning facial features) are known to severely disrupt holistic processing (Young, Hellawell, & Hay, 1987; Tanaka & Farah, 1993; Maurer et al., 2002; Bruce & Young, 2012).

We therefore applied the same criterion to the contrast chimera effect, asking whether it too would be disrupted by violating first-order configural properties by spatially separating the parts of the chimera. This was done by placing the positive eye region and outer negated face region of a chimera alongside each other (see Figure 6.3). Logically, these separate eye and head chimeras present the same information as a normal eye contrast chimera. Psychologically, though, the information is not in face-like form. If the contrast chimera effect involves the creation of a holistic percept, the recognition accuracy for a normal contrast chimera in a face-like first order configuration should be better than the recognition accuracy for its spatially separated components. If no difference in recognition is seen between these two conditions it would imply that the chimeric benefit results from a more ‘piecemeal’ processing of the constituent facial features (Carey & Diamond, 1977).

6.5.1 Method

6.5.1.1 Stimuli

The 32 normal eye chimera greyscale images from Experiments 1 and 2 were used, and spatially rearranged versions of each were created.

6.5.1.2 Participants

Sixteen new participants (Mean age = 21 years, 8 females) participated in Experiment 3. All participants fulfilled the same demographic requirements as those in Experiments 1 and 2.

6.5.1.3 Design and procedure

An independent measures design was used, with each participant allocated to one of two image manipulation conditions (8 participants per group):

(i) Separate eye and head chimera: each face seen by participants in this condition consisted of a contrast negated face with a homogeneous grey eye region presented adjacent to a positive contrast eye region from the same image. The distance between the eye region alone and the blank eye region was that of the distance from the ear to the edge of the positive eye region in the individual. Hence the stimuli comprised the same components as a standard contrast chimera, but placed alongside each other.

(ii) Normal eye chimera: the same stimuli as the ‘eye chimera’ condition in Experiment 1 and the ‘normal eye chimera’ condition in Experiment 2.

All other aspects of design and procedure were as for Experiments 1 and 2. Again an 'original image' control condition was seen by all participants after their allocated image manipulation condition, to determine which faces each participant could recognise from a normal greyscale image.

6.5.2 Results

Recognition accuracy was calculated using the same method as in Experiments 1 and 2 and can be seen in Figure 6.3. The percentage correct recognition accuracies in each condition were arcsine transformed for statistical analysis. For the analysis by participants an independent t test was used to probe the difference between the separate eyes and head chimera and normal eye chimera conditions. The t test showed that there was a significant difference in recognition accuracy between the normal eye chimera (M = 71.46%, SD = 13.3, radians = .95) and the separate eye and head chimera (M = 52.9%, SD = 11.7, radians =.60) conditions, $t\ (_{14}) = 2.81$, $p=0.014$, $r^2=.36$.

An items analysis using an independent t test revealed a pattern of results that matched the main analysis by participants. The t test showed that there was a significant difference in recognition accuracy between the normal eye chimera and the separate eye and head chimera conditions, $t\ (_{62}) = 2.61$, $p=0.011$, $r^2=.10$.

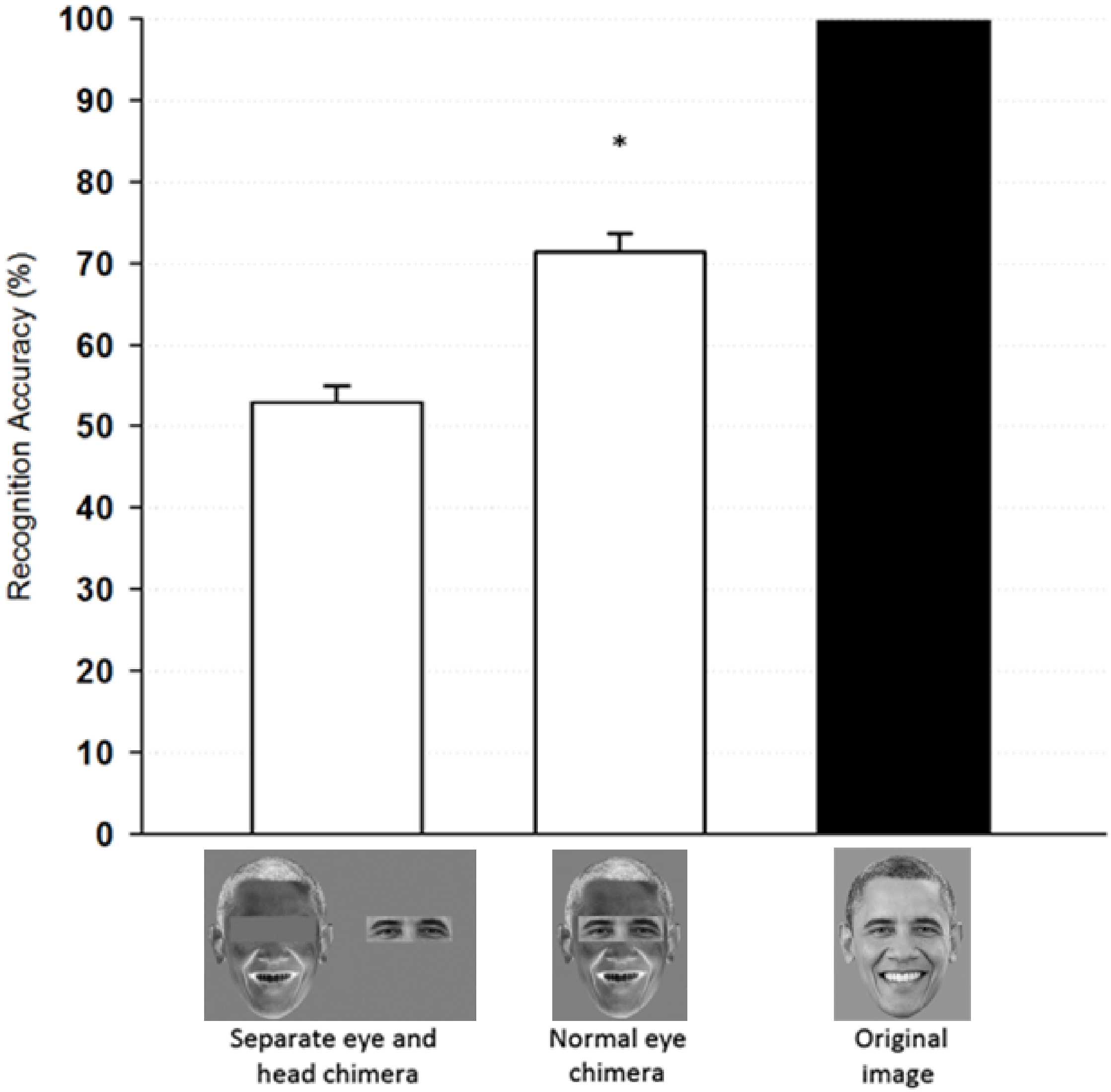

Figure 6.3: Histogram displaying recognition accuracy in each condition of Experiment 3, with standard error. Asterisk marks normal eye chimera showing a significantly higher recognition rate than the separate eye and head chimera.

6.5.3 Discussion

The contrast chimera benefit was greater when the positive contrast eye region was presented within a face-like configuration, demonstrating that it involves some form of holistic face processing. Having established that holistic processing of the contrast positive and contrast negated parts of the chimera takes place, we turned our attention to how precise this holistic processing might be, addressing Gilad et al. (2009) hypothesis of the importance of ordinal contrast relationships in the eye region.

6.6 Experiment 4

As Experiments 1-3 all pointed to the importance of interaction between information from the positive and negative parts of the chimera in the eye region, we sought to clarify how this happens by looking more closely at the role played by ordinal contrast relationships. As already noted, Gilad et al. (2009) account of the contrast chimera benefit is that the pattern of ordinal contrast relations around the eyes forms such a stable feature of the faces we see that the brain incorporates this stability into its perceptual representation of faces, leading to a severe decrement in recognition when the critical eye region is contrast negated. This account is couched in terms of a general property of faces, raising the question of whether any transform that can restore a normal set of ordinal contrast relationships to the eye region of a contrast negated face might make it more recognisable. Note that Gilad et al.'s (2009) account does not predict this would be the case - it only emphasises the stability of ordinal contrast relationships as inherent to the importance of the eye region. We sought to test how precisely the eye region must match the negated part of a contrast chimera for a benefit in recognition to accrue.

In a standard contrast chimera, as used here in Experiments 1-3, the chimera is created from a single face photograph. Because the positive eye region and the contrast negated part of the chimera both come from the same photograph, they necessarily match each other in two important ways - in identity (the parts come from images of the same person, even though one part is negated) and in the pattern of illumination (the parts come from images that are equivalently lit, even though one part is negated). In Experiment 4, we sought to unpack these influences as far as practicable, by including conditions that took the positive eye region of the chimera from a different photograph of the same person as that used for the contrast negated part (creating a difference in pattern of illumination between the parts of the chimera) or that took the positive eye region of the chimera from a photograph of a different person from that used for the contrast negated part (creating a difference in identity and in pattern of illumination between the parts of the chimera). In both of these conditions any overall ordinal contrast relations that are common to all face images will remain unchanged.

6.6.1 Method

6.6.1.1 Stimuli

The contrast negated and contrast chimera images of the 32 famous individuals used in Experiments 1-3 were used as the 'contrast negated' and 'normal eye chimera' conditions for Experiment 4.

Two additional chimera conditions were also created. First, using the positive eye regions from a second photograph of each of the faces of the 32 famous people in the stimulus set and combining these with the contrast negated images from the original set of photographs to create 'same eye identity but different image' chimeras in which the positive and negative parts came from different images of the same person. Second, by using the positive eye regions from photographs of the faces of 32 famous people who were not in the original stimulus set and combining these with the contrast negated images from the original set to create 'different eye identity' chimeras.

A pilot study was used to check that the two images of each face used to create the 'same eye identity but different image' chimeras were clearly perceptibly different in the eye region. In this pilot, 10 participants (Mean age years = 23, 5 females) completed a simple ‘same/different’ task where they were required to judge whether 2 sequentially presented pairs of eyes were from the same or different images. In each case the identity of the eyes remained unchanged across images; the only possible change was that the 2 images were or were not identical. Participants had to distinguish whether eye pairs were same or different image pairs for each of the 32 famous faces used. Results indicated that the eyes used to create the ‘same eye identity but different image' chimera stimuli could be easily distinguished from the ‘normal eye chimera’ eyes based on similar performance in both the same eye and different eye conditions (Different accuracy = 97%, SD=7.9; same accuracy = 96%, SD = 6.6).

6.6.1.2 Participants

Thirty-two participants (Mean age = 21 years, 16 females) participated in Experiment 4. All participants fulfilled the same demographic requirements as those in Experiments 1, 2 and 3.

6.6.1.3 Design and procedure

An independent measures design was used with each participant allocated to one of four image manipulation conditions (8 participants per group):

(i) Contrast negated: each image seen by participants in this condition was a full contrast negated image, as in Experiments 1 and 2.

(ii) Different eye identity (ID) chimera: each image seen by participants in this condition involved a contrast negated version of one of the individuals from the standard image set combined with a positive contrast eye region of the same size as that used for the normal eye chimera condition, but with this eye region belonging to one of four famous individuals (Orlando Bloom for Caucasian males, Sienna Miller for Caucasian females, Tyra Banks for Non- Caucasian females, Tiger Woods for non-Caucasian males) of the same sex and ethnicity.

(iii) Same eye identity (ID) different image chimera: each image seen by participants in this condition involved a contrast negated version of one of the individuals from the standard image set combined with a positive contrast eye region of the same size as that used for the normal eye chimera condition, but with this eye region taken from a different photograph of the same individual.

(iv) Normal eye chimera: the images seen by participants in this group were of the same individuals manipulated the same way as for the ‘eye chimera’ condition in Experiment 1 and the ‘normal eye chimera’ condition in Experiments 2 and 3.

All other aspects of design and procedure were as for Experiments 1 to 3, with an 'original image' control condition at the end of the session, to determine which of the 32 faces in the main set of stimuli could be recognised by each participant from a normal greyscale image.

6.6.2 Results

Recognition accuracies calculated using the same method as in Experiments 1-3 can be seen in Figure 6.4. Note that in the different eye ID chimera condition there are a variety of possible responses that might occur, based on identifying the image in terms of the positive eye region, in terms of the negated rest of the face, recognising it as an entirely unrelated person, or complete inability to recognise it as anyone familiar. In practice, the majority of responses involved either recognition based on the negated part (38%) or a complete inability to recognise it as anyone familiar (40%), whereas incorrect identification where an answer was produced (20%) and identification based on the eye region alone (2%) was relatively rare. In light of this, the current experiment allowed responses based on either the negated part of the image or on the positive eye region as 'correct'. This is a valid tactic because either of the eye or the negated face regions each contain one identity and therefore depending on which facial areas are used by the viewer for identification either answer can be viewed as based on information present in the stimulus.

For the analysis by participants, percentage correct recognition accuracies in each condition were arcsine transformed (results also stated in arcsine radians) and submitted to a one way between groups ANOVA. This indicated a significant effect of the image manipulation condition on recognition accuracy with a large effect size ($F_{4, 35}$ = 15.67,

$p<.001$, $n^2 = .63$). A Tukey's HSD post hoc test (corrected alpha value a= .05, critical Tukey's value =.28) was used to compare performance between conditions. This showed that the mean performance in the normal eye chimera condition (M= 81%, SD= 8.4, radians = .96) was significantly higher than for the contrast negated (M = 45.5%, SD= 11.1, radians = .48) and different eye ID chimera (M = 39.9%, SD= 11.1, radians = .44) conditions. However, performance in the normal eye chimera condition did not differ significantly from the same eye ID different image chimera condition (M = 71.7%, SD =16.8, radians =.84). Instead, the same eye ID different image chimera condition also showed significantly higher recognition accuracy than the different eye ID chimera and contrast negated conditions. The comparison between the different eye ID chimera and contrast negated conditions showed no significant difference.

