Article:

Andrews, Timothy John orcid.org/0000-0001-8255-9120, Baseler, Heidi orcid.org/0000-0003-0995-8453, Jenkins, Robert orcid.org/0000-0003-4793-0435 et al. (2 more authors) (2016) Contributions of feature shapes and surface cues to the recognition and neural representation of facial identity. Cortex; a journal devoted to the study of the nervous system and behavior. pp. 280-291. ISSN 1973-8102

https://doi.org/10.1016/j.cortex.2016.08.008

CONTRIBUTIONS OF FEATURE SHAPES AND SURFACE CUES TO THE RECOGNITION AND NEURAL REPRESENTATION OF FACIAL IDENTITY

Timothy J Andrews*, Heidi Baseler, Rob Jenkins, A. Mike Burton and Andrew W Young

Department of Psychology, University of York, York, YO10 5DD, UK

*Corresponding author:
timothy.andrews@york.ac.uk

Figures: 7

Key words: FFA, familiar, shape, surface, identity

ABSTRACT

A full understanding of face recognition will involve identifying the visual information that is used to discriminate different identities and how this is represented in the brain. The aim of this study was to explore the importance of shape and surface properties in the recognition and neural representation of familiar faces. We used image morphing techniques to generate hybrid faces that mixed shape properties (more specifically, second order spatial configural information as defined by feature positions in the 2d-image) from one identity and surface properties from a different identity. Behavioural responses showed that recognition and matching of these hybrid faces was primarily based on their surface properties. These behavioural findings contrasted with neural responses recorded using a block design fMRI adaptation paradigm to test the sensitivity of Haxby et al.'s (2000) core face-selective regions in the human brain to the shape or surface properties of the face. The fusiform face area (FFA) and occipital face area (OFA) showed a lower response (adaptation) to repeated images of the same face (same shape, same surface) compared to different faces (different shapes, different surfaces). From the behavioural data indicating the critical contribution of surface properties to the recognition of identity, we predicted that brain regions responsible for familiar face recognition should continue to adapt to faces that vary in shape but not surface properties, but show a release from adaptation to faces that vary in surface properties but not shape. However, we found that the FFA and OFA showed an equivalent release from adaptation to changes in both shape and surface properties. The dissociation between the neural and perceptual responses suggests that, although they may play a role in the process, these core face regions are not solely responsible for the recognition of facial identity.

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 et al., 1997; Haxby, Hoffman and 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 and colleagues (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.

A distinction between *shape* and *surface* properties is widely used in face perception research (Bruce & Young, 1998, 2012). 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) 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. Shape properties arise from the 3d geometrical description of a face, and how that is projected onto a 2d image.

There are a number of different ways of operationalizing the distinction between shape and surface properties, which allow them to be manipulated quasi-independently. For example O'Toole et al. (1999) use 3d scans to derive 'surface texture' and 'surface shape' descriptions of a particular face. Other approaches derive shape descriptions from

2d images based on the spatial location of fiducial points that correspond to key features (corners of eyes, mouth etc; e.g. Tiddemann, Burt & Perrett, 2001). Of course, within any of these schemes, manipulations of shape and surface properties cannot be fully independent, 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. So, a change in the shape of the cheeks will involve both a shape change (the spatial position of key features) and a surface change (created by the altered pattern of shading) in the same region. Nonetheless, image manipulation methods allow us to hold some aspects of face shape or surface properties fixed as closely as possible. This allows a direct test of the relative contributions of these components of shape and surface information.

In this paper, we examine the 'second order configural properties' of face shape (Maurer et al, 2002), i.e. those defined by the spatial layout of features. These properties are held by many researchers to underlie recognition of the identity of faces (for reviews see McKone & Yovel, 2009; Piepers & Robbins, 2012). Richler, Mack, Gauthier & Palmeri (2009) put this very clearly, writing 'subtle differences in spatial relations between face features being encoded [are] particularly useful for successful recognition of a given face" (p. 2856). More recently, there have been challenges to the notion that identity is perceived through this aspect of face shape (Burton et al, 2015). Nevertheless, 'configural processing' remains a very popular account of face recognition. We therefore investigate the contribution of this aspect of face shape in what follows. We shall use the term 'shape' throughout, while noting that that our analysis is limited to shape as delivered by second order configural properties.

When comparing the relative contributions of shape and surface cues, previous studies have reported that both can contribute to judgements of unfamiliar facial identity

(O'Toole et al., 1999; Jiang et al., 2006). These studies differ from the experiments below in two ways. First, they manipulate shape through 3d surface descriptions, and second they examine *unfamiliar* face identification. Here, we are interested in the perception of familiar face identity, which differs markedly from unfamiliar face perception because the participant has previous experience of seeing familiar faces across many different viewing conditions (Hancock et al. 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, 1994; Burton et al. 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 et al., 2002; Burton, Jenkins, Hancock & White, 2005; Russell et al., 2007; 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 et al., 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). Together, these studies suggest that surface properties of the face are the dominant cue in the recognition of familiar faces.

The aim of this study was to investigate the relative importance of shape and surface properties in the recognition and neural representation of familiar faces. Within the core system of face-selective regions, the fusiform face area (FFA) is thought to be particularly important for the representation of invariant facial characteristics that are important for face recognition (Haxby, Hoffman, & Gobbini, 2000; Grill-Spector, Knouf, & Kanwisher, 2004; Rotshtein et al., 2005). Consistent with the role of the FFA in processing facial identity, fMRI studies have shown a reduced response (adaptation) to repeated images of the same face in the FFA (Grill-Spector et al., 1999; Andrews & Ewbank, 2004). These findings imply that the identity of the face is represented at some level in the FFA and this representation is being adapted by repeated presentations. However, a much stronger test for a link between neural activity and the recognition of facial identity is to determine whether this adaptation is still evident when images of the same identity vary along a dimension that is not important for face recognition (i.e. image-invariant adaptation to identity). In a previous study using this logic, Jiang and colleagues (Jiang et al., 2009; see also Caharel et al., 2009; Itz et al., 2016a) found a release from adaptation in the FFA to changes in both shape and surface properties, suggesting that both properties are represented in this region. Although this runs counter to the behavioral studies of familiar faces, which suggest a greater sensitivity to changes in surface properties and more invariance to changes in shape, Jiang et al's (2009) findings might be explained either by differences in how they manipulated shape (see above) or by their use of unfamiliar faces.

In this study, we therefore introduce a method for investigating the contribution of shape and surface information in familiar face recognition by creating hybrid images that had the 2d shape properties from one identity and the surface properties from a different identity. To remove as much irrelevant variation as much as possible, we followed Burton et al.'s (2005) use of averaged images of each face to minimise the idiosyncrasies of particular photographs. We then used adaptation to determine the sensitivity of face-selective regions to changes in the shape or surface properties of familiar faces. Our predictions were that regions directly responsible for familiar face recognition should adapt to face images that have the same surface properties but vary in shape, and show a release from adaptation to face images that have different surface properties but the same shape.

MATERIALS AND METHODS

Participants

Participants were recruited separately for behavioural and fMRI experiments (Behavioural experiments: n = 22, female = 11, mean age = 20.4 years, SD = 5.4 years; fMRI experiment: n = 20, female = 9, mean age = 26.6 years, SD = 5.0 years). All participants had normal or corrected to normal vision and were drawn from an opportunity sample of students and staff at the University of York. All participants gave their written informed consent. The study was approved by the York Neuroimaging Centre Ethics Committee. Prior to taking part, participants were tested to make sure that they were able to recognize each of the familiar identities used in this study. Participants viewed images (~ 6 x 8 deg) at a distance of 57 cm.

