CORTEX 103 (2018) 199–210

Available online at www.sciencedirect.com

ScienceDirect

Journal homepage: www.elsevier.com/locate/cortex

Research report

Patterns of neural response in face regions are predicted by low-level image properties

Katja Weibert, Tessa R. Flack, Andrew W. Young and Timothy J. Andrews[*]

Department of Psychology and York Neuroimaging Centre, University of York, York, United Kingdom

ARTICLE INFO

Article history:
Received 13 September 2017
Reviewed 2 November 2017
Revised 26 January 2018
Accepted 13 March 2018
Action editor Stefan Schweinberger
Published online 23 March 2018

Keywords:
Face
Viewpoint
Identity
fMRI
MVPA

ABSTRACT

Models of face processing suggest that the neural response in different face regions is selective for higher-level attributes of the face, such as identity and expression. However, it remains unclear to what extent the response in these regions can also be explained by more basic organizing principles. Here, we used functional magnetic resonance imaging multivariate pattern analysis (fMRI-MVPA) to ask whether spatial patterns of response in the core face regions (occipital face area – OFA, fusiform face area – FFA, superior temporal sulcus – STS) can be predicted across different participants by lower level properties of the stimulus. First, we compared the neural response to face identity and viewpoint, by showing images of different identities from different viewpoints. The patterns of neural response in the core face regions were predicted by the viewpoint, but not the identity of the face. Next, we compared the neural response to viewpoint and expression, by showing images with different expressions from different viewpoints. Again, viewpoint, but not expression, predicted patterns of response in face regions. Finally, we show that the effect of viewpoint in both experiments could be explained by changes in low-level image properties. Our results suggest that a key determinant of the neural representation in these core face regions involves lower-level image properties rather than an explicit representation of higher-level attributes in the face. The advantage of a relatively image-based representation is that it can be used flexibly in the perception of faces.

1. Introduction

Recognising the identity or expression of a face is a simple and relatively effortless process for most human observers. However, the size and shape of a face image on the retina changes frequently in natural conditions as the face is seen from different viewpoints. The visual system must ignore these sources of variation due to change in viewpoint to facilitate the recognition of identity or expression, yet at the same time be able to process the implications of these viewpoint changes because of their role in social communication. Understanding the way that information about faces is represented in the brain is central to understanding the processes involved in face perception (Quiroga, 2017).

Neural models of face perception propose that different brain regions are involved in processing different information

---

[*] *Corresponding author*. Department of Psychology and York Neuroimaging Centre, University of York, York, United Kingdom.
E-mail address: timothy.andrews@york.ac.uk (T.J. Andrews).

https://doi.org/10.1016/j.cortex.2018.03.009
0010-9452/© 2018 Elsevier Ltd. All rights reserved.

from the face (Andrews & Ewbank, 2004; Bruce & Young, 2012; Davies-Thompson & Andrews, 2012; Duchaine & Yovel, 2015; Gobbini & Haxby, 2007; Haxby, Hoffman, & Gobbini, 2000; Ishai, 2008). The widely used model of Haxby et al. (2000) proposes a core system comprising regions in the occipital and temporal lobes; the occipital face area (OFA), the fusiform face area (FFA), and the superior temporal sulcus (STS). The OFA is proposed to have feedforward projections to both the STS and the FFA. The connection between the OFA and STS is thought to be important in processing dynamic changes in the face (such as expression) that are important for social interactions, whereas the connection between the OFA and FFA is important for the representation of invariant facial characteristics that are used for the recognition of identity.

Patterns of response in the core face regions to identity and expression provide some support for Haxby et al.'s (2000) model. The role of the FFA in the neural representation of identity is evident in studies that have shown distinct spatial patterns of response for faces from different identities (Anzellotti, Fairhall, & Caramazza, 2013; Axelrod & Yovel, 2015; Guntupalli, Wheeler, & Gobbini, 2017; Verosky, Todorov, & Turk-Browne, 2013; Zhang et al., 2016). Conversely, distinct spatial patterns of response in the STS have been related to different facial expressions of emotion (Said, Moore, Engell, Todorov, & Haxby, 2010; Sormaz, Watson, Smith, Young, & Andrews, 2016; Wegrzyn et al., 2015; Zhang et al., 2016). A key feature of some of these studies is that the discrimination of identity or expression is evident despite changes in the viewpoint of the image, implying some degree of view-invariance.