A separate items analysis revealed a pattern of results that matched the main analysis by participants. There was a significant effect of the image manipulation condition on recognition accuracy with a medium effect size ($F_{3, 127} = 15.91$, $p<.001$, $n^2 = .28$). A Tukey's HSD post hoc test (corrected alpha value a= .05) showed that the mean score in the normal eye chimera condition was significantly higher than for the contrast negated and different eye ID chimera conditions. However, performance in the normal eye chimera condition did not differ significantly from the same eye ID different image chimera condition. Instead, the same eye ID different image chimera condition also showed significantly higher recognition accuracy than the different eye ID chimera and contrast negated conditions. The comparison between the different eye ID chimera and contrast negated conditions showed no significant difference.

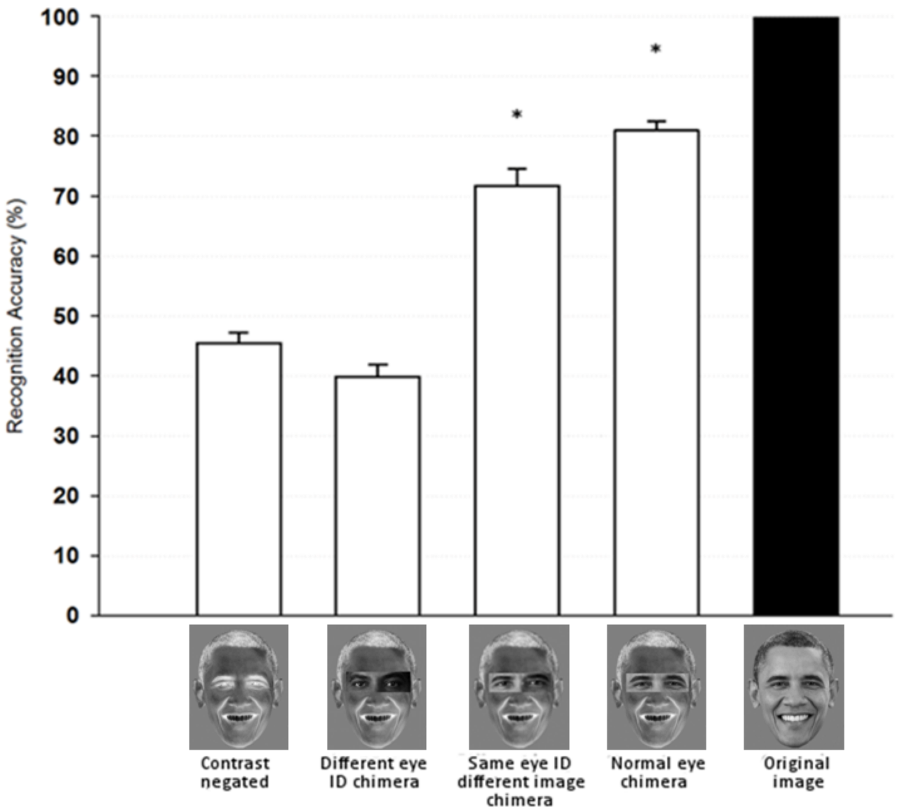

Figure 6.4: Histogram displaying recognition accuracy in each condition of Experiment 4, with standard error. Asterisks denote conditions with significantly greater recognition accuracy than contrast negated and different eye ID chimera conditions. Performance for the two asterisked conditions does not differ significantly from each other.

6.6.3 Discussion

The results showed that contrast chimeras created from two different images of the same face (same eye ID different image chimeras) were as effective in promoting recognition as chimeras made from parts of the same photograph (normal eye chimera condition). So the contrast chimera effect shows some degree of tolerance of lighting changes. However, putting someone else's eyes into a contrast chimera (the different eye ID chimera condition) nullified the chimeric benefit and brought performance back to the level found for fully negated images.

It is of course possible that an experiment with more power might reveal a statistically reliable difference between the 'same eye ID different image' chimera and the normal eye chimera conditions, but for present purposes this is not the main point, since both of these conditions led to a substantial boost in recognition compared to the different eye ID chimera condition. This clearly implies that identity information from the eye region is a critical component of the contrast chimera benefit. The mere presence of ordinal contrast relationships in the eye region that approximate those of any normal face (as in the different eye ID chimera condition) is not sufficient. If ordinal contrast relationships are the key determinant of the effect, they are sufficiently precisely specified as to be specific to the relevant identity. This suggests some level of interdependency of ordinal contrast relationships and shape cues in the eye region for representations of identity.

6.7 General Discussion

The aim of this study was to use the contrast chimera effect to explore the importance of the eye region for face recognition. The results clearly confirm Gilad et al. (2009) observation that restoring the correct contrast polarity to the eye region of a contrast negated face image confers a considerable benefit to recognition. However, the study's findings go further and demonstrate a number of important properties of this contrast chimera effect. First, the benefit to recognition of contrast negated face images from creating a contrast chimera containing a positive contrast eye region is a robust, easily replicable phenomenon. Second, the eye region does seem critical, since only contrast chimeras based around the eye region produced any statistically significant benefit for identification. Third, the contrast chimera benefit does not arise because the positive eyes are themselves intrinsically highly recognisable; it reflects a genuine interaction between the positive and negated regions of the chimera. Fourth, the contrast chimera effect reflects holistic processing of the image; the components are most effective in promoting recognition when arranged in a face-like configuration. Fifth, this holistic processing of contrast chimeras uses precise information from the eye region; there is no benefit from incorrect eyes that none the less preserve many of the ordinal contrast relationships typical of faces seen in everyday life.

These characteristics of the contrast chimera effect suggest that at least three interdependent properties contribute to the role of the eye region in face recognition. First, the eye region has important contrast polarity relations; Experiments 1-4 show how reversing these impacts on recognition (cf. Gilad et al., 2009). Second, information from the eye region is used in a precise fashion when these ordinal contrast relationships are correct (Experiment 4). Third, the eye region is combined holistically with other parts of the face (Experiment 3).

Three aspects of these findings merit more general discussion. These involve their implications for understanding normal face recognition mechanisms, why the eye region might be so important, and implications for the debate about whether faces can be considered a ‘special’ class of objects.

Considering the implications for understanding normal recognition mechanisms, this has of course been why the effects of contrast negation have been considered so intriguing. The paradox of contrast negation is that, strictly speaking, all of the information is still present in the negated image, yet the visual system seems unable to use it. The striking benefit from putting a relatively small region around the eyes back into photo positive therefore has important implications, particularly since (as we have demonstrated here) the benefit accrues from the participant's now being able to use information from the contrast negated region itself. As noted in the Introduction, this does not sit easily with accounts of the impact of contrast negation based only on the reversal of shape from shading or pigmentation cues, since these accounts offer no explanation as to why information from the negated region should become usable under the particular condition that the image has a positive eye region. This doesn't eliminate any role for such cues, but at the least it implies they are only part of the story.

Interestingly, previous studies have shown that contrast negation does not itself preclude holistic face processing (Hole, George, & Dunsmore, 1999; Taubert & Alais, 2011). Thus the finding that the eye region contrast chimera is processed holistically with the contrast negated part of the image must fit into this more general background. Further studies might therefore usefully seek to clarify how the positive eye region facilitates the usefulness of information in the negated part of the chimera, which seems to us the enigma at the core of the contrast chimera effect.

One way to approach this enigma is to think more carefully about the information in a greyscale face photograph. The image involves a complex pattern of light and shade resulting from how light falls across the 3D structure of the face and the surface pigmentation of the facial features, skin and hair. Some of this pattern is potentially informative about the face's identity, but much of it also results from identity-irrelevant lighting variations and camera characteristics. Burton, Jenkins, Hancock, & White (2005) have shown how averaging a number of different images of the same person's face can eliminate most of this irrelevant variation, leading to an averaged pigmentation pattern with good recognisability. Moreover, Burton et al. (2005) noted that these averaged pigmentation patterns are sufficiently characteristic of the face in question that most of them remain recognisable even if they are themselves morphed into an average face shape. This shows that pigmentation patterns contribute sufficiently strongly to the perception of face identity that they can override 2D shape cues involving feature positioning (the 'second-order configuration').

A promising direction for future work on understanding the contrast chimera effect therefore seems to us to explore what aspects of this pigmentation pattern can be used from the contrast reversed part of the image. For instance, whether the information accessed from the negated region involves relatively coarse or fine spatial frequencies, and whether or not it is 2D-shape dependent. Such approaches could also be used to ask whether the benefit of a correctly rendered eye region is specific to overcoming the deleterious effect of contrast negation per se, or whether it applies equally in any circumstances that create problems for face recognition. For example the addition of visual noise or some other degradation of the image might be used to reduce recognition of a face to a level comparable to that created by contrast negation, and the benefit of restoring the eye region could then be determined. A direct comparison between the benefit of restoring the eyes in a contrast chimera and in other types of degraded image could help to establish the extent to which the contrast chimera effect reflects the more general diagnostic value of the eye region in face recognition.

Why should the eye region have such privileged status? Various lines of evidence converge on the importance of the eye region in the perception and recognition of faces (McKelvie, 1976; Haig, 1986; O'Donnell & Bruce, 2001; Barton, Radcliffe, Cherkasova, Edelman, & Intriligator, 2006; Peterson & Eckstein, 2012), but the contrast chimera benefit makes the point particularly effectively. Prior to the discovery of the chimera effect evidence already existed showing a substantial effect of eye inversion on face recognition (Leder, Candrian, Huber, & Bruce, 2001; Rakover & Teucher, 1997). Similarly, the role of holistic processing in the contrast chimera benefit matches findings showing that prosopagnosic individuals who do not process faces holistically fail to fixate on and extract diagnostic information from the eyes when trying to recognise faces (Caldara et al., 2005; Orban de Xivry, Ramon, Lefèvre, & Rossion, 2008).

There are a number of factors that may be involved in this pre-eminence for the eyes. Structurally the positive contrast eye region contains a number of characteristic pigmentation variations in a small area; these include the sclera, iris and pupil within the eyeball, the border between the nose and the brow, and the eyebrows, with the latter being implicated as also contributing to familiar face recognition (Sadr, Jarudi, & Sinha, 2003). These structural properties are what leads to the presence of substantial information from horizontal spatial frequencies in the eye region described by Dakin & Watt (2009), and they create the stable ordinal contrast relationships noted by Gilad et al. (2009) which potentially offer a useful anchor point around which to construct facial identity representations. Moreover, the eyes are of course also important to a wide range of social signals (Bruce & Young, 2012), so human perceivers have many reasons for looking at this region of the face. Studies in evolutionary biology suggest that the importance of signals communicated through eye gaze is such that the human eye may have evolved a distinct pigmentation pattern that can communicate gaze direction very effectively (Kobayashi & Kohshima, 1997).

Although the exact mechanism that produces the contrast chimera effect is unclear, there are hints from experiments carried out in the current study that hint at which regions are crucial. Firstly, experiment one demonstrates that having the correct unmanipulated surface information is important in the eye region alone and in turn experiment 2 shows that this positive contrast eye region must be presented in the context of correct facial shape information provided by facial features surrounding the eye region. Taken together these results suggest that this surface based information in the eye region is incorporated into an facial identity percept in the context of the surrounding face regions (although these only need to primarily provide the correct shape or edge information), possibly because it maintains the correct ordinal contrast relationships described by Gilad et al (2009) and earlier Sinha (2002). Experiment 4 shows that these ordinal contrast relationships are not image specific but are contingent on the correct shape based identity information also being present in the eye regions. So the ability of a viewer to accurately identify a contrast chimeric face is likely to depend on a mechanism that incorporates the correct shape and surface information in the eye region with the shape information from surrounding facial features. Moreover experiment 3 demonstrates that the eyes and facial features must be presented in their proper first order configural relationships for the effect to arise.

From this it is possible to cautiously draw at least two main inferences about how the contrast chimera effect reflects naturalistic face processing. Firstly, the contrast chimeric effect suggests that the richest identity information is contained in the eye region and expressed through both shape and surface information and through ordinal contrast relationships with surrounding face regions. Secondly, the contrast chimera effect reflects naturalistic face processing by requiring facial features to be in approximately correct first order configuration for the viewer to be able to extract the correct facial identity, presenting the eyes without surrounding feature context isn't sufficient for the viewer to reliably extract the facial identity. Taken together these findings suggest that it possible that maintaining precise shape and surface information in a highly informative face region (in this case the eye region) and the correct feature configuration in a face allows viewers to reconstruct facial identities in images that are someway visually impoverished or manipulated. Therefore the contrast chimera effect may highlight the key rules a face image must adhere to so that a viewer is still able to accurately extract facial identity information.

A point that should not be neglected, though, is that the importance of the eye region isn't the only thing that needs to be explained. Both our results and those of Gilad et al. (2009) show that the recognition rate for eye region contrast chimeras is none the less lower than that for a full positive face. So we cannot totally discount the usefulness of other face areas and in fact (Sinha, 2002) suggested that important relative contrast relationships may also exist in other face regions (e.g. nose-cheek) as well as the eyes. A possibility for future studies would be to test whether chimeric stimuli involving combinations of regions provide significantly better recognition than a single region chimera alone.