Stimuli

Figure 1 shows the familiar face stimuli used in this study. The images were based on grayscale average images that were generated by combining 12 different images from each of 8 identities familiar to our UK participants (Alan Sugar, Chris Moyles, Derren Brown, Gary Lineker, Jeremy Paxman, Jeremy Kyle, Louis Walsh). Grayscale, average images were used because these provide an estimate of each face's shape and surface properties that removes idiosyncrasies introduced by pose and lighting conditions specific to a particular photograph. The averaging was performed in a graphics program in which key fiducial points (e.g., corners of the mouth, of the eyes, etc.) were manually located in each image, and these were connected to form a grid representing the shape (i.e. the second-order relational properties) of the image (for details see Burton et al, 2015). The average shape for each identity was then determined by combining the location of each point on the grid across all images. To generate the average surface for each identity, each image was deformed (morphed) to a standard shape. In this way, the same part of each image will contain the mouth, the eyes, and so forth. The average surface for each identity is then generated by combining all these images. Finally, the average surface properties can be morphed back to the average shape to create the average image for each identity (see Burton et al, 2005, for full details of this procedure). We should note that the photos used to derive these stimuli were ‘ambient images’ (Jenkins et a, 2011), i.e. they were selected from an internet image search on the celebrities’ names, with no selection criteria except that the full face be visible in high resolution. Arbitrary sampling of such image sets has been shown to give stable averages, even over rather small set sizes (Burton, Kramer, Ritchie & Jenkins, 2016; Jenkins, Burton & White, 2006). The images on the diagonal (top left to bottom right) in Figure 1 depict the combination of average shape and surface properties of the faces of 8 familiar identities. Because the shape and surface information is determined independently, however, it is also possible to combine the shape and surface properties from different identities to create hybrid images. These hybrid faces are represented by the 56 remaining off-diagonal images in Figure 1; images in each row have the same surface information, and images in each column have the same shape (the same fiducial positions).

Behavioural Recognition Experiment

To compare the role of surface and shape cues in the recognition of familiar faces, we asked participants to report the perceived identity of the hybrid familiar faces shown in Figure 1, in which the surface properties were from one identity and the shape properties were from a different identity. Participants viewed one of the hybrid face images (6 x 8 deg) for 5 seconds, after which the image disappeared and was replaced by a fixation cross for 3 seconds. The names of the eight possible familiar faces were presented at the bottom of the screen throughout the entire duration of the 8-second trial, along with a corresponding number from 1 to 8. Participants used a button press to perform an 8-AFC recognition task. Participants could respond at any time during the trial.

Behavioural Matching Experiment

In a complementary behavioural experiment, participants performed a matching task (Fig. 2). In this task, participants viewed one of the eight possible veridical average familiar face images (6 x 8 deg) for 5 seconds, after which the image disappeared and was replaced by a fixation cross for 3 seconds. Eight possible hybrid images were presented at the bottom of the screen throughout the entire duration of the 8-second trial, along with a corresponding number from 1 to 8. One of the faces was the same average face, but the other faces either varied in shape (same surface; the rows in Figure 1) or varied in surface (same shape; the columns in Figure 1). Participants performed a simultaneous matching task in which they had to indicate which of the 8 lower images was identical (i.e. shared both shape and texture) to the original average face image shown at the top of the screen. Participants could respond at any time during the trial.

fMRI Experiment

The fMRI experiment used a block design with 4 different stimulus conditions (Fig. 3): (1) no change (same shape, same surface); (2) shape change (different shape, same surface); (3) surface change (same shape, different surface); (4) shape+surface change (different shape, different surface). A block (rather than event-related) design was chosen because it offers optimal power to detect differences between conditions, and because previous studies using closely related designs have shown differential adaptation effects in our region of principal interest, the FFA (Andrews et al., 2010; Harris et al., 2012; Davies Thompson et al., 2013).

To meaningfully interpret differences in neural adaptation, it is useful to know how different were the images in each block condition. Suppl. Fig. 1 shows the mean change in image intensity across images. This was calculated by taking the average of the absolute differences in gray value at each pixel for successive pairs of images within a block. Suppl. Fig. 1 also shows the correlation across corresponding pixel values for successive pairs of images within a block. Shape and surface changes had a similar effect on these low-level image measures. The largest change in low-level properties was found when both shape and surface properties changed.

Within each stimulus block in the fMRI experiment, each image was presented for 975ms followed by a 150ms blank screen. Eight images were shown per block, resulting in a block length of 9s. Each stimulus condition was repeated 8 times. This gave a total of 32 blocks, which were presented in a counterbalanced order. Blocks were separated by a 9s fixation screen (a white fixation cross on a mean gray background).

To maintain attention during the scan, participants performed a red dot detection task in which they were required to press a button when a red dot appeared on any of the images. Mean accuracy was 88% across all conditions (no change: 87%, shape change: 88%, surface change: 90%; shape+surface change: 90%). Mean response time (RT) across conditions was 494 msec (no change: 480, shape change: 488, surface change: 490; shape+surface change: 515). The use of this incidental red dot detection task means that any differential effort resulting from trying explicitly to recognise the images in each type of block would not create a confound with the patterns of neural adaptation in face-selective regions. The same incidental task has been used in other studies for similar reasons (Andrews et al., 2010; Harris et al., 2012). None the less, it is well known that familiar face recognition is a mandatory process (Young et al., 1986; Lavie et al., 2003) in which the identities of familiar faces cannot be ignored, and to this extent any influence of normal, automatic face recognition will be evident despite its not being explicitly required.

Face-selective regions were identified from an independent localiser scan. This localiser used a block design with 7 different conditions: same identity faces, different identity faces, bodies, inanimate objects, places, and scrambled images. The faces in the localizer scan were different to those used in the main experiment. Each block consisted of 10 images from each condition; each image was presented for 700 ms and followed by a 200-ms blank screen. Stimulus blocks were separated by a 9-s gray screen with a central fixation cross. Each condition was repeated 4 times in a counterbalanced design.

Data from the fMRI experiment were collected using a GE 3 Tesla HD Excite MRI scanner at the York Neuroimaging Centre at the University of York. A T1-weighted structural MRI (2.25 x 2.25 x 3mm voxel) and a gradient-echo EPI were acquired for each participant. A gradient-echo EPI sequence with a radio-frequency coil tuned to 127.4 MHz was used to acquire 38 axial slices (TR = 3s, TE = 33ms, flip angle = 90°, FOV = 260mm, matrix size = 128 x 128, slice thickness = 3mm, voxel size: 2.25 x 2.25 x 3mm). Data were analysed with FEAT version 4.1 (http://www.fmrib.ox.ac.uk/fsl). The first 9 seconds (3 volumes) from each scan were discarded, and MCFLIRT motion correction, spatial smoothing (Gaussian, FWHM 6mm), and temporal high-pass filtering (cutoff 0.0093Hz) were applied. The BOLD response for each condition was modelled with a boxcar function convolved with a standard haemodynamic response function.

Face-selective regions were individually defined in each individual using the localiser scan. Face-selective voxels were defined by comparing the response to faces with the response to each non-face condition. These statistical maps were averaged and thresholded at $p<0.001$ (uncorrected). Contiguous clusters of voxels located within the occipital and temporal lobes were defined as the FFA, OFA and pSTS in each participant. We were not able to localize other face regions such as anterior face patch (Rajimehr et al., 2009). This may reflect signal dropout in this region of the brain with the EPI sequence used in this experiment.

To analyse the data from the experimental scan, the time-course of response from each voxel within each ROI was converted from units of image intensity to percentage signal change. Voxels were then averaged to give a single time series for each ROI in each participant. Individual stimulus blocks were normalized by subtracting the response at the start of the block from the response at every time point and then averaged to produce a mean time series for each condition for each participant. The peak response was calculated as the average of the percent BOLD signal change at 9 and 12 seconds post-stimulus. In order to determine significant differences in the peak response to each stimulus condition, repeated measures ANOVAs were conducted across participants.

RESULTS

Behavioural Recognition Task

In the recognition task participants had to report the perceived identity of hybrid face images that contained the shape from one familiar individual and the surface from another individual. For each trial, we determined whether the identity reflected the shape information in the hybrid image, the surface information in the hybrid image or neither the shape nor the surface information. The results shown in Figure 4 reveal that participants reported the identity based on the surface properties (mean = 90.4%, SE = 2.2%) on more trials compared to when they used the shape properties (mean = 4.4%, SE = 0.8%; $t(17) = 30.32$, $p < 0.001$). Similarly, response times for trials in which the identity was reported based on the surface properties (mean = 2156 ms, SE = 121ms) were significantly shorter than response times on trials in which the identity was reported based on the shape (mean = 2938 ms, SE = 248 ms; $t(17) = -3.28$, $p = 0.004$).