Other studies, however, have shown that the neural response to faces is sensitive to changes in the image, particularly to changes in viewpoint. Electrophysiological studies have shown that sensitivity to the view of a face varies across regions, with more posterior face regions being view-selective and more anterior face regions being more view-invariant (Dubois, de Berker, & Tsao, 2015; Freiwald & Tsao, 2010; Perrett et al., 1991). A similar pattern of results is evident from human neuroimaging studies. View-dependent spatial patterns of response have been found in the core face regions (OFA, FFA and STS), whereas more anterior regions have a more invariant representation (Axelrod & Yovel, 2012; Carlin, Calder, Kriegeskorte, Nili, & Rowe, 2011; Guntupalli et al., 2017; Kietzmann, Swisher, König, & Tong, 2012; Ramirez, Cichy, Allefeld, & Haynes, 2014).

Although previous studies have shown distinct patterns of response in the core face-selective regions to the relatively high-level facial characteristics of identity and expression, it is unclear to what extent more basic image properties might explain the topographic organization of these regions. Recent studies have suggested that patterns of response to images from different object categories can be predicted by the low-level properties of the image (Andrews, Watson, Rice, & Hartley, 2015; Bracci & de Beeck, 2016; Coggan, Liu, Baker, & Andrews, 2016; Rice, Watson, Hartley, & Andrews, 2014; Watson, Hartley, & Andrews, 2017; Watson, Young, & Andrews, 2016). Here, we ask whether the neural representation in core face-selective regions might also involve simpler organizing principles that are grounded in the statistical properties of the image. To achieve this, we investigated patterns of neural response in core regions that were consistent across different participants and evaluated the extent to which these consistent patterns of response were driven by high-level characteristics such as identity and expression or by relatively low-level image properties. We show that for face images with varying viewpoints, such as are commonly encountered in everyday life, much of the pattern of response in face regions of the human brain can be predicted from image properties without requiring an explicit representation of high-level characteristics such as identity and expression. We argue that this relatively image-based representation may provide a more flexible code for the perception and recognition of faces (see Chang & Tsao, 2017).

2. Methods

2.1. *Participants*

For Experiment 1, data were collected from 19 participants (7 males, mean age = 22.8 ± 1.0 years). For Experiment 2, data were collected from 24 participants (9 males, mean age 23.5 ± 2.5 years). All participants were right-handed with normal or corrected to normal vision. Written informed consent was obtained from all participants. The study was approved by the York Neuroimaging Centre Ethics Committee. Images were placed onto a 1/f amplitude noise mask to ensure all images stimulated the same amount of the visual field despite changes in orientation. Face images were back-projected onto a screen, approximately 57 cm from the participant at a height of ~8°. Stimuli were presented using PsychoPy2 (Peirce, 2007).

2.2. *Experiment 1*

Experiment 1 compared the effects of identity and viewpoint in face-selective regions, using an ambient images approach (Burton, Jenkins, & Schweinberger, 2011; Jenkins, White, Van Montfort, & Burton, 2011) in which images of each identity across each viewpoint could themselves be different in the many other ways encountered in everyday life. Many behavioural studies have shown that the ability to recognise identity across such image changes is a key characteristic of familiar face recognition (Burton, 2013; Kramer, Young, & Burton, 2018; Young & Burton, 2017). Image-invariant face recognition is also evident in event related potential (ERP) studies (Johnston, Overell, Kaufman, Robinson, & Young, 2016; Schweinberger, Pickering, Jentzsch, Burton, & Kaufmann, 2002), yet is less commonly investigated with functional magnetic resonance imaging (fMRI) (Davies-Thompson, Gouws & Andrews, 2009; Davies-Thompson, Newling & Andrews, 2013).