Finally, the findings have implications for the debate about the status of faces as 'special' stimuli for the visual system. This debate has been bedevilled by the different meanings and interpretations given to the word 'special' (Hay & Young, 1982; Ellis & Young, 1989), but the dependence of the contrast chimera benefit on the eye region of the face and the fact that few stimuli other than faces have eyes gives it a clearly unusual status that qualifies as 'special' within most of the definitions that have been used. Indeed a number of studies have shown that in the perception of faces is actually unlike many other classes of object in that it is differentially severely impaired by contrast negation (e.g. Subramanian & Biederman, 1997; Nederhouser et al., 2007).

Overall, the findings of the current study provide a robust replication of the contrast chimera effect first found by Gilad et al. (2009) and support the importance of contrast relationships in the eye region to face recognition. When correct contrast relationships are present, information from the eye region is used in a precise but holistic fashion.

Chapter 7 - General Discussion

7.1 Summary of research questions and findings

The aim of the current thesis was to investigate the extent to which facial shape and surface properties represent facial expression and identity information using a combination of fMRI, behavioural and model fitting methods. The roles of shape and surface properties have been previously studied with regard to facial expression in terms of both PCA-based measures of image similarities between expressions (Andrew J. Calder et al., 2001) and manipulations of surface-based properties through contrast negation (White, 2001; Harris, Young, & Andrews, 2014) or line drawings (e.g. Etcoff & Magee, 1992; Kirita & Endo, 1995; Magnussen, Sunde, & Dyrnes, 1994). Results from these studies show that facial shape cues are important to the perception of facial expression. But there has until now been a lack of evidence to suggest that human observers are able to extract expression information from surface cues alone. The roles of shape and surface properties in facial identity perception have also been investigated in a number of studies that either disrupt facial surface cues (e.g. Bruce & Langton, 1994, Kemp, Pike, White, & Musselman, 1996) or aim to separate shape and surface cues to identity (e.g. Russell, Sinha, Biederman, & Nederhouser, 2006; Russell & Sinha, 2007). However, the neural basis of these behavioural findings has not been fully established.

In Chapter 3, the relative role of shape and surface cues in the processing of facial expression was addressed using an expression categorisation task on images that varied in either shape or surface cues. Chapter 4 used a combination of a perceptual similarity task, image property model fitting and fMRI MVPA to test the extent to which expression cues based on facial shape or surface properties reflect the similarity in the perceived expression similarity and linked this expression similarity to neural representations of expressions in face selective regions of the brain. Chapter 5 moved the research question about the roles of facial shape and surface into the domain of facial identity. Using images created using the same averaging methods used by Burton et al (2005) to convey either unique shape or surface cues to facial identity this question was addressed with an fMRI MVPA paradigm. Finally, Chapter 6 examined the roles of shape and surface information in facial identity perception using the contrast chimera effect (Gilad et al, 2009).

The empirical work presented in the current thesis has two main implications for the understanding of facial expression and identity perception. Firstly, the findings suggest a revised role for facial surface information in the perception and neural representation of facial expression. Secondly, they provide a better understanding of how facial identity is represented in the brain and what previous studies using contrast negation actually reveal about image cues to facial identity. The following sections provide a brief summary of the results from all four empirical chapters. Subsequently, the findings are discussed in the broader context of current research on face processing. The discussion concludes with a summary of the main findings and their significance.

7.1.1 Chapter 3 - Both shape and surface cues are important to the categorisation of facial expressions

Theoretical accounts of face processing often emphasise feature shapes as the primary visual cue to the recognition of facial expressions (Bruce & Young, 2012). However, changes in facial expression also affect the surface properties of the face (Calder et al, 2001). Chapter 3 investigated whether this surface information could also be used in the recognition of facial expression. In Experiment 1 results indicated that participants were able to identify 5 facial expressions (fear, anger, disgust, sadness, happiness) from the Ekman and Friesen (1976) face set when the images that had been manipulated such that they varied primarily in either shape or surface properties. It was found that the accurate categorization of facial expression is possible based on varying shape or surface properties, but that different expressions are relatively dependent on surface or shape properties.

Experiment 2 then tested the extent to which the categorisation of facial expressions could be disrupted by contrast negation. Experiment 2 contained contrast negated versions of the original images, shape varying and surface varying images used in Experiment 1. Results showed that experimental conditions where facial surface varied (e.g. surface varying or original images) showed significantly decreased recognition accuracy when contrast negated compared to their positive contrast counterparts. In addition, shape varying images showed a significant decrease in recognition accuracy when contrast negated likely reflecting that either these images did contain some surface varying expression information or that shape cues can be impaired to some extent by contrast negation. This finding shows any image which contains surface varying expression information (that Experiment 1 showed can be used to categorise facial expression) is impaired when contrast negated as the contrast negation disrupts these surface varying cues to facial expression.

Finally, Chapter 3 introduced a new method to understand the relative contributions of shape and surface information to the categorization of facial expressions. Stimuli created by combining the surface properties of one expression with the shape properties from a different expression showed that the categorization of facial expressions in these hybrid images was equally dependent on the surface and shape properties. Together, these findings provided a direct demonstration that both feature shape and surface information make significant contributions to the recognition of facial expressions.

The finding that expressions can be accurately categorised based on facial surface alone is in line with a recent study by Benton (2009) who used a behavioural adaptation paradigm and showed that the adaptation after effect to contrast negated expressions was significantly smaller that to normal contrast expression images. This suggests that facial surface information is part of the visual representation of facial expressions as well as edge based information. In the context of previous work by Calder et al (2001) who used

PCA to show that there exists useful surface based information conveying facial expression, Chapter 3 provides the first clear demonstration that surface varying expression information can be used by human observers to categorise expressions. With this in mind future models of expression perception should take into account that surface based information can be useful in expression categorisation alongside shape information.

7.1.2 Chapter 4 - Perceived similarity of facial expressions is based on image shape and surface similarities

The aim of Chapter 4 was to understand the extent to which the perception and neural representation of facial expressions is based on shape and surface information. A recent study by Said et al (2010) has shown that neural populations in the STS may contain a perceptual similarity-based representation of facial expressions. However, prior to Chapter 4 it was not clear to what extent representations of perceptual similarity are based on different sources of visual image similarities between different expressions. To answer this question Chapter 4 measured the perceptual similarity between different facial expressions in a behavioural task and correlated this with objective image-based models that could represent either facial shape or surface similarity.

The results in Chapter 4 showed that the perceptual similarity of different facial expressions (fear, anger, disgust, sadness, happiness) can be predicted by models representing both surface and feature shape similarities of facial expressions. Following this, a block design fMRI MVPA study was employed to test whether the perceptual similarity of expressions could also be predicted from the patterns of neural response in face-selective brain regions. Results showed that neural response patterns in the posterior superior temporal sulcus (STS) and occipital face area (OFA) but not in the fusiform face area (FFA) could significantly predict perceptual similarity response patterns. This shows that patterns of neural response in face responsive brain regions thought to be involved in facial expression perception (STS and OFA) do, to some extent reflect the perceptual similarity of facial expressions. The findings in Chapter 4 thus demonstrate that the perceptual similarity of facial expressions reflects the shape and surface properties of the image and correlates with the activity of specific face-selective regions.

The findings of Chapter 4 build on those in Chapter 3 by showing that both facial shape and surface contribute to the representation of facial expression similarity. These findings provide an advance on previous work (e.g. Said et al, 2010) in two main ways. Firstly, by showing that perceptual similarity judgements of facial expressions are based on both facial shape and surface similarities. Secondly, by showing that patterns of neural response to facial expression can be discriminated in the OFA as well as in the STS. These results are also consistent with a recent study showing that patterns of neural representation in the STS but not the FFA are able to discriminate different facial expressions (Zhang et al., 2016).

7.1.3 Chapter 5 - A shape based contribution to the neural representation of facial identity

A range of behavioural studies have shown the importance of surface compared to shape information for the recognition of facial identity (Burton, Jenkins, Hancock, & White, 2005; Russell et al., 2006; Russell & Sinha, 2007). The aim of Chapter 5 was to test whether face responsive brain regions contained shape and surface information based representations of familiar faces. Images used in Chapter 5 were created using the same averaging methods described by Burton et al (2005) with faces that had the shape properties from one identity and the surface properties from a different identity were employed in this experiment. A block design fMRI experiment combined with a MVPA was used to test the sensitivity of face-responsive brain regions to the shape or surface properties of each face identity. Based on previous behavioural evidence it was predicted that face responsive brain regions sensitive to facial identity (i.e. the FFA and OFA) would show consistent patterns of response (i.e. consistent neural response patterns that could discriminate within category facial identity over between category facial identity) to face blocks that had surface based cue to the identity, but not show consistent neural response patterns to face blocks that had the same shape properties. Results did not fit this prediction and instead indicated that within category identity neural response patterns in the FFA and OFA were discriminable from between category identity responses to facial identities based on facial shape properties but not facial surface properties. This finding provides the first multi voxel patterns analysis based evidence that consistent neural response patterns can be found to facial identities on the basis facial shape information.

The fact that face responsive brain regions contained consistent neural responses to facial identity based on facial shape properties is consistent with previous suggestions that apparent category selectivity within visual cortex (such as that for faces in the FFA) may emerge from neural activation in visually mapped regions containing lower level image based representations (e.g. shape, surface or contrast) (Op de Beeck, Haushofer, & Kanwisher, 2008; Andrews, Watson, Rice, & Hartley, 2015). Recent findings from Rice et al. (2014) have shown that category selective neural responses in object selective regions of visual cortex arose from the low level image properties of each object category, suggesting that the ventral visual stream may be organised around representations of the image properties of what are usually considered higher level stimuli (such as faces).

7.1.4 Chapter 6 – Contrast negation and the importance of the eye region for holistic representations of facial identity

Chapter 6 sought to employ the recently discovered contrast chimera effect (Gilad et al, 2009) to further understand the process of facial identity recognition. As outlined in the introduction studies of contrast negation have provided clear evidence that this manipulation dramatically reduced facial identity recognition. Competing accounts of have been given to explain why contrast negation causes such a large decrement in facial identity recognition. Bruce & Langton (1994) suggest the contrast chimera effect arises from the disruption of pigmentation patterns crucial to facial identity whilst Kemp, Pike, White, & Musselman (1996) suggesting that contrast negation disrupts crucial shape from shading cues to facial identity. Chapter 6 aimed to understand why the contrast chimera managed to support good rates of facial identity recognition and what this effects means for the aforementioned theories of why contrast negation disrupts facial identity recognition.

In Chapter 6 a series of experiments were carried out aiming to test the causes of the contrast chimera effect (Gilad et al, 2009). The first experiment successfully replicated the contrast chimera effect whilst showing that it is specific to the eye region and no benefit is seen when returning the nose, mouth or forehead regions to positive contrast. The second experiment showed that presenting a positive contrast eye region alone was not enough to produce a good rate of identity recognition. This demonstrates that the contrast chimera benefit does not rely purely on the identity information in the positive eye region. A further experiment probed this and showed that when presenting a contrast negated outer face region separately alongside a positive eye region, there was no recognition benefit despite of the fact that the image contains all of the feature parts present in a contrast chimera. This shows that much like the processing of a normal face, the perception of a contrast chimeric face is holistic, in the sense that the parts of the image must be presented in a face-like configuration for the positive eye region and negated rest of the face to be useful. The final experiment aimed to address the main claim of Gilad et al., namely that ordinal contrast relationships in the eye region are important in representations of facial identity. To do this, contrast chimeras were created with incorrect eye identity but correct ordinal contrast relationships with surrounding regions. This condition showed no recognition benefit, suggesting that the contrast chimera benefit is as reliant on the correct shape information in the eye region as much as it depends on correct ordinal contrast relationships in and around the eye region.

Importantly, Chapter 6 found that neither the positive contrast eye region nor the contrast negated outer face region support a high recognition rate on their own, yet they can be used together to produce identity recognition rates close to those for normal face photographs. The findings of Chapter 6 provided a series of informative tests of which specific properties are important in conveying facial identity and which of these are met by a contrast chimeric face.

7.2 The role of shape and surface properties in the representation of facial expression

The findings of Chapter 3 are different from a number of previous contrast negation based expression perception studies which reported that contrast negation or surface cue disruption does not substantially impair perception of expressions (White, 2001; Pallett & Meng, 2013; Harris et al., 2014a). This had been taken to mean that the shape (edge based) information that remains intact once a face image has been contrast negated is sufficient to accurately identify an expression to nearly the same level as when the face image is not manipulated.

Whilst the findings of Chapter 3 are in line with the suggestion that shape information alone can be sufficient to support good rates of expression recognition (e.g. Chapter 3 shows that expressive faces varying in shape only can support high rates of recognition), the results show that equivalent image manipulations that leave varying surface information as the main cue to facial expression can support equally high recognition accuracy. Therefore previous assertions that shape information is more useful than surface information for expression perception made by White (2001) and others more recently must be revised. In fact results in Chapter 3 Experiment 2 indicated that negated facial expression images were categorised significantly less accurately than the original non-manipulated facial expression images (albeit not to the extent seen with facial identity judgements). This suggests that whatever degradation of surface based information is caused by contrast negation is sufficient to significantly impair accurate expression perception and qualifies previous literature showing no effect of contrast negation on facial expression perception (e.g. White, 2001; Pallett & Meng, 2013; Harris et al., 2014a).