Behavioural Matching Task

The matching experiment determined participants' ability to discriminate face images that varied in either shape or surface properties. Figure 5 shows the number of errors when judging faces that varied in shape or in surface cues. These results show that participants responded more accurately and more quickly when the task involved faces that varied in surface properties (accuracy: mean = 96.9%, SE = 1.1%; response time: mean = 2782ms, SE = 85ms) compared to when they varied in shape (accuracy: mean = 67.3%, SE = 3.5%; response time: mean = 4088 ms, SE = 152ms) (surface vs. shape differences: accuracy: $t(21) = 9.19$, $p < 0.001$; response time: $t(21) = -9.90$, $p < 0.001$). Together, the results from these behavioural recognition and matching experiments consistently show the relative importance of surface compared to shape properties in the recognition of these familiar face stimuli.

fMRI Experiment

A localiser scan revealed face-selective regions that corresponded to the fusiform face area (FFA), occipital face area (OFA) and posterior superior temporal sulcus (pSTS). The location of these regions is shown in Figure 6 and Table 1. Next, we determined how these regions responded to changes in shape and surface properties. First, a 2 × 4 × 2 ANOVA with the factors Region (FFA, OFA) Condition (no change; shape change; surface change; shape+surface change) and Hemisphere (right, left) was conducted to determine whether the corresponding regions of interest in the two hemispheres responded differently. The STS was not included in this part of the analysis as it was only identified in the right hemisphere. There was no main effect of Hemisphere ($F(1,19) = 0.48$, $p = 0.50$). There was also no interaction between Hemisphere*Region ($F(1,19) = 0.27$, $p = 0.61$), Hemisphere*Condition ($F(3,57) = 2.26$, $p = 0.09$) or Hemisphere*Region*Condition ($F(3,57) = 0.65$, $p = 0.59$). As there were no significant effects of hemisphere, the time courses were averaged across hemispheres for each region in all further analyses.

Figure 7 shows the time course of response to different conditions in the different REGIONs. A 3 x 4 ANOVA, with the factors Region (FFA, OFA, pSTS) and Condition (no change, shape change; surface change; shape+surface change), was then performed on the data. This showed a significant effect of Region ($F(2,36) = 23.82$, $p < 0.001$), Condition ($F(3,54) = 14.96$, $p < 0.001$) and an interaction between Region*Condition ($F(6,108) = 7.24$, $p < 0.001$). To explore these effects, we focused on the pattern of response in each region.

In the FFA, there was a lower response (adaptation) to the no change condition compared to the shape change ($t(19) = -4.78$, $p < 0.001$), surface change ($t(19) = -5.79$, $p <$

0.001) and shape+surface change ($t(19) = -6.70$, $p < 0.001$) conditions. However, there was no difference in response between the shape+surface change condition and either the shape change ($t(19) = 1.33$, $p = 0.20$) or surface change ($t(19) = 1.42$, $p = 0.17$) conditions.

The OFA showed a similar pattern of response to the FFA. There was a lower response (adaptation) to the no change condition compared to the shape change ($t(19) = -4.39$, $p < 0.001$), surface change ($t(19) = -4.86$, $p < 0.001$) and shape+surface change ($t(19) = -5.19$, $p < 0.001$) conditions. However, there was no difference in response between the shape+surface change condition and either the shape change ($t(19) = 0.07$, $p = 0.95$) or surface change ($t(19) = -0.03$, $p = 0.97$) conditions.

In contrast to the FFA and OFA, the pSTS did not show a lower response (adaptation) to the no change condition compared to the shape+surface change condition ($t(18) = -0.40$, $p = 0.69$). However, the pSTS responded more to the shape change condition compared to all other conditions (no change: $t(18) = 4.42$, $p < 0.001$; surface change: $t(18) = 2.48$, $p = 0.02$; shape+surface change: $t(18) = 2.81$, $p = 0.01$).

Although our choice of FFA, OFA and pSTS as REGIONs was determined *a priori* from Haxby et al.'s (2000) neural model, we also used a whole-brain group analysis to investigate responses in the ventral stream outside these face-selective regions. Supplementary Figure 2 shows the response to place-selective regions (blue) and face-selective regions (red). Adaptation to the no change condition compared to the shape+surface condition did not show significant overlap with place-selective regions, but did overlap with face-selective regions. There was also no evidence of adaptation to the shape change or surface change conditions across the ventral visual pathway.

Finally, we determine the extent to which the data could reflect processing in early visual areas. A control region, which was visually responsive but not face selective, was also defined for each participant by transforming the anatomical occipital pole region from the Harvard-Oxford Cortical Structural Atlas in the MNI standard brain into the participant's functional data space. Supplementary Figure 3 shows the time-course of response in this region. There were no differences in response to the different conditions (F(3,57) = 2.06, p = 0.115).

DISCUSSION

The behavioural part of our study used matching and recognition tasks to investigate the type of visual information that is important for the recognition of facial identity. Specifically, we focussed on the roles of shape and surface properties. To address this issue, we created hybrid images that combined aspects of shape from one identity and surface from a different identity. Our results from both behavioural tasks clearly show that these surface properties are more important than the shape properties for the recognition of facial identity.

The fMRI experiment built on these behavioural results by using the logic that a neural region that is directly responsible for the recognition of facial identity should show a corresponding differential sensitivity to surface over shape information. This was achieved with a block design paradigm in which neural responses to changes in surface, shape, or both surface and shape were compared to a 'no change' baseline that would create maximal adaptation. We found that the neural responses in the FFA were equally sensitive to changes in shape *and* surface properties of faces. This difference between behavioural findings and the pattern of neural responses in FFA implies that the FFA does not contain an image-invariant representation of identity that could contribute directly to face recognition. Similar findings held for the OFA.

Our behavioural findings show that when participants were asked to recognize hybrid familiar faces that contained the surface properties from one identity and the shape properties from a different identity, they used the surface properties on more than 90% of trials and the shape properties on less than 5% of trials. Similarly, in a matching task in which the faces differed in only shape or in only surface properties, participants were more accurate and faster at making discriminations based on changes in the surface properties.

These results contrast with some previous studies that suggest both shape and surface properties are important for the recognition of facial identity (O'Toole et al., 1999; Jiang et al., 2006). One key difference between these studies and our own is that we have defined shape explicitly in terms of second order configural relations – exactly those aspects of shape that are generally held to underlie face recognition (Maurer et al, 2002; Tanaka & Gordon, 2011). Our results show quite clearly that shape, defined in this way, delivers rather poor identity information by comparison to surface properties. On the other hand, studies using full 3d shape information appear to show a greater influence of shape in recognition. This suggests that the standard definitions of spatial feature layout, commonly used in configural accounts of face recognition, are inadequate (Burton et al, 2015).

A further source of discrepancy is that most previous studies have used unfamiliar faces and constrained the range of image variability by taking the images from relatively standardised or artificially generated sets. Because it is not possible to ask participants to recognize an unfamiliar face, these studies typically use matching tasks in which participants are asked to determine whether face images belong to the same or a different identity, and often treat any difference between images as if these constituted different faces. It is possible, therefore, that in experiments with unfamiliar faces participants may use features that are specific to particular image sets, but are not stable across a wider range of viewing conditions. The recognition of familiar faces is based on the experience of seeing many different exemplars from the same identity. This allows observers to discern invariant features of the face that are common across previous images. Our findings suggest that the surface properties that we have defined play a key role in the invariant representation that leads to the recognition of familiar faces (see also, Itz et al., 2014; 2016b). The reason shape information may not be such a reliable cue for the recognition of identity is that it is less stable across images (Burton, 2013). For example, rigid and non-rigid movements of the head can have a marked effect on the perceived 2d (and often the 3d) shape of the face. Other studies have shown that familiar face recognition is relatively unimpaired if the surface properties are presented on a standardized shape (Burton et al., 2005) or when the shape is distorted by stretching (Hole et al., 2002; Sandford & Burton, 2014).

None of the core face-selective regions showed any bias toward representing the surface properties of the face. Consistent with previous studies (Grill-Spector et al., 1999; Andrews & Ewbank, 2004), we found a lower response (adaptation) in the FFA and OFA to repeated images of the same face compared to faces that differed in shape, surface, or both. Our predictions were that face-selective regions responsible for the recognition of facial identity should be insensitive or invariant to changes in shape when the surface properties are held constant, but sensitive to changes in surface properties when the shape properties are held constant. In contrast to these predictions, we found a release from adaptation in the FFA and OFA to changes in both shape and surface properties. The lack of adaptation to the change in surface properties could not be explained by greater low-level image differences in this condition, as these were similar for both the shape change and surface change conditions (see Suppl. Fig. 1). Indeed, the fact that the low-level image properties showed similar variation demonstrates a dissociation with the representation involved in face recognition. So, although familiar faces with the same surface properties but different shapes were recognized as belonging to the same identity, the response in these core face-selective regions did not show any adaptation.

How do these findings sit with the idea that the FFA is involved in processing invariant characteristics of faces (Haxby et al., 2000; Gobbini and Haxby, 2007)? Identity is obviously central to these invariant characteristics, making it tempting to equate the FFA with an area primarily responsible for face recognition. However, previous neuroimaging studies have reported mixed results about whether the FFA has an image-invariant representation of identity. Some studies have reported image invariance (Loffler et al. 2005; Rotshtein et al. 2005; Eger et al. 2005; Ewbank and Andrews, 2008; Davies-Thompson et al., 2013), whereas others have reported image dependence (Grill-Spector et al. 1999; Andrews and Ewbank 2004; Pourtois et al. 2005; Davies-Thompson et al. 2009; Xu et al. 2009; Weibert and Andrews, 2015). Indeed, our results are similar to data reported by Jiang and colleagues who showed a release from adaptation to changes in shape and surface properties of unfamiliar faces in the OFA and FFA (Jiang et al., 2009). A similar sensitivity to shape and surface properties has been reported in EEG response to faces (Schulz et al., 2012; Itz et al., 2014; Dzhelyova & Rossion, 2014). More generally, these results are consistent with previous studies that have shown patterns of response in face regions are sensitive to the image properties (Yue et al., 2006; Xu al., 2009; Rice et al., 2014; Watson et al., 2016). For example, patterns of response in the fusiform gyrus to faces can be predicted by their image properties (Rice et al., 2014). Moreover, equivalent changes in the image statistics that result in either a change in identity or no change in identity lead to an equivalent release from adaptation in regions such as the OFA and FFA (Yue et al., 2006; Xu et al., 2009).

It seems, then, that the involvement of FFA in processing invariant characteristics of faces such as identity is not at the level where full image invariance is achieved. This might for instance happen because the FFA contributes to the early stages of face recognition, perhaps being involved in some form of image normalisation. This would be consistent with studies of congenital prosopagnosia which report normal patterns of response in face regions can occur despite disrupted face recognition (Furl et al., 2011; Avidan and Behrmann, 2014). These findings should not undermine the important role of core regions such as the FFA and OFA in face processing. For example, patterns of response in the FFA have been linked with individual differences in familiar face recognition (Furl et al., 2011; Weibert and Andrews, 2015). The contribution of these core regions is further supported by lesion studies (Rossion et al., 2003; Barton, 2008) and by the finding that direct electrical stimulation of these regions selectively disrupts face perception (Parvizi et al., 2012; Jonas et al., 2012). Our point is simply that the FFA does not itself show the functional properties that explicitly characterise familiar face recognition; this point is not intended to deny its importance to the process. Indeed, this is consistent with the idea that interactions between core regions such as the FFA and regions in the extended face processing network, particularly in the anterior temporal lobe, are important for the explicit representation of facial identity (Collins and Olson, 2014).

Neural models of human face perception also propose a pathway leading to the posterior superior temporal sulcus (pSTS) that is responsible for processing changeable aspects of faces such as gaze and expression (Haxby et al., 2000). The pSTS showed a different pattern of response compared to the OFA and FFA. This is consistent with previous studies that have shown a differential sensitivity to shape and surface properties in these brain regions (Harris et al., 2014). The pSTS did not show any adaptation to identity, but showed a larger response to changes in shape when the surface properties were unchanged. One interpretation of these findings is that the changes in shape within a block are being interpreted as dynamic transformations of an individual face (see Lee et al., 2010;

Pitcher et al., 2011). This finding is consistent with previous studies that have shown that the pSTS responds more to sequences of faces varying in gaze and expression in which the identity was constant compared with sequences in which the identity varied (Andrews and Ewbank, 2004; Davies-Thompson et al., 2009; Baseler et al., 2014). Indeed, to be socially meaningful, changes in expression and gaze must be tracked across an individual face. The larger response to changes in shape compared to changes in both the shape and texture implies that activity in the pSTS is modulated by the identity of the face images.

CONCLUSIONS

Our results show that the recognition of facial identity is dominated the surface properties of the image, by comparison to the spatial layout of features. When participants were shown hybrid faces that contained the shape from one identity and the surface properties from another identity, they reported the identity based on the surface properties. Based on the behavioural data, we predicted that a region responsible for face recognition should show adaptation to faces with the same surface properties, and this should be apparent even across differences in shape. However, we found that face-selective regions, such as the FFA, showed an equal sensitivity to both shape and surface properties. This difference between the neural and perceptual responses to facial identity suggests that while the FFA may contribute to the early stages of analysis of invariant characteristics of faces, it is not itself responsible for the recognition of facial identity.

REFERENCES

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., Davies-Thompson, J., Kingstone, A., & Young, A. W. (2010). Internal and external features of the face are represented holistically in face-selective regions of visual cortex. Journal of Neuroscience, 30, 3544-3552.

Avidan, G., & Behrmann, M. (2014). Structural and functional impairment of the face processing network in congenital prosopagnosia. Frontiers in Bioscience, 6, 236-257.

Baseler, H. A., Harris, R. J., Young, A. W., & Andrews, T. J. (2014). Neural responses to expression and gaze in the posterior superior temporal sulcus interact with facial identity. Cerebral Cortex, 24, 737-744.

Barton, J. J. (2008). Structure and function in acquired prosopagnosia: Lessons from a series of 10 patients with brain damage. Journal of Neuropsychology, 2, 197–225.

Bruce, V. (1994). Stability from variation: The case of face recognition. Quarterly Journal of Experimental Psychology, 47, 5-28.

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

Bruce, V., & Young, A. (2012). Face perception. Hove, East Sussex: Psychology Press.

Burton, A. M. (2013). Why has research in face recognition progressed so slowly? The importance of variability. Quarterly Journal of Experimental Psychology, 66, 1467-1485.

Burton, A. M., Bruce, V., & Hancock, P. J. (1999). From pixels to people: A model of familiar face recognition. Cognitive Science, 23(1), 1-31.

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

Burton, A. M., Kramer, R. S. S., Ritchie, K. L., & Jenkins, R. (2016). Identity from variation: Representations of faces derived from multiple instances. Cognitive Science, 40(1), 202–223.

Burton, A. M., Schweinberger, S. R., Jenkins, R., & Kaufmann, J.M., (2015). Arguments against a configural processing account of familiar face recognition. Perspectives on Psychological Science, 10(4), 482–496.

Caharel, S., Jiang, F., Blanz, V., & Rossion, B. (2009). Recognizing an individual face: 3D shape contributes earlier than 2D surface reflectance information. Neuroimage, 47, 1809–1818.

Collins, J. A., & Olson, I. R. (2014). Beyond the FFA: The role of the ventral anterior temporal lobes in face processing. Neuropsychologia, 61, 65-79.

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

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.

Davies-Thompson, J., Newling, K., & Andrews, T. J. (2013). Image-invariant responses in face-selective regions do not explain the perceptual advantage for familiar face recognition. Cerebral Cortex, 23, 370-377.

Dzhelyova M, & Rossion B. (2014). Supra-additive contribution of shape and surface information to individual face discrimination as revealed by fast periodic visual stimulation. Journal of Vision, 14(14).

Eger, E., Schweinberger, S., Dolan, R., & Henson, R. (2005). Familiarity enhances invariance of face representations in human ventral visual cortex: fMRI evidence. Neuroimage, 26(4), 1128-1139.

Ewbank, M. P., & Andrews, T. J. (2008). Differential sensitivity for viewpoint between familiar and unfamiliar faces in human visual cortex. Neuroimage, 40(4), 1857-1870.

Furl, N., Garrido, L., Dolan, R. J., Driver, J., & Duchaine, B. (2011). Fusiform gyrus face selectivity relates to individual differences in facial recognition ability. Journal of Cognitive Neuroscience, 23 (7), 1723–1740.

Gobbini, I. M., & Haxby, J. V. (2007). Neural systems for recognition of familiar faces. Neuropsychologia, 45(1), 32-41.

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., Kushnir, T., Edelman, S., Avidan, G., Itzchak, Y., & Malach, R. (1999). Differential processing of objects under various viewing conditions in the human lateral occipital complex. Neuron, 24(1), 187-203.

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

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, 109, 21164–21169.

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

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

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

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, 736–47.

Itz, M. L., Schweinberger, S. R., & Kaufmann, J. M. (2016a). Effects of caricaturing in shape or color on familiarity decisions for familiar and unfamiliar faces. PLoS ONE, 11(2), e0149796.

Itz, M. L., Golle, J., Luttmann, S., Schweinberger, S. R., & Kaufmann, J. M. (2016b). Dominance of texture over shape in facial identity processing is modulated by individual abilities. British Journal of Psychology, DOI: 10.1111/bjop.12199.

Jenkins, R., Burton, A. M., & White, D. (2006). Face recognition from unconstrained images: Progress with prototypes. In Proceedings of 7th IEEE Conference on Automatic Face and Gesture Recognition. (p.25-30). Los Alamitos, CA: IEEE.