This inconsistency may be due to the use of different tasks between Chapter 3 in this thesis and that of the aforementioned past studies. Whereas previous studies investigating the effects of contrast negation on facial expression perception typically employed an expression pair matching paradigm and used only two or three differing expressions (e.g. White, 2001; Pallett & Meng, 2013; Harris et al., 2014a), the experiments in Chapter 3 employed a 5 alternative forced choice categorisation, which is a task that requires more accurate identification of each expression (theoretically an observer may be able to match 2 expressions having miscategorised them both). Thus the use of a stricter test in Chapter 3 of this thesis appears have revealed that contrast negation does impair facial expression perception, albeit to a lesser extent than it impairs facial identity recognition. These results are supported by a recent finding suggesting that contrast negation of facial expressions may hinder their perception. Specifically, Benton (2009) found that participants who viewed contrast negated faces showed significantly reduced adaptation after effects compared to normal contrast expression images. This result, taken together with the findings of Chapter 3, suggests that there exists some degree of surface based representation of facial expressions.

Subsequently Chapter 4 investigated whether the perceptual similarity between expression pairs is based on the physical image properties of facial expressions. Previous studies such as Said, Moore, Engell, & Haxby (2010) have compared the perceptual similarity ratings of facial expressions to neural response patterns in STS, finding that they were significantly correlated. Chapter 4 extended these findings in two main ways. Firstly, since Chapter 3 showed that both shape and surface information could be used by observers to identify facial expressions, Chapter 4 extended the question to investigate whether perceptual similarity ratings of expressions might also be based on facial shape and surface similarity. Secondly, since previous studies focused on measuring neural response patterns to facial expressions only in the STS, Chapter 4 aimed to expand the measurement of neural response to facial expression across all core face responsive brain regions.

The results in Chapter 4 demonstrated three main findings. Firstly, patterns of facial expression perceptual similarity could be predicted by both facial shape and surface similarity models of facial expression. Secondly, both the OFA and STS contained consistent patterns of neural response to each facial expression whilst the FFA did not. Finally, these ratings of perceptual similarity of facial expressions were significantly correlated to neural response patterns to facial expressions in OFA and STS but not in FFA.

The fact that patterns of facial expression perceptual similarity were reflected by both facial shape and surface based image properties shows that the graded dimensional (i.e. continuous) representations of facial expression previously outlined by Woodworth & Schlosberg (1954) and Russell (1980) can be conveyed by similarities in physical image properties. This result echoes the findings from Chapter 3 which had earlier shown that categorical representations of facial expression are likely to be based on both facial shape and surface information.

The Chapter 4 finding that it is possible to discriminate between different patterns of neural response to different expressions in STS replicates a number of previous studies that were also able to discriminate between different facial expressions based on neural response patterns in STS (Said, et al., 2010a, Said, Moore, Norman, Haxby, & Todorov, 2010 , Zhang et al., 2016). Additionally both Said et al 2010 studies showed that patterns of neural response in STS correlate to patterns of perceptual similarity ratings. Furthermore, Chapter 4 provided another novel result by showing that patterns of neural response to facial expression in the OFA are consistent and that these neural responses correlate to perceptual similarity of expressions much like they had in the STS.

Although the Chapter 4 finding that the FFA could not discriminate between facial expressions is in line with predictions made by the Haxby et al. (2001) model's suggestion that the FFA is not involved in facial expression perception, further discussion is merited due to recent inconsistencies in the neuroimaging literature. In this context the results in Chapter 4 differed to studies by Harry, Williams, Davis, & Kim (2013), Skerry & Saxe (2014) and Wegrzyn et al. (2015) who found that the FFA could decode positive and negative facial expressions. More recent results were consistent with the findings in Chapter 4 ,for example Zhang et al. (2016) also showed that there were no consistent neural patterns of response to facial expression in the FFA.

It is difficult to know with certainty why MVPA studies inconsistently find that the FFA is able to decode facial expressions across different studies, but there are at least two plausible explanations. Firstly, the two of the aforementioned studies (Harry et al, 2013 and Skerry & Saxe, 2014) that found an FFA ability to decode facial expressions included experimental conditions with non face stimuli and the neural response to these conditions was included in the global normalisation of neural response patterns. Practically this means that the mean that was subtracted from the neural response would be lower in face selective regions, possibly therefore removing less of the neural response specific to facial characteristics. Therefore it is possible that some of the neural response patterns in FFA that appear to be evoked by expression actually included neural responses to other facial characteristics (e.g. identity, gender, etc.). This may be reflected in the authors admitting that although expressions could be decoded significantly above chance in the FFA effect sizes were small (e.g. Harry et al, 2013). This possibility may explain the difference in findings as experiments where normalisation was based only on neural response to faces (e.g. Chapter 4 and Zhang et al, 2016) show no decoding of facial expressions in FFA.

7.3 The role of shape and surface properties in the representation of facial identity

7.3.1 A shape based contribution to the neural representation of facial identity

Having found compelling evidence to suggest that facial shape and surface properties were of approximately equal utility and salience in facial expression perception, chapters 5 and 6 asked to what extent shape and surface information were important in neural and perceptual representations of facial identity. Chapter 5 utilised familiar face stimuli in which either shape or surface information had been independently manipulated to allow either to be the only varying identity cue. Multi voxel pattern analysis showed that consistent neural responses to facial identity could be found in FFA and OFA but not STS on the basis of face shape but not face surface properties. This finding contrasts with a number of previous behavioural studies which showed that surface properties play a preeminent role in determining facial identity (Bruce & Langton, 1994; Vuong, Peissig, Harrison, & Tarr, 2005; Burton et al., 2005; Russell, Sinha, Biederman, & Nederhouser, 2006. The results in Chapter 5 provide novel MVPA based evidence that consistent neural responses to facial identity can be found in core face responsive regions based on facial shape information.

Interestingly some aspects of the results in Chapter 5 can explain previous findings about the role facial features play in representations of facial identity. For example, it is clear in figure 5.1 that one of the main cues to facial identity in the shape invariant conditions is the consistency of the positioning and shape of facial features. Consequently, it is possible that the consistent neural response patterns in OFA and FFA to shape cues to identity are based on these invariant facial characteristics. This fits with previous work by Yovel & Kanwisher (2004) and Maurer et al. (2007) who compared neural activation in face responsive regions to conditions, where either facial feature positioning or feature shape was informative about facial identity. Both studies found that the FFA showed approximately equal sensitivity for facial feature or feature positioning relationships. However Maurer and colleagues also found that there was an area in fusiform gyrus (not overlapping with the FFA, likely the OFA) that showed increased sensitivity to feature positioning. Leading on from this it is possible that the ability of the OFA to discriminate between facial identities (as shown in Chapter 5) is based on its sensitivity to facial feature shapes and their positioning.

Similarly the finding in Chapter 5 that the FFA contained consistent neural response to facial identity on the basis of shape information. Previous studies attempting to use patterns of neural response in the FFA to discriminate between different identities have had mixed results, with some finding discriminable patterns of response to facial identity (e.g. Liu, Harris, & Kanwisher, 2010; Natu et al., 2010; Nestor, Plaut, & Behrmann,2011; Zhang et al., 2016) and others not (Kriegeskorte, Formisano, Sorger, & Goebel, 2007; Nestor, Vettel, & Tarr, 2013; Nestor, Plaut, & Behrmann, 2016). However, none of these studies systematically measured changes in, or manipulated physical image features that convey facial identity.

The absence of a discriminable MVPA neural pattern based on surface information does not necessarily mean that these patterns do not exist at some (perhaps more fine grained) scale. Rather, it shows only that the attempts in Chapter 5 to discriminate neural response patterns on the basis of surface information did not succeed. There is evidence from univariate fMRI studies to suggest that both the OFA and FFA respond to facial shape and surface cues to facial identity (F Jiang et al., 2009). This evidence does not necessarily conflict with the findings of Chapter 5; rather it answers a different question. This is because univariate fMRI identifies neural populations that are responsive to some aspect of a stimulus on the basis of significant activation in individual voxels, whereas MVPA tests large scale neural response patterns across distributed regions for consistency and whether the stimuli that evoke those responses can be discriminated from each other based on these patterns. As such the findings in Chapter 5 do not mean that there is no neural response to surface based representations of facial identity, but rather that facial identities may be discriminated based on facial shape information.

Although Chapter 5 makes a promising start by showing that it is possible to find consistent neural responses to facial identity in the FFA and OFA based on their response to facial shape cues, there are a few interesting avenues to explore leading from this result. In future it may be interesting to carry out an fMR-adaptation experiment (K Grill-Spector & Malach, 2001) using the Burton, Jenkins, Hancock, & White (2005) stimulus set. Such a design has been employed successfully in the past to help disentangle two covarying facial cues, notably this was recently done elegantly by Harris, Young, & Andrews (2012) to help disentangle the extent to which different face responsive regions were involved in expression or identity processing.

A potential design could mirror the one used in Chapter 4 and involve an fMRI MVPA experiment using familiar faces that vary in shape, surface or both facial shape and surface information. In addition to this, behavioural data based models of perceptual similarity and identity categorisation and image property models based on shape and surface similarity could be developed. Rather than testing purely whether there is enough shape or surface based neural response in face responsive regions to discriminate facial identity on the basis of either cue, this would answer a different question to the one asked in Chapter 5. Specifically, it would probe whether there are neural responses in different face responsive regions that respond to facial identity correlate to the perceptual similarity or categorisation of facial identity and in turn whether this is based on shape and surface cues that convey facial identity. This would also be significantly different to the adaptation study carried out by Jiang et al (2006) and discussed earlier, as they used unfamiliar, computer generated faces which may have weaker neural representations and represent less of the natural variation in face images. This might resolve the discrepancy between previous behavioural evidence showing that facial surface information is critical to facial identity perception and the findings in Chapter 5 showing that some neural representations of facial identity are shape based.

7.3.2 The contrast chimera effect suggests an important role for the eye region in representations of facial identity

The dramatic effect of contrast negation on facial identity recognition is a classic phenomenon that has been widely replicated. Long-standing theories have debated whether contrast negation exerts its influence through disruption of 3D shape from shading cues (Kemp, Pike, White, & Musselman, 1996) or because it reverses regular patterns of pigmentation that convey facial identity (Bruce & Langton, 1994). The recently discovered contrast chimera effect Gilad, Meng, & Sinha (2009) investigated in Chapter 6 has highlighted the limitations of previous explanations about how contrast reversal disrupts face recognition as it violates aspects of both types of explanation whilst producing high rates of identity recognition. The contrast chimera effect is a particularly interesting case as it identifies the eye region as critical to the creation of an overall representation that can be used to recognise facial identity.

Whilst discussion of how the contrast chimera effect arises is interesting, it is also important to consider what the contrast chimera effect discovered by Gilad and colleagues (2009) and replicated here in Chapter 6 means for previous explanations for the role of facial shape and surface properties in the perception of facial identity. One explanation for the contrast negation effect offered by Bruce & Langton (1994) suggested that contrast negation reduced facial identity recognition because this manipulation reverses the brightness of important pigmented regions of the face. The contrast chimera effect does not completely rule out this explanation as it is possible that these important pigmented regions are actually in the eye region. However, the fact that the majority of the image is contrast negated in a contrast chimera does cause problems for Bruce and Langton's suggestions that other pigmented regions of the face (not in the eye region) contain crucial identity information. As if it is true that important identity information is contained in these pigmented non-eye regions, it may not necessarily be 'crucial' to the percept of facial identity. The significance of the findings in Chapter 6 is that it expands on this original explanation by confirming that it is reversing the brightness relationships between the eyes that primarily causes the reduction in facial identification accuracy.

With regards to Kemp et al's (1996) alternative explanation that the difficulty in recognising contrast negated faces is due to the reversal of 3D shape from shading cues, the contrast chimera effect clearly shows the need for a different explanation. Kemp et al's underlying premise was that disrupting depth cues through contrast reversal interferes with the use of information about 3D shape. Whilst it is true that contrast negation is likely to disrupt 3D shape from shading cues across the whole face, the contrast chimera effect shows that it is possible to restore a high level of identification accuracy by only returning a small region of the face (the eyes) to positive contrast. This suggests one of two possible things; either that 3D shape from shading information is mainly a critical cue to facial identity in the eye region relative to other face regions, or that restoring positive contrast to the eye regions restores other important cues as well as 3D shape from shading that are important for facial identity recognition. The latter argument may seem to be more likely as Liu, Collin, & Chaudhuri (2000) showed that facial identity recognition was poor for faces with intact 3D information but missing facial surface information.

A property of the human eye that may be important in both the percept of facial identity and the contrast chimera effect is the large amount of white sclera around the darkened iris, which has been found to be unique to the human species (Kobayashi & Kohshima, 1997). Kobayashi & Kohshima (1997) suggest that this gives the human eye an ability to communicate direction of gaze without movement of the head. Since being able to read gaze direction and non-verbally convey intention is an important facet of human social behaviour, it seems logical that the perceptual system would choose the eye regions as a good area to centre attention on faces. Prior to the discovery of the contrast chimera effect evidence already existed showing a substantial effect of eye inversion on facial identity recognition (Leder, Candrian, Huber, & Bruce, 2001; Rakover & Teucher, 1997).

The findings in Chapter 6 provided a successful and clear replication of the contrast chimera effect originally discovered by Gilad and colleagues (2009) and indicate that the maintained percept of facial identity in a contrast chimeric image requires both precise identity-related information and the preservation of correct ordinal contrast relationships in the eye region. The importance of ordinal contrast relationships in the eye region had been earlier demonstrated by Tomalski, Csibra, & Johnson (2009) who found that gaze cueing to schematic faces in normal contrast was quicker than cueing to 'contrast reversed' schematic faces (ones in which eyes were brighter than the rest of the face). While this may seem intuitive, a further manipulation revealed that restoring correct contrast polarity in the eye regions of a contrast negated face (creating the illusion of a contrast chimera) restored normal eye region gaze cueing. Other past experiments focusing on the mechanism by which we use facial information for recognition of identity suggest human observers may be drawn to the eye region due to its consistency across different visual properties (e.g. contrast, high level of spatial frequency information, colouration) (e.g. Sinha, 2002; Gilad et al., 2009).

As well as maintaining vital ordinal contrast relationships around the eye region the positive contrast eye region in a contrast chimera also contains accurate pigmentation patterns, 3D shape from shading information (Gilad, Meng, & Sinha, 2009), and a high frequency shape information that has been shown to be important to the perception of identity (Dakin & Watt, 2009). Therefore simply maintaining a positive contrast eye region (as is done in a contrast chimera) maintains a number of different converging physical cues to facial identity contained in the eye region. Consequently it may help to understand the findings in Chapter 6 in terms of the contrast chimera effect arising through an interaction of information in the positive eye region and information from the negated parts of the chimera, rather than from either alone.

The frequent utilisation of identity information contained in the eye region by human observers has been demonstrated in a recent study by Peterson & Eckstein (2012). Peterson & Eckstein (2012) analysed face images to identify the most informative face regions in terms of how much unique identity information was conveyed in each image pixel. Pixels around the eyes were shown to convey a large amount of unique identity information, with some identity information also conveyed by the nose and mouth regions. In their experiment participants' early fixation gaze paths to each face showed a fixation preference for a region in the centre of the eyes near the bridge of the nose. Petersen & Eckstein also created a 'foveated ideal observer'; this involved modelling which regions of the face image show the most variation in shape and surface information whilst accounting for the fact that as eccentricity increases (in the visual field) visual acuity decreases. Their findings showed that the ideal fixation point for optimal integration of the largest amount of varying facial identity information (in high acuity) across the visual field lies between the centre of the eyes and bridge of the nose. Interestingly this is exactly the region in which human observers' early fixations fell in the identity judgement task in their study. Taken in combination with the findings of Chapter 6, this study provides a useful context for understanding why the eye region plays such an important role in facial identity recognition.

Whilst manipulation of contrast and luminance offers one way to investigate the roles of shape and surface in facial identity, other types of image degradations that disrupt either cue have been shown to impair recognition of faces such as blurring (Gilad-Gutnick, Yovel, & Sinha, 2012) and removal of orientation-specific spatial frequency information (Ruiz-Soler & Beltran, 2006; Dakin & Watt, 2009). One interesting future direction would involve creating chimeric faces by using one of these manipulation techniques on the outer face region whilst keeping a normal, positive contrast eye region. If a blurred or orientation-filtered chimera supported good rates of identification it would suggest that the interaction between the information contained in the eyes does not need to be combined with exact high spatial frequency shape information from other facial features.

7.4 Are face responsive regions organised by social function or by visual cues? What image properties reveal about the Haxby model

7.4.1 Evidence for the organisation of face responsive brain regions by social function

Haxby's (2001) model of face processing suggests functional subdivisions for different face responsive regions (e.g. STS processes variant features, FFA processes invariant features). This information must be conveyed through information in the face image. This thesis has shown that different face regions depend differentially on shape or surface properties.

The findings in Chapter 4 showed that in the OFA and STS (but not the FFA) neural response patterns could discriminate between different facial expressions and in turn these patterns correlated to perceptual similarity ratings of expressions (which had earlier been shown to correlate to physical image characteristics). In Chapter 5, there were consistent patterns of neural response to facial identity based on facial shape characteristics in the OFA and FFA but not the STS. Taken together these findings do suggest that in early visual face responsive regions (i.e. the OFA) information common to both facial expression and identity is processed. But results also show that patterns of neural response could only discriminate between expressions in the STS, whilst patterns of neural response could only discriminate between identities in the FFA.

7.4.2 Evidence for the organisation of face responsive brain regions by image properties

The link between image properties and patterns of neural response suggests that the neural representations in face regions may be more related to image properties than higher-level representations of expression or identity. This fits with a review by Op de Beeck et al. (2008) who suggested that areas appearing to respond to stimulus categories (e.g. selective FFA response to faces) may emerge from neural activation in nearby visually mapped regions containing lower level image based representations of the face stimulus. It is suggested that topographical organisations of lower level image maps overlap to form what appear to be discrete category selective modules. Recent work by Rice et al (2014) found that category selectivity in higher visual cortex arose from the low level image properties of each category as evidenced by the correlation of lower level image properties to neural response patterns in these regions; this included neural response to face stimuli in face responsive brain regions. A recent review by Andrews et al. (2015) also provided evidence drawn across a number of studied suggesting that there is a link between the organisation of category selective regions in ventral visual cortex and the image properties of each type of image.

This approach of understanding the perception of higher level objects on the basis of their mid-level (cf. Marr, 1982) characteristics provides a useful perspective to understanding the organisation of face responsive regions within the Haxby model of face processing. Faces arguably do occupy a special position in the visual perception sphere for a number of reasons. Whilst being relatively uniform in their characteristics they can simultaneously convey a wide range of social (e.g. emotion, mood) and non-social (e.g. race, identity) signals that are often useful and relevant to the observer. This means that any approach that seeks to parse the ways different visual characteristics affect different judgements about faces provides a useful framework for understanding the underlying process by which that judgement occurs. Indeed recent approaches related to the one taken in this thesis successfully elucidated which underlying image features convey more high level social cues such as attractiveness (Menzel, Hayn-Leichsenring, Langner, Wiese, & Redies, 2015) and dominance (Vernon et al., 2014).

Interestingly, however, the MVPA results across chapters 4 and 5 showing that the OFA can discriminate between different facial expressions and identities (although identity mainly on shape based information) highlight the limitations of an image based approach. Based on these result it may be unclear why there are supposed facial identity specific (FFA) and facial expression specific (STS) regions beyond the OFA. It is important to keep in mind that whilst an image statistic based approach provides a useful way to understand the organising principles of face responsive brain regions, there is strong evidence that extended regions within the distributed face network include the STS which contains cross modal representations of expressions of emotion (Peelen, Atkinson, & Vuilleumier, 2010), more frontal semantic regions such as the inferior anterior temporal lobe that contain amodal representations of facial identity (Anzellotti & Caramazza, 2015; Zhang et al., 2016). Conversely there is little evidence to suggest that the OFA contains representations of facial identity or expression abstracted from their visual properties, however, based on the OFA's proximity lower level visual regions suggests it's representations of facial identity and expression are likely to be heavily dependent on image statistics. As such the utility of extended regions of the Haxby model of face perception (.e.g. STS,FFA, ATL) is likely to lie in their role in representing higher level, conceptual information about facial expressions and identities that are more abstracted from the physical characteristics of the face image. Therefore although future or current models of face perception should account for lower level organising principles in areas such as the OFA, FFA and STS, authors should exercise restraint about how much of the wider face processing network can be explained purely on the basis of face based visual characteristics.

7.5. Conclusions

This thesis aimed to investigate the extent to which facial shape and surface properties play a role in representations of facial expression and identity. Results in Chapters 3 showed that facial expression categorisation can be carried using both facial shape and surface properties of the face image, rather than only shape information, as had been previously assumed. Chapter 4 combined a perceptual similarity rating task, image model regression and fMRI multi voxel pattern analysis methods to demonstrate that the perceptual similarity of expressions correlated to their shape and surface based similarity. Furthermore it was found that the OFA as well as the STS could discriminate between different expressions of emotion. These chapters produced novel findings by consistently showing that human observers can use surface based information in both categorisation of expressions and similarity judgements between expression pairs. Chapter 5 revealed that there are consistent patterns of neural response to facial identity based on facial shape information. In response to faces that varied in either facial shape or surface both the FFA and OFA could discriminate between facial identities but only for shape varying rather than surface varying familiar face stimuli. Although this was the first attempt at measuring the ability of face responsive regions to discriminate between identities based on image properties alone, the result was striking as previous behavioural evidence suggests that surface characteristics form a more salient cue in facial identity perception. Finally, results in Chapter 6 showed that information in the eye region is crucial to maintaining the percept of facial identity.

Overall the fMRI based chapters within this thesis show the potential of Haxby et al.'s conception of facial expression and identity processing pathways that overlap at the OFA, an area that can discriminate both facial identities and expressions. These pathways separate later, with the FFA primarily concerned with discriminating facial identity and the STS primarily involved in discriminating facial expression. Taken together behavioural results showed that facial shape and surface information both play important roles in the perception of facial expression, but that they interact particularly closely in the eye region to support the recognition of facial identity.

Bibliography

Akiyama, T., Kato, M., Muramatsu, T., Saito, F., Umeda, S., & Kashima, H. (2006). Gaze but not arrows: A dissociative impairment after right superior temporal gyrus damage. *Neuropsychologia*, *44*(10), 1804–1810.

Allison, T., Puce, A., & McCarthy, G. (2000). Social perception from visual cues: role of the STS region. *Trends in Cognitive Sciences*, *4*(7), 267–278.

Ambadar, Z., Schooler, J. W., & Conn, J. F. (2005). Deciphering the enigmatic face: the importance of facial dynamics in interpreting subtle facial expressions. *Psychological Science*, *16*(5), 403–410.

Andrews, T. J., & Ewbank, M. P. (2004). Distinct representations for facial identity and changeable aspects of faces in the human temporal lobe. *NeuroImage*, *23*(3), 905–913.

Andrews, T. J., & Schluppeck, D. (2004). Neural responses to Mooney images reveal a modular representation of faces in human visual cortex. *NeuroImage*, *21*(1), 91–98.

Andrews, T. J., Schluppeck, D., Homfray, D., Matthews, P., & Blakemore, C. (2002). Activity in the fusiform gyrus predicts conscious perception of Rubin's vase-face illusion. *NeuroImage*, *17*(2), 890–901.

Andrews, T. J., Watson, D. M., Rice, G. E., & Hartley, T. (2015). Low-level properties of natural images predict topographic patterns of neural response in the ventral visual pathway visual pathway. *Journal of Vision*, *15*(2015), 1–12.

Anzellotti, S., & Caramazza, A. (2014). The neural mechanisms for the recognition of face identity in humans. *Frontiers in Psychology*, *5*(JUN), 1–6.

Anzellotti, S., & Caramazza, A. (2015). From Parts to Identity: Invariance and Sensitivity of Face Representations to Different Face Halves. *Cerebral Cortex (New York, N.Y. : 1991)*, *35*(25), 9252–9254.

Arcurio, L. R., Gold, J. M., & James, T. W. (2012). The response of face-selective cortex with single face parts and part combinations. *Neuropsychologia*, *50*(10), 2454–2459.

Bar, M. (2004). Visual objects in context. *Nature Reviews. Neuroscience*, *5*(8), 617–629.

Barton, J. J. S., Press, D. Z., Keenan, J. P., & O'Connor, M. (2002). Lesions of the fusiform face area impair perception of facial configuration in prosopagnosia. *Neurology*, *58*(1), 71–8.

Barton, J. J. S., Radcliffe, N., Cherkasova, M. V, Edelman, J., & Intriligator, J. M. (2006). Information processing during face recognition: The effects of familiarity, inversion, and morphing on scanning fixations. *Perception*, *35*(8), 1089–1105.

Baseler, H. A., Harris, R. J., Young, A. W., & Andrews, T. J. (2013). Neural Responses to Expression and Gaze in the Posterior Superior Temporal Sulcus Interact with Facial Identity. *Cerebral Cortex (New York, N.Y. : 1991)*, *24*(3), 737–744.

Baylis, G. C., Rolls, E. T., & Leonard, C. M. (1985). Selectivity between faces in the responses of a population of neurons in the cortex in the superior temporal sulcus of the monkey. *Brain Research*, *342*, 91–102.

Beauchamp, M. S., Nath, A. R., & Pasalar, S. (2010). fMRI-Guided Transcranial Magnetic Stimulation Reveals That the Superior Temporal Sulcus Is a Cortical Locus of the McGurk Effect. *Journal of Neuroscience*, *30*(7), 2414–2417.

Bentin, S., Allison, T., Puce, A., Perez, E., & McCarthy, G. (1996). Electrophysiological Studies of Face Perception in Humans. *Journal of Cognitive Neuroscience*, *8*(6), 551–565.

Benton, C. P. (2009). Effect of photographic negation on face expression aftereffects. *Perception*, *38*(9), 1267–1274.

Berman, M. G., Park, J., Gonzalez, R., Polk, T. a., Gehrke, A., Knaffla, S., & Jonides, J. (2010). Evaluating functional localizers: The case of the FFA. *NeuroImage*, *50*(1), 56–71.

Betts, L. R., & Wilson, H. R. (2010). Heterogeneous structure in face-selective human occipito-temporal cortex. *Journal of Cognitive Neuroscience*, *22*(10), 2276–88.

Bötzel, K., Schulze, S., & Stodieck, S. R. (1995). Scalp topography and analysis of intracranial sources of face-evoked potentials. *Experimental Brain Research. Experimentelle Hirnforschung. Experimentation Cerebrale*, *104*, 135–143.

Bruce, V. (1982). Changing faces: visual and non-visual coding processes in face recognition. *British Journal of Psychology*, *73*(1), 105–116.

Bruce, V. (1994). Stability from variation: the case of face recognition. *The Quarterly Journal of Experimental Psychology. A*, *47*(1), 5–28.

Bruce, V., & Langton, S. (1994). The use of pigmentation and shading information in recognising the sex and identities of faces. *Perception*, *23*(7), 803–22.