Jenkins, R., White, D., Van Montfort, X., & Burton, A. M. (2011). Variability in photos of the same face. Cognition, 121(3), 313–323.

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, 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, 1078–1091

Jonas, J., Descoins, M., Koessler, L., Colnat-Coulbois, S., Sauvée, M., Guye, M., Vignal, J. P., Vespignani, H., Rossion, B., & Maillard, L. (2012) Focal electrical intracerebral stimulation of a face-sensitive area causes transient prosopagnosia. Neuroscience, 222, 281–288, Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. Journal of Neuroscience, 17(11), 4302-4311.

Lavie, N., Ro, T., & Russell, C. (2003). The role of perceptual load in processing distractor faces. Psychological Science, 14, 510-515.

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

Lee, L. C., Andrews, T. J., Johnson, S. J., Woods, W., Gouws, A., Green, G. G. R., & Young, A. W. (2010). Neural responses to rigidly moving faces displaying shifts in social attention investigated with fMRI and MEG. Neuropsychologia, 48, 477-90.

Loffler, G., Yourganov, G., Wilkinson, F., & Wilson, H. R. (2005). fMRI evidence for the neural representation of faces. Nature Neuroscience, 8(10), 1386-1391.

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

McKone, E., & Yovel, G. (2009). Why does picture-plane inver- sion sometimes dissociate perception of features and spac- ing in faces, and sometimes not? Toward a new theory of holistic processing. Psychonomic Bulletin & Review, 16, 778–797.

O'Toole, A. J., Price, T., Vetter, T., Bartlett, J. C., & Blanz, V. (1999). 3D shape and 2D textures of human faces: The role of "averages" in attractiveness and age. Image Vision and Computing, 18, 9-19.

Parvizi, J., Jacques, C., Foster, B. L., Witthoft, N., Rangarajan, V., Weiner, K. S., & Grill-Spector, K. (2012) Electrical stimulation of human fusiform face-selective regions distorts face perception. Journal of Neuroscience 32, 14915–14920.

Piepers, D. W., & Robbins, R. A. (2012). A review and clarifica- tion of the terms “holistic,” “configural,” and “relational” in the face perception literature. Frontiers in Psychology, 3, Article 559.

Pitcher, D., Dilks, D. D., Saxe, R. R., Triantafyllou, C., & Kanwisher, N. (2011). Differential selectivity for dynamic versus static information in face selective cortical regions. NeuroImage, 56, 2356-2363

Pourtois, G., Schwartz, S., Seghier, M. L., Lazeyras, F., & Vuilleumier, P. (2005). Portraits or people? Distinct representations of face identity in the human visual cortex. Journal of Cognitive Neuroscience, 17(7), 1043-1057.

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, 2381-2395

Rajimehr, R., Young, J. C., & Tootell, R. B. H. (2009). An anterior temporal face patch in human cortex predicted by macaque maps. Proceedings of the National Academy of Sciences , 106, 1995-2000.

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. Journal of Neuroscience, 34, 8837–8844.

Richler, J. J., Mack, M. L., Gauthier, I., & Palmeri, T. J. (2009). Holistic processing of faces happens at a glance. Vision Research, 49, 2856–2861.

Rotshtein, P., Henson, R. N. A., Treves, A., Driver, J., & Dolan, R. J. (2004). Morphing Marilyn into Maggie dissociates physical and identity face representations in the brain. Nature Neuroscience, 8(1), 107-113.

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

Russell, R., Biederman, I., Nederhouser, M., & Sinha, P. (2007). The utility of surface reflectance for the recognition of upright and inverted faces. Vision Research, 47, 157–165.

Sandford, A., & Burton, A. M. (2014). Tolerance for distorted faces: Challenges to a configural processing account of familiar face recognition. Cognition, 132, 262-8.

Schulz, C., Kaufmann, J. M., Walther, L., & Schweinberger, S. R. (2012). Effects of anticaricaturing vs. caricaturing and their neural correlates elucidate a role of shape for face learning. Neuropsychologia, 50, 2426–34.

Tanaka, J. W., & Gordon, I. (2011). Features, configuration and holistic face processing. In A. J. Calder, G. Rhodes, M. H. Johnson, & J. V. Haxby (Eds.), The Oxford handbook of face perception (pp. 15–30). Oxford, United Kingdom: Oxford University Press.

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

Watson, D. M., Young, A. W., & Andrews, T. J. (2016). Spatial properties of objects predict patterns of neural response in the ventral visual pathway. Neuroimage, 126, 173-183.

Weibert, K., & Andrews, T. J. (2015). Activity in the right fusiform face area predicts the behavioural advantage for the perception of familiar faces. Neuropsychologia, 75, 588-596.

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), 2800-2807.

Young, A. W., Ellis, A. W., Flude, B. M., McWeeny, K. H., & Hay, D.C. (1986). Face-name interference. Journal of Experimental Psychology: Human Perception and Performance. 12, 466–475.

Yue, X., Tjan, B. S., & Biederman, I. (2006). What makes faces special? Vision Research, 46, 3802-3811.

Region	Mean coordinates (Standard Error)		
	x	***y***	***z***
FFA			
L	-41.15 (0.86)	-55.76 (1.93)	-21.35 (0.63)
R	41.28 (0.61)	-54.95 (1.58)	-21.26 (0.76)
OFA			
L	-40.37 (1.11)	-81.24 (1.17)	-12.96 (1.01)
R	40.96 (1.05)	-82.34 (1.18)	-12.70 (0.94)
STS			
R	50.67 (1.42)	-51.43 (1.85)	6.93 (1.08)

Table 1 MNI coordinates of face-selective regions of interest defined in the localizer scan.