Bruce, V., & Young, A. (1986). Understanding face recognition. *British Journal of Psychology*, *77*(3), 305–327.

Bruce, V., & Young, A. (1998). *In the eye of the beholder : the science of face perception*. Oxford University Press.

Bruce, V., & Young, A. (2012). *Face perception*. Psychology Press.

Burt, D. M., & Perrett, D. I. (1995). Perception of age in adult Caucasian male faces: computer graphic manipulation of shape and colour information. *Proceedings. Biological Sciences / The Royal Society*, *259*(1355), 137–143.

Burton, A. M. (2013). Why has research in face recognition progressed so slowly? The importance of variability. *Quarterly Journal of Experimental Psychology (2006)*, *66*(8), 1467–85.

Burton, A. M., Bruce, V., & Hancock, P. (1999). From Pixels to People: A Model of Familiar Face Recognition. *Cognitive Science*, *23*(1), 1–31.

Burton, A. M., Bruce, V., & Hancock, P. J. B. (1990). From pixels to people: A model of familar face recognition. *Cognitive Science*, *23*(1), 1–31.

Burton, A. M., Jenkins, R., Hancock, P. J. B., & White, D. (2005). Robust representations for face recognition: the power of averages. *Cognitive Psychology*, *51*(3), 256–84.

Butler, A., Oruc, I., Fox, C. J., & Barton, J. J. S. S. (2008). Factors contributing to the adaptation aftereffects of facial expression. *Brain Research*, *1191*, 116–126.

Caldara, R., Schyns, P., Mayer, E., Smith, M. L., Gosselin, F., & Rossion, B. (2005). Does prosopagnosia take the eyes out of face representations? Evidence for a defect in representing diagnostic facial information following brain damage. *Journal of Cognitive Neuroscience*, *17*, 1652–1666.

Calder, A. J., Burton, A. M., Miller, P., Young, A. W., & Akamatsu, S. (2001). A principal component analysis of facial expressions. *Vision Research*, *41*(9), 1179–1208.

Calder, A. J., & Young, A. W. (2005). Understanding the recognition of facial identity and facial expression. *Nature Reviews. Neuroscience*, *6*(8), 641–651.

Calder, A. J., Young, A. W., Keane, J., & Dean, M. (2000). Configural information in facial expression perception. *Journal of Experimental Psychology. Human Perception and Performance*, *26*(2), 527–551.

Calder, A. J., Young, A. W., Perrett, D. I., Etcoff, N. L., & Rowland, D. (1996). Catgorical perceptions of morphed facial expressions. *Visual Cognition*, *3*(2), 81–117.

Calvert, G. A., Hansen, P. C., Iversen, S. D., & Brammer, M. J. (2001). Detection of audio-visual integration sites in humans by application of electrophysiological criteria to the BOLD effect. *Neuroimage*, *14*(2), 427–38.

Carey, S., & Diamond, R. (1977, January 21). From piecemeal to configurational representation of faces. *Science (New York, N.Y.)*.

Carlson, T. A., Schrater, P., & He, S. (2003). Patterns of activity in the categorical representations of objects. *Journal of Cognitive Neuroscience*, *15*(5), 704–717.

Clifford, C. W., & Rhodes, G. (2005). Fitting the Mind to the World: Adaptation and Aftereffects in High-Level Vision. In *Advances in Visual Cognition* (pp. 213–240).

Cox, D. D., & Savoy, R. L. (2003). Functional magnetic resonance imaging (fMRI) “brain reading”: Detecting and classifying distributed patterns of fMRI activity in human visual cortex. *NeuroImage*, *19*(2), 261–270.

Dakin, S., & Watt, R. (2009). Biological “bar codes” in human faces. *Journal of Vision*, *9*, 1–

10.

Davies, G., Ellis, H. D., & Shepherd, J. (1978). Face recognition accuracy as a function of mode of representation. *Journal of Applied Psychology*, *63*(2), 180–187.

Davies-Thompson, J., & Andrews, T. J. (2012). Intra- and interhemispheric connectivity between face-selective regions in the human brain. *Journal of Neurophysiology*, *108*(11), 3087–95.

Davies-Thompson, J., Gouws, A., & Andrews, T. J. (2009). An image-dependent representation of familiar and unfamiliar faces in the human ventral stream. *Neuropsychologia*, *47*(6), 1627–1635.

Dricot, L., Sorger, B., Schiltz, C., Goebel, R., & Rossion, B. (2008). The roles of “face” and “non-face” areas during individual face perception: Evidence by fMRI adaptation in a brain-damaged prosopagnosic patient. *NeuroImage*, *40*(1), 318–332.

Du, S., Tao, Y., & Martinez, A. M. (2014). Compound facial expressions of emotion. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(15), E1454–62.

Ekman, P. (1972). Universals and cultural differences in facial expressions of emotion. *Nebraska Symposium on Motivation*, *19*, 207–283.

Ekman, P., & Friesen, W. V. (1978). Facial Action Coding System: A Technique for the Measurement of Facial Movement. In *Consulting Psychologists Press*.

Ekman, P., & Friesen, W. V. (1976). *Pictures of Facial Affect*. *Palo Alto, CA: Consulting Psychologists Press*.

Ellis, H. D., Davies, G. M., & Shepherd, J. W. (1978). Remembering Pictures of Real and Unreal Faces - Some Practical and Theoretical Considerations. *British Journal of Psychology*, *69*(4), 467–474.

Ellis, H. D., & Young, A. W. (1989). *“Are faces special?” In A. W. Young & H. D. Ellis (Eds.), Handbook of research on face processing (pp. 1-26). Amsterdam: North Holland.*

Engell, A. D., & Haxby, J. V. (2007). Facial expression and gaze-direction in human superior temporal sulcus. *Neuropsychologia*, *45*(14), 3234–3241.

Etcoff, N. L., & Magee, J. J. (1992). Categorical perception of facial expressions. *Cognition*, *44*(3), 227–240.

Etzel, J. A., Zacks, J. M., & Braver, T. S. (2013). Searchlight analysis: Promise, pitfalls, and potential. *NeuroImage*, *78*, 261–269.

Flack, T. R., Andrews, T. J., Hymers, M., Al-Mosaiwi, M., Marsden, S. P., Strachan, J. W. A., … Young, A. W. (2015). Responses in the right posterior superior temporal sulcus show a feature-based response to facial expression. *Cortex*, *69*, 14–23.

Flack, T. R., Watson, D. M., Harris, R. J., Hymers, M., Gouws, A., Young, A. W., & Andrews, T. J. (2014). Distinct Representations for Rigid and Non-Rigid Facial Movements in Face-Selective Regions of the Human Brain. *Journal of Vision*, *14*(10), 1383–1383.

Fox, C. J., & Barton, J. J. S. (2007). What is adapted in face adaptation? The neural representations of expression in the human visual system. *Brain Research*, *1127*, 80 – 89.

Fox, C. J., Moon, S. Y., Iaria, G., & Barton, J. J. S. (2009). The correlates of subjective perception of identity and expression in the face network: An fMRI adaptation study. *NeuroImage*, *44*(2), 569–580.

Fox, C. J., Oruç, I., & Barton, J. J. S. (2008). It doesn't matter how you feel. The facial identity aftereffect is invariant to changes in facial expression. *Journal of Vision*, *8*(3), 11.1–13.

Freiwald, W. A., & Tsao, D. Y. (2010). Functional compartmentalization and viewpoint generalization within the macaque face-processing system. *Science (New York, N.Y.)*, *330*(6005), 845–51.

Friston, K. J., Holmes, A. P., Poline, J. B., Grasby, P. J., Williams, S. C., Frackowiak, R. S., & Turner, R. (1995). Analysis of fMRI time-series revisited. *NeuroImage*, *2*(1), 45–53.

Galper, R. E. (1970). Recognition of faces in photographic negative. *Psychonomic Science*, *19*(4), 207–208.

Galper, R., & Hochberg, J. (1971). Recognition memory for photographs of faces. *The American Journal of Psychology*, *84*(3), 351–354.

Ganel, T., Goshen-Gottstein, Y., & Ganel, T. (2004). Effects of familiarity on the perceptual integrality of the identity and expression of faces: the parallel-route hypothesis revisited. *Journal of Experimental Psychology. Human Perception and Performance*, *30*(3), 583–597.

Ganel, T., Valyear, K. F., Goshen-Gottstein, Y., & Goodale, M. A. (2005). The involvement of the "fusiform face area" in processing facial expression. *Neuropsychologia*, *43*(11), 1645–1654.

George, N., Dolan, R. J., Fink, G. R., Baylis, G. C., Russell, C., & Driver, J. (1999). Contrast polarity and face recognition in the human fusiform gyrus. *Nature Neuroscience*, *2*(6), 574–80.

Gilad, S., Meng, M., & Sinha, P. (2009). Role of ordinal contrast relationships in face encoding. *Proceedings of the National …*, *13*(2009), 5353–5358.

Gilad-Gutnick, S., Yovel, G., & Sinha, P. (2012). Recognizing degraded faces: The contribution of configural and featural cues. *Perception*, *41*(12), 1497–1511.

Goesaert, E., & Op de Beeck, H. P. (2013). Representations of Facial Identity Information in the Ventral Visual Stream Investigated with Multivoxel Pattern Analyses. *Journal of Neuroscience*, *33*(19), 8549–8558.

Gratton, C., Sreenivasan, K. K., Silver, M. A., & D'Esposito, M. (2013). Attention selectively modifies the representation of individual faces in the human brain. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *33*(16), 6979–89.

Grill-Spector, K., Knouf, N., & Kanwisher, N. (2004). The fusiform face area subserves face perception, not generic within-category identification. *Nature Neuroscience*, *7*(5), 555–562.

Grill-Spector, K., & Malach, R. (2001). fMR-adaptation: a tool for studying the functional properties of human cortical neurons. *Acta Psychologica*, *107*(1-3), 293–321.

Grill-Spector, K., & Malach, R. (2004). The human visual cortex. *Annual Review of Neuroscience*, *27*(1), 649–77.

Haig, N. D. (1986). Investigating face recognition with an image processing computer. *Perception*, *15*, 235–247.

Hancock, P. J. B., Bruce, V., & Burton, A. M. (1998). A comparison of two computer-based face identification systems with human perceptions of faces. *Vision Research*, *38*(15-16), 2277–2288.

Hancock, P. J. B., Bruce, V., & Mike Burton, A. (2000). Recognition of unfamiliar faces. *Trends in Cognitive Sciences*, *4*(9), 330–337.

Hanke, M., Halchenko, Y. O., Sederberg, P. B., Hanson, S. J., Haxby, J. V., & Pollmann, S. (2009). PyMVPA: A python toolbox for multivariate pattern analysis of fMRI data. *Neuroinformatics*, *7*(1), 37–53.

Hanson, S. J., & Halchenko, Y. O. (2008). Brain reading using full brain support vector machines for object recognition: there is no “face” identification area. *Neural Computation*, *20*(2), 486–503.

Hanson, S. J., Matsuka, T., & Haxby, J. V. (2004). Combinatorial codes in ventral temporal lobe for object recognition: Haxby (2001) revisited: Is there a “face” area? *NeuroImage*, *23*(1), 156–166.

Harris, R. J., Young, A. W., & Andrews, T. J. (2012). Morphing between expressions dissociates continuous from categorical representations of facial expression in the human brain. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(51), 21164–9.

Harris, R. J., Young, A. W., & Andrews, T. J. (2014a). Brain regions involved in processing facial identity and expression are differentially selective for surface and edge information. *NeuroImage*, *97*, 217–223.

Harris, R. J., Young, A. W., & Andrews, T. J. (2014b). Dynamic stimuli demonstrate a categorical representation of facial expression in the amygdala. *Neuropsychologia*, *56*(1), 47–52.

Harry, B., Williams, M. A., Davis, C., & Kim, J. (2013). Emotional expressions evoke a differential response in the fusiform face area. *Frontiers in Human Neuroscience*, *7*(October), 692.

Haxby, J. V, Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., & Pietrini, P. (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science (New York, N.Y.)*, *293*(5539), 2425–2430.

Haxby, J., Hoffman, E., & Gobbini, M. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences*, *4*(6), 223–233.

Hay, D. C., & Young, A. W. (1982). The human face. In *In A. W. Ellis (Ed.), Normality and pathology in cognitive functions* (pp. 173–202).

Haynes, J.-D. (2015). A Primer on Pattern-Based Approaches to fMRI: Principles, Pitfalls, and Perspectives. *Neuron*, *87*(2), 257–270.

Haynes, J.-D., & Rees, G. (2005). Predicting the stream of consciousness from activity in human visual cortex. *Current Biology : CB*, *15*(14), 1301–7.

Haynes, J.-D., & Rees, G. (2006). Decoding mental states from brain activity in humans. *Nature Reviews. Neuroscience*, *7*(7), 523–34.

Hoffman, E. A., & Haxby, J. V. (2000). Distinct representations of eye gaze and identity in the distributed human neural system for face perception. *Nature Neuroscience*, *3*(1), 80–84.

Hole, G. J., George, P. A., & Dunsmore, V. (1999). Evidence for holistic processing of faces viewed as photographic negatives. *Perception*, *28*(3), 341–359.

Hole, G. J., George, P. A., Eaves, K., & Rasek, A. (2002). Effects of geometric distortions on face-recognition performance. *Perception*, *31*(10), 1221–1240.

Huettel, S. A., Song, A. W., & McCarthy, G. (2009). *Functional Magnetic Resonance Imaging, 2nd Edition* (Vol. 23). Massachusetts: Sinauer.

Ishai, A. (2008). Let’s face it: It's a cortical network. *NeuroImage*, *40*(2), 415–419.

Ishai, A., Schmidt, C. F., & Boesiger, P. (2005). Face perception is mediated by a distributed cortical network. *Brain Research Bulletin*, *67*(1-2), 87–93.

Itz, M. L., Schweinberger, S. R., Schulz, C., & Kaufmann, J. M. (2014). Neural correlates of facilitations in face learning by selective caricaturing of facial shape or reflectance. *NeuroImage*, *102*, 736–747.

Jeffreys, D. A. (1996). Evoked potential studies of face and object processing. *Visual Cognition*, *3*(1), 1–38.

Jiang, F., Blanz, V., & O’Toole, A. J. (2006). Probing the visual representation of faces with adaptation: A view from the other side of the mean. *Psychological Science*, *17*(6), 493–500.

Jiang, F., Dricot, L., Blanz, V., Goebel, R., & Rossion, B. (2009). Neural correlates of shape and surface reflectance information in individual faces. *Neuroscience*, *163*(4), 1078–91.

Johnston, A., Hill, H., & Carman, N. (1992). Recognising faces: effects of lighting direction, inversion, and brightness reversal. *Perception*, *21*(3), 365–75.

Johnston, P., Mayes, A., Hughes, M., & Young, A. W. (2013). Brain networks subserving the evaluation of static and dynamic facial expressions. *Cortex*, *49*(9), 2462–2472.

Kamitani, Y., & Tong, F. (2005). Decoding the visual and subjective contents of the human brain. *Nature Neuroscience*, *8*(5), 679–85.

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *17*(11), 4302–4311.

Kaplan, J. T., & Meyer, K. (2012). Multivariate pattern analysis reveals common neural patterns across individuals during touch observation. *NeuroImage*, *60*(1), 204–212.

Kaufmann, J. M., & Schweinberger, S. R. (2004). Expression influences the recognition of familiar faces. *Perception*, *33*(4), 399–408.

Kemp, R., Pike, G., White, P., & Musselman, a. (1996). Perception and recognition of normal and negative faces: the role of shape from shading and pigmentation cues. *Perception*, *25*(1), 37–52.

Kirby, M., & Sirovich, L. (1990). Applications of the Karhunen-Loeve Procedure for the Characterization of Human Faces. *IEEE Trans. PAMI*, *12*, 103–108.

Kirita, T., & Endo, M. (1995). Happy face advantage in recognizing facial expressions. *Acta Psychologica*, *89*, 149–163.

Kobayashi, H., & Kohshima, S. (1997). Unique morphology of the human eye. *Nature*, *287*(1992), 767–768.

Kosslyn, S., Hamilton, S., & Bernstein, J. (1995). The perception of curvature can be selectively disrupted in prosopagnosia. *Brain and Cognition*, *27*(1), 36–58.

Kriegeskorte, N. (2009). Relating Population-Code Representations between Man, Monkey, and Computational Models. *Frontiers in Neuroscience*, *3*(3), 363–73.

Kriegeskorte, N., Formisano, E., Sorger, B., & Goebel, R. (2007). Individual faces elicit distinct response patterns in human anterior temporal cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(51), 20600–20605.

Kriegeskorte, N., & Kievit, R. A. (2013). Representational geometry: Integrating cognition, computation, and the brain. *Trends in Cognitive Sciences*, *17*(8), 401–412.

Kriegeskorte, N., Mur, M., & Bandettini, P. A. (2008a). Representational similarity analysis - connecting the branches of systems neuroscience. *Frontiers in Systems Neuroscience*, *2*(4), 1–28.

Kriegeskorte, N., Mur, M., & Bandettini, P. A. (2008b). Representational similarity analysis - connecting the branches of systems neuroscience. *Frontiers in Systems Neuroscience*, *2*(November), 4.

Krumhuber, E. G., Kappas, A., & Manstead, A. S. R. (2013). Effects of dynamic aspects of facial expressions: A review. *Emotion Review*, *5*(1), 41–46.

Lai, M., Oruç, I., & Barton, J. J. S. (2011). The role of skin texture and facial shape in representations of age and identity. *Cortex*, 1–14.

Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D. H. J., Hawk, S. T., & van Knippenberg, A. (2010). Presentation and validation of the Radboud Faces Database. *Cognition & Emotion*, *24*(8), 1377–1388.

Leder, H. (1999). Matching person identity from facial line drawings. *Perception*, *28*(9), 1171–1175.

Leder, H., Candrian, G., Huber, O., & Bruce, V. (2001). Configural features in the context of upright and inverted faces. *Perception*, *30*(1), 73–83.

Leonard, C. M., Rolls, E. T., Wilson, F. A. W., & Baylis, G. C. (1985). Neurons in the amygdala of the monkey with responses selective for faces. *Behavioural Brain Research*, *15*(2), 159–176.

Lewis, S., Thoma, R. J., Lanoue, M. D., Miller, G. A., Heller, W., Edgar, C., … Cañive, J. M. (2003). Visual processing of facial affect. *Neuroreport*, *14*(14), 1841–5.

Liu, C. H., Collin, C. A., Burton, A. M., & Chaudhuri, A. (1999). Lighting direction affects recognition of untextured faces in photographic positive and negative. *Vision Research*, *39*(24), 4003–9.

Liu, C. H., Collin, C. A., & Chaudhuri, A. (2000). Does face recognition rely on encoding of 3-D surface? Examining the role of shape-from-shading and shape-from-stereo. *Perception*, *29*(6), 729–743.

Liu, J., Harris, A., & Kanwisher, N. (2010). Perception of face parts and face configurations: an FMRI study. *Journal of Cognitive Neuroscience*, *22*(1), 203–211.

Logothetis, N. K. (2008). What we can do and what we cannot do with fMRI. *Nature*, *453*(7197), 869–878.

Logothetis, N. K., & Wandell, B. A. (2004). Interpreting the BOLD signal. *Annu. Rev. Physiol.*, *66*, 735–769.

Longmore, C. A., Liu, C. H., & Young, A. W. (2008). Learning faces from photographs. *Journal of Experimental Psychology. Human Perception and Performance*, *34*(1), 77–100.

Luria, S. M., & Strauss, M. S. (1978). Comparison of eye movements over faces in photographic positives and negatives. *Perception*, *7*(3), 349–58.

Magnussen, S., Sunde, B., & Dyrnes, S. (1994). Patterns of Perceptual Asymmetry in Processing Facial Expression. *Cortex*, *30*(2), 215–229.

Magnussen, S., Sunde, B., & Dyrnes, S. (1994). Patterns of perceptual asymmetry in processing facial expression. *Cortex*, *30*(2), 215–229.

Marr, D. (1982). *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. San Francisco, CA: W. H. Freeman.

Materna, S., Dicke, P. W., & Thier, P. (2008). Dissociable roles of the superior temporal sulcus and the intraparietal sulcus in joint attention: a functional magnetic resonance imaging study. *Journal of Cognitive Neuroscience*, *20*(1), 108–19.

Mattavelli, G., Sormaz, M., Flack, T., Asghar, A. U. R., Fan, S., Frey, J., … Andrews, T. J. (2013). Neural responses to facial expressions support the role of the amygdala in processing threat. *Social Cognitive and Affective Neuroscience*, 1–6.

Maurer, D., Grand, R. Le, & Mondloch, C. J. (2002). The many faces of configural processing. *Trends in Cognitive Sciences*, *6*(6), 255–260.

Maurer, D., O'Craven, K. M., Le Grand, R., Mondloch, C. J., Springer, M. V., Lewis, T. L., & Grady, C. L. (2007). Neural correlates of processing facial identity based on features versus their spacing. *Neuropsychologia*, *45*(7), 1438–1451.

McKelvie, S. (1976). The role of eyes and mouth in the memory of a face. *The American Journal of Psychology*, *89*(2), 311–323.

McKelvie, S. J. (1973). The meaningfulness and meaning of schematic faces. *Perception & Psychophysics*, *14*(2), 343–348.

McKelvie, S. J. (1995). Emotional expression in upside-down faces: evidence for configurational and componential processing. *The British Journal of Social Psychology / the British Psychological Society*, *34 (Pt 3)*, 325–34.

Meng, M., Cherian, T., Singal, G., & Sinha, P. (2012). Lateralization of face processing in the human brain. *Proceedings. Biological Sciences / The Royal Society*, *279*(1735), 2052–61.

Menzel, C., Hayn-Leichsenring, G. U., Langner, O., Wiese, H., & Redies, C. (2015). Fourier power spectrum characteristics of face photographs: Attractiveness perception depends on low-level image properties. *PLoS ONE*, *10*(4).

Montgomery, D. C., Peck, E. A., & Vining, G. G. (2012). *Introduction to Linear Regression Analysis, 5th Edition*. New York: Wiley.

Natu, V. S., Jiang, F., Narvekar, A., Keshvari, S., Blanz, V., & O'Toole, A. J. (2010). Dissociable neural patterns of facial identity across changes in viewpoint. *Journal of Cognitive Neuroscience*, *22*(7), 1570–1582.

Nederhouser, M., Yue, X., Mangini, M. C., & Biederman, I. (2007). The deleterious effect of contrast reversal on recognition is unique to faces, not objects. *Vision Research*, *47*(16), 2134–42.

Nestor, A., Plaut, D. C., & Behrmann, M. (2011). Unraveling the distributed neural code of facial identity through spatiotemporal pattern analysis. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(24), 9998–10003.

Nestor, A., Plaut, D. C., & Behrmann, M. (2015). Feature-based face representations and image reconstruction from behavioral and neural data. *Proceedings of the National Academy of Sciences*, *113*(2), 201514551.

Nestor, A., Vettel, J. M., & Tarr, M. J. (2013). Internal Representations for Face Detection : An Application of Noise-Based Image Classification to BOLD Responses. *Human Brain Mapping*, *34*(11), 3101–3115.

Nguyen, V. T., Breakspear, M., & Cunnington, R. (2013). Fusing concurrent EEG-fMRI with dynamic causal modeling: Application to effective connectivity during face perception. *NeuroImage*.

Nichols, D. F., Betts, L. R., & Wilson, H. R. (2010). Decoding of faces and face components in face-sensitive human visual cortex. *Frontiers in Psychology*, *1*(JUL), 1–13.

Nili, H., Wingfield, C., Walther, A., Su, L., Marslen-Wilson, W., & Kriegeskorte, N. (2014). A toolbox for representational similarity analysis. *PLoS Computational Biology*, *10*(4), e1003553.

Norman, K. A., Polyn, S. M., Detre, G. J., & Haxby, J. V. (2006). Beyond mind-reading: multi-voxel pattern analysis of fMRI data. *Trends in Cognitive Sciences*, *10*(9), 424–30.

O'Craven, K., & Kanwisher, N. (2000). Mental Imagery of Faces and Places Activates Corresponding Stimulus-Specific Brain Regions. *Journal of Cognitive Neuroscience*, 1013–1023.

O'Donnell, C., & Bruce, V. (2001). Familiarisation with faces selectively enhances sensitivity to changes made to the eyes. *Perception*, *30*(6), 755–764.

O'Toole, A. J., Deffenbacher, K. A., Valentin, D., & Abdi, H. (1994). Structural aspects of face recognition and the other race effect. *Visual Cognition*, *7*(4), 437–463.

O'Toole, A. J., Deffenbacher, K. A., Valentin, D., McKee, K., Huff, D., & Abdi, H. (1998). The perception of face gender: the role of stimulus structure in recognition and classification. *Memory & Cognition*, *26*, 146–160.

O'Toole, A. J., Jiang, F., Abdi, H., & Haxby, J. V. (2005). Partially distributed representations of objects and faces in ventral temporal cortex. *Journal of Cognitive Neuroscience*, *17*(4), 580–90.

O'Toole, A. J., Vetter, T., & Blanz, V. (1999). Three-dimensional Shape and Two-Dimensional Surface Reflactance Contributions to Face Recognition: An Application of Three-Dimensional Morphing. *Vision Research*, *39*, 3145–3155.

Oatley, K., & Johnson-Laird, P. N. (1987). Towards a cognitive theory of emotions. *Cognition & Emotion*, *1*(1), 29–50.

Ohayon, S., Freiwald, W. A., & Tsao, D. Y. (2012). What makes a cell face selective? The importance of contrast. *Neuron*, *74*(3), 567–81.

Op de Beeck, H. P., Haushofer, J., & Kanwisher, N. G. (2008). Interpreting fMRI data: maps, modules and dimensions. *Nature Reviews. Neuroscience*, *9*(2), 123–135.

Orban de Xivry, J.-J., Ramon, M., Lefèvre, P., & Rossion, B. (2008). Reduced fixation on the upper area of personally familiar faces following acquired prosopagnosia. *Journal of Neuropsychology*, *2*(Pt 1), 245–268.

Pallett, P. M., & Meng, M. (2013). Contrast negation differentiates visual pathways underlying dynamic and invariant facial processing. *Journal of Vision*, *13*(14), 1–18.

Peelen, M. V, Atkinson, A. P., & Vuilleumier, P. (2010). Supramodal representations of perceived emotions in the human brain. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *30*(30), 10127–34.

Peirce, J. W. (2008). Generating stimuli for neuroscience using PsychoPy. *Frontiers in Neuroinformatics*, *2*, 10.

Peterson, M. F., & Eckstein, M. P. (2012). Looking just below the eyes is optimal across face recognition tasks. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(48), E3314–23.

Phillips, R. (1972). Why are faces hard to recognize in photographic negative? *Attention, Perception, & Psychophysics*.