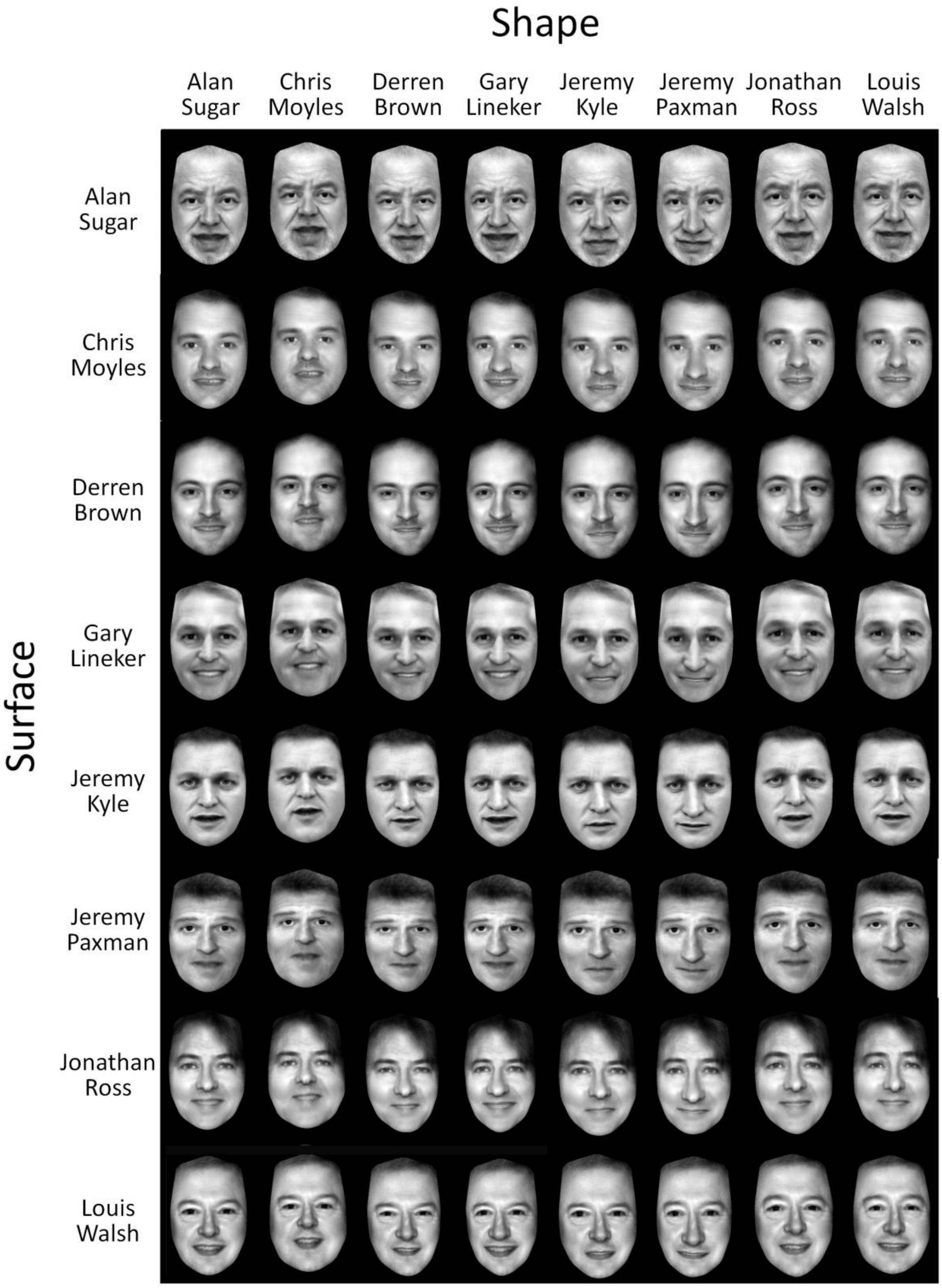


Figure 1 Images used in behavioural and fMRI experiments. Hybrid faces were generated by creating an average image of each individual (shown along the diagonal from top left to bottom right of the Figure) and then combining the 2d shape (image fiducials) from one identity with the surface from another identity (leading to the off-diagonal images in the Figure). Images in each row have the same surface information, and images in each column have the same shape.

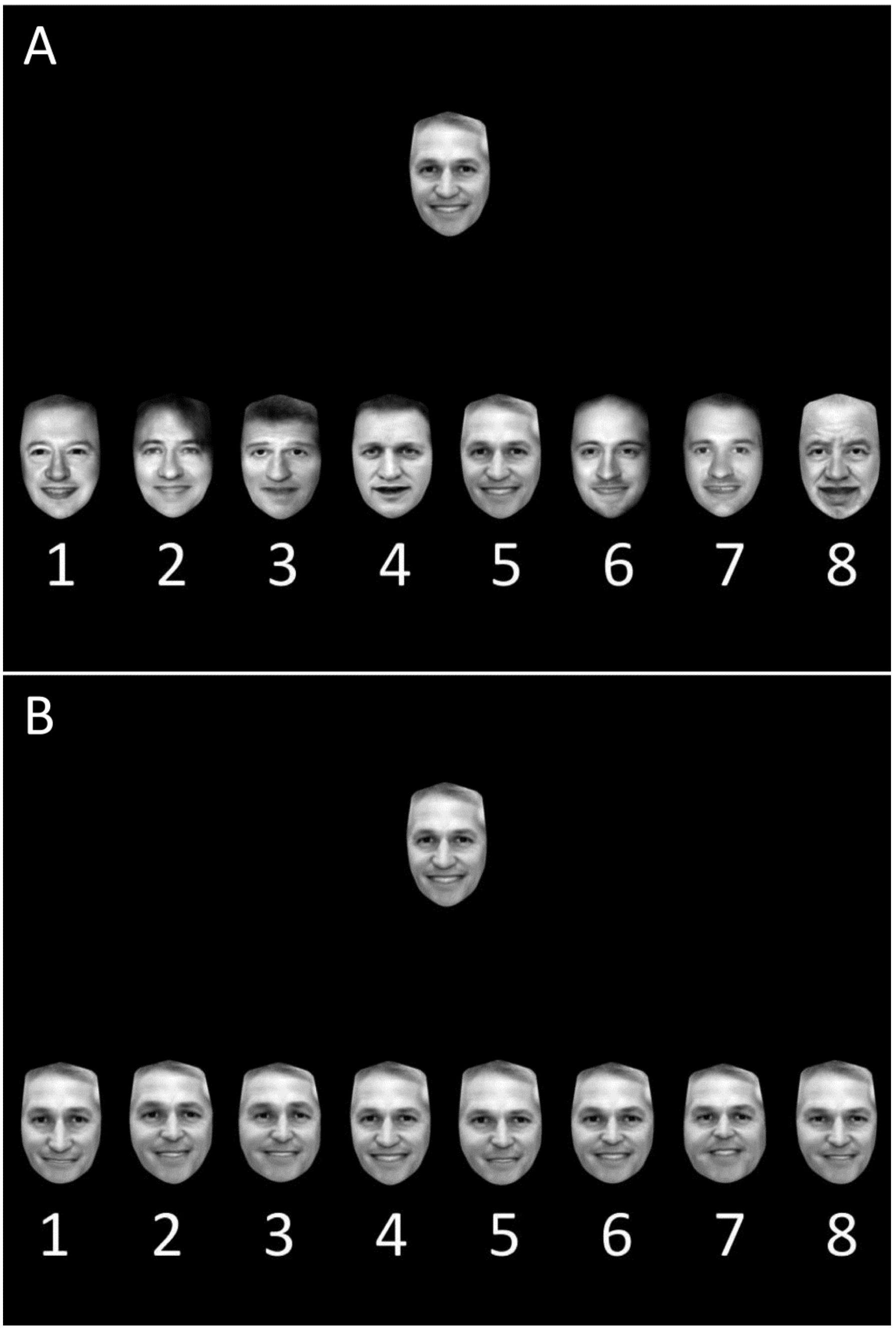


Figure 2 Examples of trials from the behavioural matching experiment. Participants were presented with a face with the shape and surface from a familiar identity (top of each panel). They were then presented with an array of faces (bottom of each panel) that either had (A) the same 2d shape, but different surfaces or (B) the same surface, but different shapes. The task was to match the original image to the identical image in the test array.

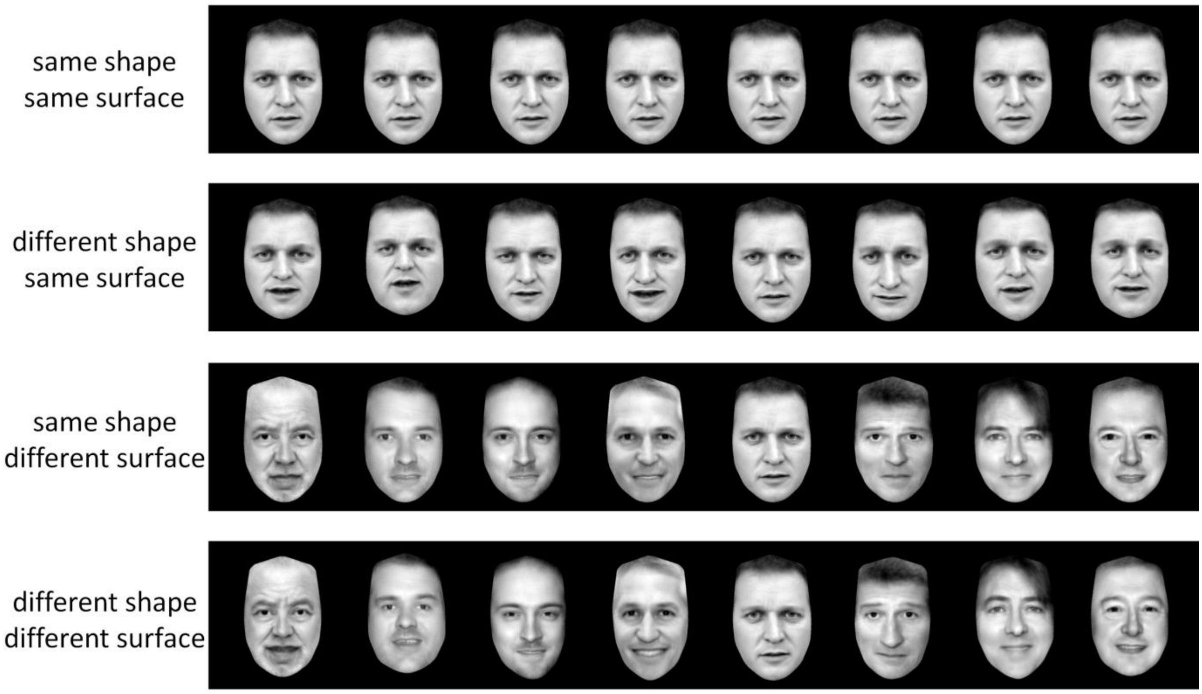


Figure 3 Design of fMRI experiment. There were 4 conditions, illustrated here with images in the different rows: (1) same shape, same surface; (2) different shape, same surface; (3) same shape, different surface; (4) different shape, different surface.