Pitcher, D. (2008). Transcranial magnetic stimulations studies of face processing.

Pitcher, D. (2014). Facial expression recognition takes longer in the posterior superior temporal sulcus than in the occipital face area. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *34*(27), 9173–7.

Pitcher, D., Walsh, V., Yovel, G., & Duchaine, B. (2007). TMS evidence for the involvement of the right occipital face area in early face processing. *Current Biology : CB*, *17*(18), 1568–73.

Poldrack, R. a., Halchenko, Y. O., & Hanson, S. J. J. (2009). Decoding the large-scale structure of brain function by classifying mental states across individuals. *Psychological Science*, *20*(11), 1364–1372.

Pourtois, G., Sander, D., Andres, M., Grandjean, D., Reveret, L., Olivier, E., & Vuilleumier, P. (2004). Dissociable roles of the human somatosensory and superior temporal cortices for processing social face signals. *European Journal of Neuroscience*, *20*(12), 3507–3515.

Psalta, L., Young, A. W., Thompson, P., & Andrews, T. J. (2014). The thatcher illusion reveals orientation dependence in brain regions involved in processing facial expressions. *Psychological Science*, *25*(1), 128–136.

Puce, A., Allison, T., Bentin, S., Gore, J. C., & McCarthy, G. (1998). Temporal cortex activation in humans viewing eye and mouth movements. *The Journal of Neuroscience*, *18*(6), 2188–2199.

Puce, A., & Perrett, D. I. (2004). Electrophysiology and brain imaging of biological motion. *The Neuroscience of Social Interaction*, 1–22.

Rakover, S. S., & Teucher, B. (1997). Facial inversion effects: parts and whole relationship. *Perception & Psychophysics*, *59*(5), 752–61.

Ramachandran, V. S. (1988). Perception of shape from shading. *Nature*, *331*(6152), 163–166.

Ramon, M., Dricot, L., & Rossion, B. (2010). Personally familiar faces are perceived categorically in face-selective regions other than the fusiform face area. *European Journal of Neuroscience*, *32*(9), 1587–1598.

Reddy, L., & Kanwisher, N. (2006). Coding of visual objects in the ventral stream. *Current Opinion in Neurobiology*, *16*(4), 408–414.

Rhodes, G., Michie, P. T., Hughes, M. E., & Byatt, G. (2009). The fusiform face area and occipital face area show sensitivity to spatial relations in faces. *The European Journal of Neuroscience*, *30*(4), 721–33.

Rice, G. E., Watson, D. M., Hartley, T., & Andrews, T. J. (2014). Low-Level Image Properties of Visual Objects Predict Patterns of Neural Response across Category-Selective Regions of the Ventral Visual Pathway. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *34*(26), 8837–44.

Rossion, B. (2008). Constraining the cortical face network by neuroimaging studies of acquired prosopagnosia. *NeuroImage*, *40*(2), 423–426.

Rossion, B., Caldara, R., Seghier, M., Schuller, A. M., Lazeyras, F., & Mayer, E. (2003). A network of occipito-temporal face-sensitive areas besides the right middle fusiform gyrus is necessary for normal face processing. *Brain*, *126*(11), 2381–2395.

Rotshtein, P., Henson, R. N. A., Treves, A., Driver, J., & Dolan, R. J. (2005). Morphing Marilyn into Maggie dissociates physical and identity face representations in the brain. *Nature Neuroscience*, *8*(1), 107–13.

Rotshtein, P., Henson, R. N., Treves, A., Driver, J., & Dolan, R. J. (2005). Morphing Marilyn into Maggie dissociates physical and identity face representations in the brain. *Nat Neurosci*, *8*(1), 107–113.

Rozin, P., Lowery, L., & Ebert, R. (1994). Varieties of disgust faces and the structure of disgust. *Journal of Personality and Social Psychology*, *66*(5), 870–881.

Ruiz-Soler, M., & Beltran, F. S. (2006). Face perception: an integrative review of the role of spatial frequencies. *Psychological Research*, *70*(4), 273–92.

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, *39*(6), 1161–1178.

Russell, R. (2003). Sex, beauty, and the relative luminance of facial features. *Perception*, *32*(9), 1093–1107.

Russell, R., & Sinha, P. (2007). Real-world face recognition: The importance of surface reflectance properties. *Perception*, *36*(9), 1368–1374.

Russell, R., Sinha, P., Biederman, I., & Nederhouser, M. (2006). Is pigmentation important for face recognition? Evidence from contrast negation. *Perception*, *35*(6), 749–59.

Sadr, J., Jarudi, I., & Sinha, P. (2003). The role of eyebrows in face recognition. *Perception*, *32*(3), 285–293.

Said, C. P., Moore, C. D., Engell, A. D., & Haxby, J. V. (2010). Distributed representations of dynamic facial expressions in the superior temporal sulcus. *Journal of Vision*, *10*, 1–12.

Said, C. P., Moore, C. D., Norman, K. A., Haxby, J. V., & Todorov, A. T. (2010). Graded representations of emotional expressions in the left superior temporal sulcus. *Frontiers in Systems Neuroscience*, *4*(March), 6.

Santos, I. M., & Young, A. W. (2008). Effects of inversion and negation on social inferences from faces. *Perception*, *37*(7), 1061–1078.

Schiltz, C., & Rossion, B. (2006). Faces are represented holistically in the human occipito-temporal cortex. *NeuroImage*, *32*(3), 1385–94.

Schönemann, P. H. (1966). A generalized solution of the orthogonal procrustes problem. *Psychometrika*.

Schwarzlose, R. F., Swisher, J. D., Dang, S., & Kanwisher, N. (2008). The distribution of category and location information across object-selective regions in human visual cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(11), 4447–4452.

Shinkareva, S. V, Mason, R. A., Malave, V. L., Wang, W., Mitchell, T. M., & Just, M. A. (2008). Using fMRI Brain Activation to Identify Cognitive States Associated with Perception of Tools and Dwellings. *PLoS ONE*, *3*(1), e1394.

Sinha, P. (2000). Here's looking at you, kid. *Perception*, *29*(8), 1005–1008.

Sinha, P. (2002). Qualitative representations for recognition. In *In Lecture Notes for Computer Science Berlin: Springer-Verlag* (pp. 249–262).

Sinha, P., Balas, B. J., Ostrovsky, Y., & Russell, R. (2006). Face Recognition by Humans. *Face Recognition: Models and Mechanisms*, 1–30.

Skerry, A. E., & Saxe, R. (2014). A common neural code for perceived and inferred emotion. *The Journal of Neuroscience : The Official Journal of the Society for*

Neuroscience, *34*(48), 15997–6008.

Sormaz, M., Watson, D. M., Smith, W. A. P., Young, A. W., & Andrews, T. J. (2016). Modelling the perceptual similarity of facial expressions from image statistics and neural responses. *NeuroImage*, *129*, 64–71.

Spiridon, M., & Kanwisher, N. (2002). How distributed is visual category information in human occipito-temporal cortex? An fMRI study. *Neuron*, *35*, 1157–1165.

Steeves, J. K. E., Culham, J. C., Duchaine, B. C., Pratesi, C. C., Valyear, K. F., Schindler, I., … Goodale, M. A. (2006). The fusiform face area is not sufficient for face recognition: Evidence from a patient with dense prosopagnosia and no occipital face area. *Neuropsychologia*, *44*(4), 594–609.

Subramanian, S., & Biederman, I. (1997). Does the contrast reversal affect object identification? *Invest Ophthalmol Vis Sci*, *38*, 998.

Tanaka, J. W., & Farah, M. J. (1993). Parts and wholes in face recognition. *The Quarterly Journal of Experimental Psychology. A, Human Experimental Psychology*, *46*(2), 225–45.

Taubert, J., & Alais, D. (2011). Identity aftereffects, but not composite effects, are contingent on contrast polarity. *Perception*, *40*(4), 422–436.

Tiddeman, B., Burt, M., & Perrett, D. (2001). Prototyping and transforming facial textures for perception research. *IEEE Computer Graphics and Applications*, *21*(5), 42–50.

Tiddeman, B. P., Burt, D. M., & Perrett, D. I. (2001). Prototyping and transforming facial textures for perception research. *IEEE Computer Graphics and Applications*, *21*(October), 42–50.

Tiddeman, B. P., Stirrat, M. R., & Perrett, D. I. (2005). Towards Realism in Facial Image Transformation: Results of a Wavelet MRF Method. *Computer Graphics Forum*, *24*(3), 449–456.

Tomalski, P., Csibra, G., & Johnson, M. H. (2009). Rapid orienting toward face-like stimuli with gaze-relevant contrast information. *Perception*, *38*(4), 569–578.

Tootell, R., Hadjikhani, N., Vanduffel, W., Liu, A., Mendola, J., Sereno, M., & Dale, A. (1998). Functional analysis of primary visual cortex (V1) in humans. *Nature*, *95*(February), 811–817.

Torrance, J. S., Wincenciak, J., Hahn, A. C., DeBruine, L. M., & Jones, B. C. (2014). The relative contributions of facial shape and surface information to perceptions of attractiveness and dominance. *PLoS ONE*, *9*(10), e104415.

Tovée, M. J. (1996). *An Introduction to the Visual System*. *Cambridge: Cambridge University Press*.

Troje, N. F., & Bülthoff, H. H. (1996). Face recognition under varying poses: The role of texture and shape. *Vision Research*, *36*(12), 1761–1771.

Turk, M. A., & Pentland, A. P. (1991). Face recognition using eigenfaces. *Proceedings. 1991 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*.

Vass, L. K., & Epstein, R. A. (2013). Abstract representations of location and facing direction in the human brain. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *33*(14), 6133–42.

Vernon, R. J. W., Sutherland, C. A. M., Young, A. W., & Hartley, T. (2014). Modeling first impressions from highly variable facial images. *Proceedings of the National Academy of Sciences of the United States of America*, *2014*, 1409860111–.

Vuong, Q. C., Peissig, J. J., Harrison, M. C., & Tarr, M. J. (2005). The role of surface pigmentation for recognition revealed by contrast reversal in faces and Greebles. *Vision Research*, *45*(10), 1213–23.

Waller, B. M., Cray, J. J., & Burrows, A. M. (2008). Selection for universal facial emotion. *Emotion*, *8*(3), 435–439.

Watson, D. M., Hartley, T., & Andrews, T. J. (2014). Patterns of response to visual scenes are linked to the low-level properties of the image. *NeuroImage*, *99*, 402–410.

Webster, M. A., Kaping, D., Mizokami, Y., & Duhamel, P. (2004). Adaptation to natural facial categories, *428*(April), 357–360.

Wegrzyn, M., Riehle, M., Labudda, K., Woermann, F., Baumgartner, F., Pollmann, S., … Kissler, J. (2015). Investigating the brain basis of facial expression perception using multi-voxel pattern analysis. *Cortex*, *69*, 131–140.

White, M. (2000). Parts and Wholes in Expression Recognition. *Cognition & Emotion*, *1*(14), 39–60.

White, M. (2001). Effect of photographic negation on matching the expressions and identities of faces. *Perception*, *30*(8), 969–981.

White, M., & Li, J. (2006). Matching faces and expressions in pixelated and blurred photos. *American Journal of Psychology*, *119*(1), 21–28.

Winston, J. S., Henson, R. N. A., Fine-Goulden, M. R., & Dolan, R. J. (2004). fMRI-adaptation reveals dissociable neural representations of identity and expression in face perception. *Journal of Neurophysiology*, *92*(3), 1830–1839.

Woodworth, R. S., & Schlosberg, H. (1954). *Experimental psychology (revised edition).* New York: Henry Holt.

Xu, X., & Biederman, I. (2010). Loci of the release from fMRI adaptation for changes in facial expression, identity, and viewpoint. *Journal of Vision*, *10*, 1–13.

Xu, X., Yue, X., Lescroart, M. D., Biederman, I., & Kim, J. G. (2009). Adaptation in the fusiform face area (FFA): image or person? *Vision Research*, *49*(23), 2800–7.

Yoshizaki, K., & Kato, K. (2011). Contrast polarity of eyes modulates gaze-cueing effect. *Japanese Psychological Research*, *53*(3), 333–340.

Young, A. W., Hellawell, D., & Hay, D. C. (1987). Configurational information in face perception. *Perception*, *16*(6), 747–59.

Young, A. W., Perrett, D. I., Calder, A. J., Sprengelmeyer, R., & Ekman, P. (2002). Facial Expressions of Emotion: Stimuli and Tests(FEEST). *(Thames Valley Test Company, Bury St Edmunds, England)*, (January).

Yovel, G., & Kanwisher, N. (2004). Face perception: Domain specific, not process specific. *Neuron*, *44*(5), 889–898.

Yue, X., Nasr, S., Devaney, K. J., Holt, D. J., & Tootell, R. B. H. (2013). FMRI analysis of contrast polarity processing in face-selective cortex in humans and monkeys. *NeuroImage*, *76*, 57–69.

Yue, X., Tjan, B. S., & Biederman, I. (2006). What makes faces special? *Vision Research*, *46*(22), 3802–11.

Zhang, H., Japee, S., Nolan, R., Chu, C., Liu, N., & Ungerleider, L. G. (2016). Face-selective regions differ in their ability to classify facial expressions. *NeuroImage*, *130*, 77–90.

Zhang, J., Kriegeskorte, N., Carlin, J. D., & Rowe, J. B. (2013). Choosing the rules: distinct and overlapping frontoparietal representations of task rules for perceptual decisions. *The Journal of Neuroscience*, *33*(29), 11852–62.