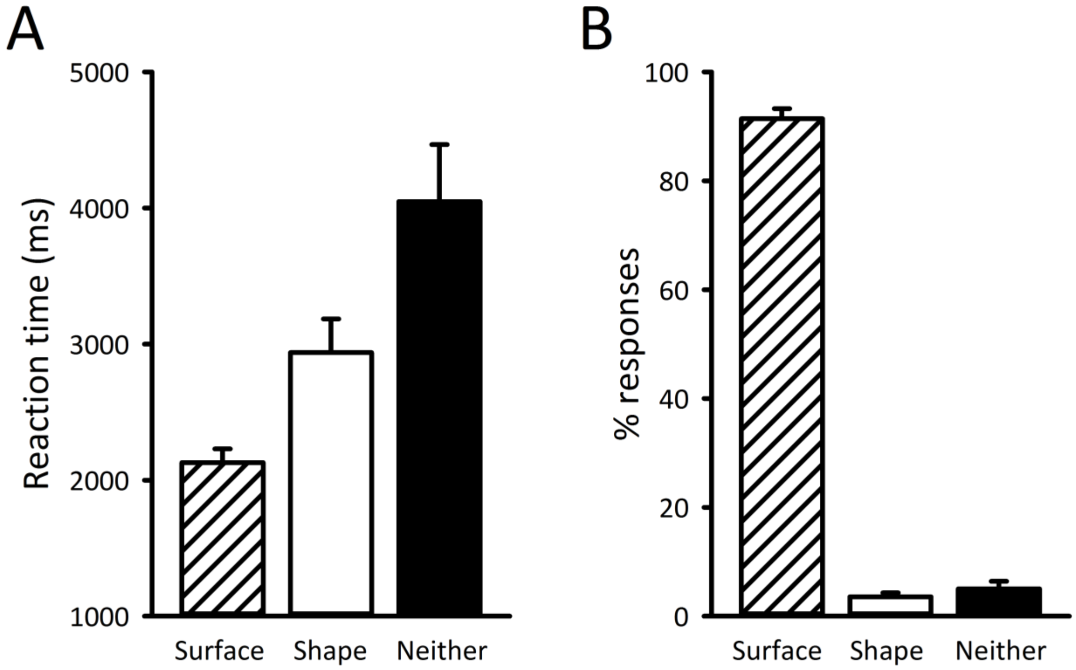


Figure 4 Performance on the behavioural recognition task. Participants were asked to perform an 8AFC for hybrid familiar faces that contained the 2d shape (fiducial positions) from one identity and the surface properties of another identity. The data show that participants used the surface properties more than the shape properties in their judgements of facial identity.

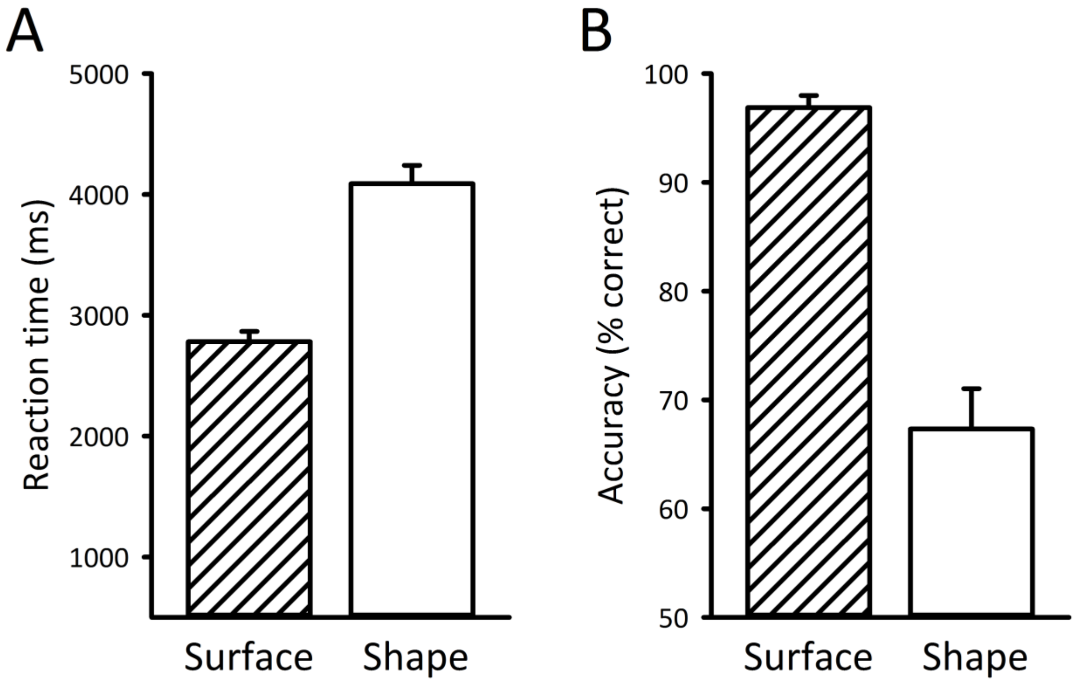


Figure 5 Performance on the behavioural matching task (see Fig. 2). Participants had (A) faster reaction times and were (B) more accurate when judgements had to be made using the surface properties compared to the shape properties of the face.

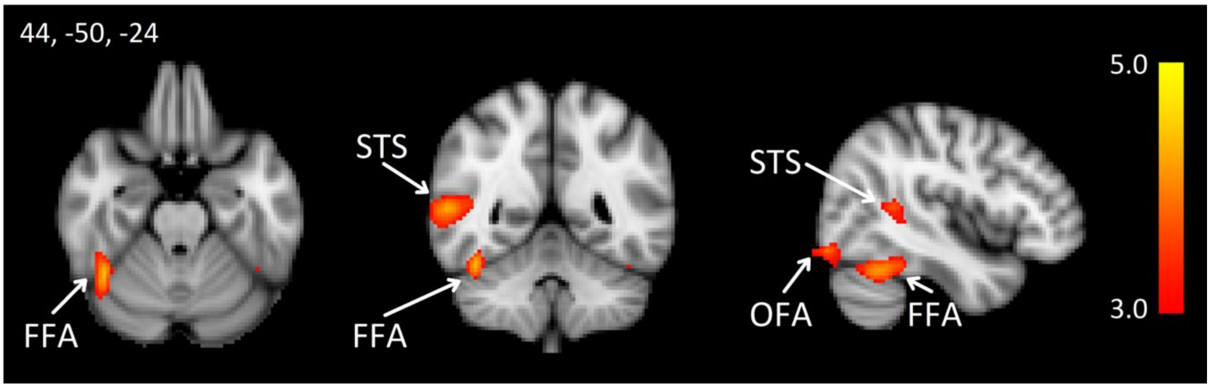


Figure 6 Average location of face selective regions. Regions of interest were defined at the individual level from an independent functional localiser scan. Images are shown in radiological convention and coordinates are given in MNI space (mm). FFA: fusiform face area, OFA: occipital face area, STS: posterior superior temporal sulcus.

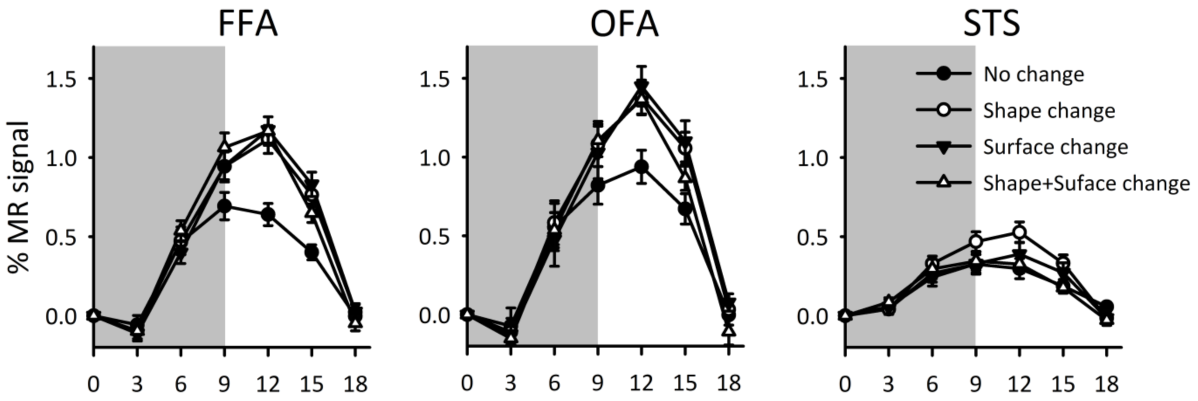
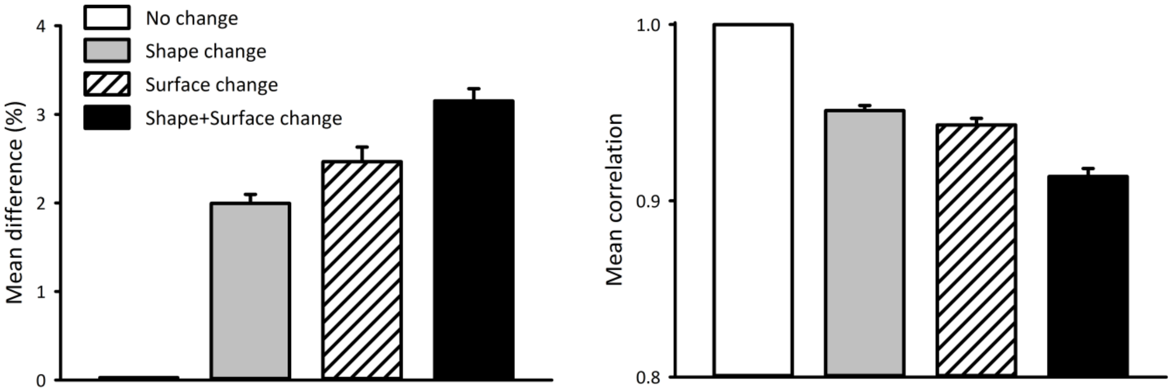
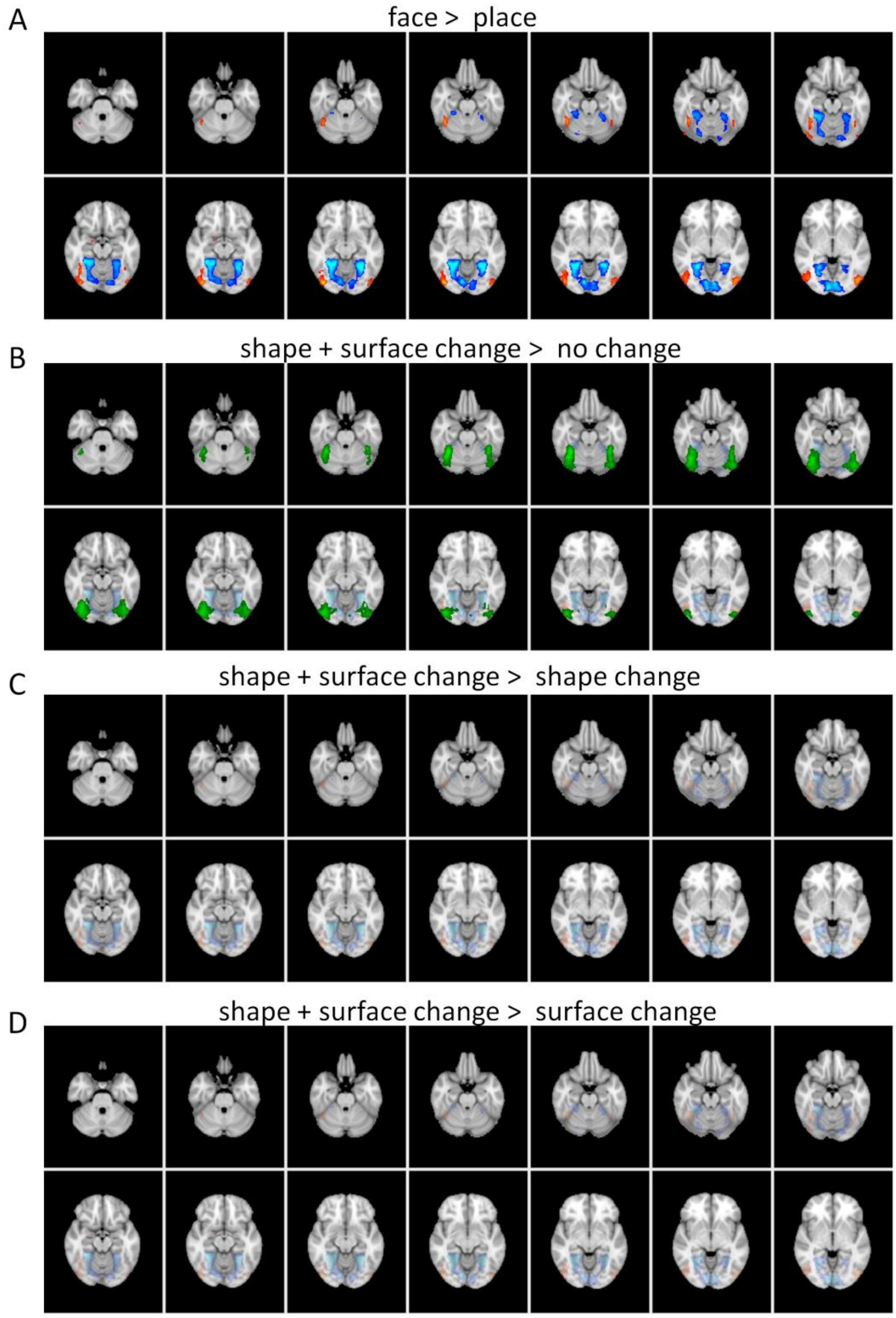


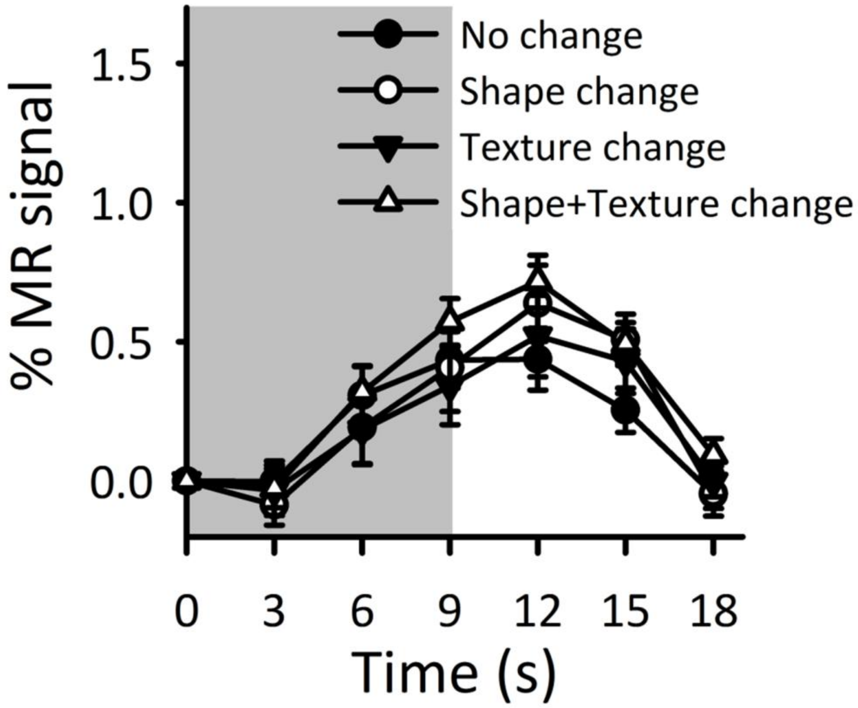
Figure 7 The average time course of neural response to each condition in the FFA, OFA and posterior STS across all participants. The shaded area indicates the duration of the stimulus block. Error bars represent ± SE.



Supplementary Figure 1 Comparison of image statistics for difference conditions. The mean change in intensity was calculated by taking the average of the absolute differences in gray value at each pixel for successive pairs of images within a block in the fMRI experiment. Correlations were calculated across corresponding pixel values for successive pairs of images within a block. Shape and surface changes had a similar effect on these low-level image measures. The largest change in low-level properties was found when both shape and surface properties changed.



Supplementary Figure 2 Group analysis of the fMRI experiment. (A) Place-selective voxels (place> face) are indicated in blue-light blue and face-selective voxels (face>place) are shown in red/yellow. (B-D) Voxels showing adaptation (lower response compared to shape+texture change) are shown in green. Significant adaptation is only evident for the contrast of shape+surface change > no change. This adaptation overlaps with face-selective regions. Place-selective and face-selective activation are shown in reduced contrast for comparison. Images are thresholded at $p<0.05$ (corrected FWE) and are superimposed on the MNI152 brain. Slices are taken from $z = -30$ (top left) to $z = -4$ (bottom right).



Supplementary Figure 3 The average time course of neural response to each condition in the Occipital Pole across all participants. The shaded area indicates the duration of the stimulus block. Error bars represent ± SE.