Fig. 1a shows the nine conditions from Experiment 1. These were based around a block design that included three different face identities (Brad Pitt (BP), David Beckham (DB), and Justin Timberlake (JT)) shown from three different viewpoints; left three-quarters (which we term −45°), front (0°), and right three-quarters (45°). Ambient image stimuli were taken from a variety of internet sources. Importantly, the use of famous identities well-known in the UK ensured participants could easily recognise the faces across different viewpoints (Davies-Thompson, Gouws, & Andrews, 2009; Young & Burton, 2017). This allowed us to select different images for each

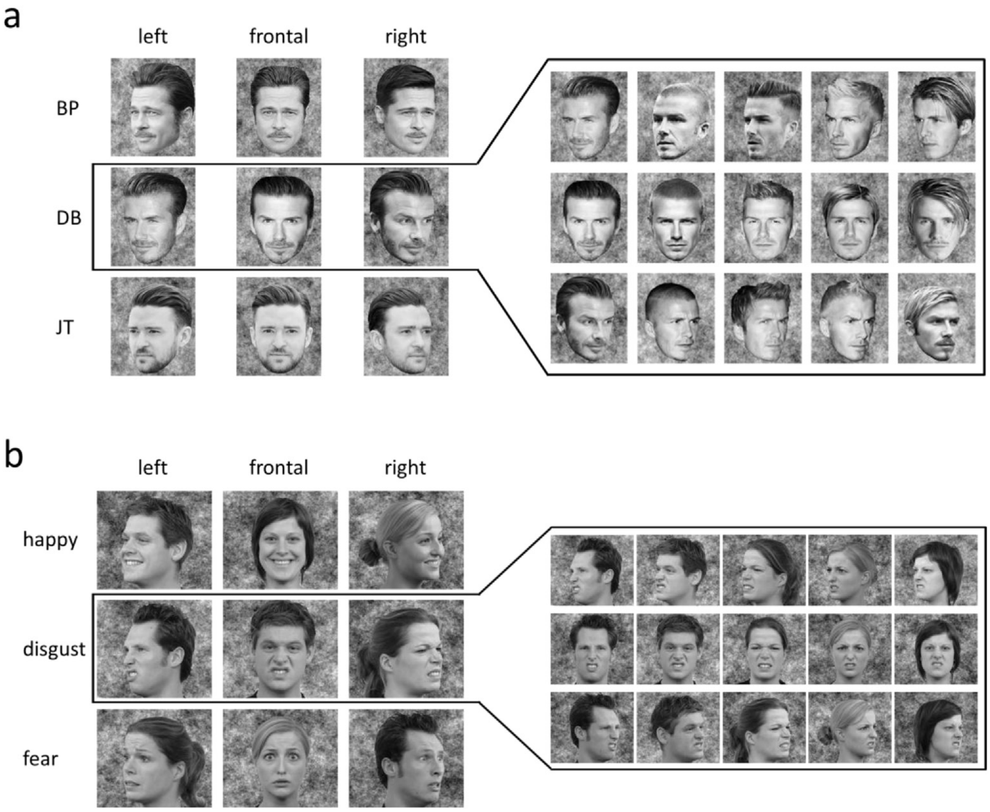

Fig. 1 – Image conditions used to create trial blocks in Experiments 1 and 2. (a) Experiment 1: There were three familiar identities (Brad Pitt (BP), David Beckham (DB), Justin Timberlake (JT)) shown at three viewpoints (left ¾, frontal, right ¾) giving a total of 9 conditions. Within each block images had the same identity and viewpoint, but varied in appearance. The images used for DB are shown on the right. (b) Experiment 2: There were three facial expressions (happy, disgust, fear) shown at three viewpoints (left ¾, frontal, right ¾) giving a total of 9 conditions. Within each block images had the same expression and viewpoint, but varied in identity. The images used for disgust are shown on the right.

combination of identity and viewpoint, so that any effects of identity could not be attributed to a specific image (Supplementary Fig. 1). It also allowed us to choose from a wide range of potential images, so we could match the selected images in appearance and in low-level properties (Table 1). The general appearance (e.g., hairstyle) of images from each identity/viewpoint combination was comparable across conditions. For example, in Supplementary Fig. 1, image 1 in each condition (row) has facial hair and styled hair. Image 2 is clean shaven with cropped hair. Images across viewpoint were also controlled. For example, compare Image 1 of JT for left, frontal and right views. [It is important to note that the order of the images within each block (condition) was randomized – the ordering here is just to highlight the way in which images were selected]. Finally, it was also important that the images within each condition varied in appearance. This ensures that any representation of identity is not image-dependent.

Table 1 – Mean image statistics (SD) of images in each condition in Experiment 1.

		Left	Frontal	Right
Grey value intensity	BP	147.17 (6.96)	149.98 (7.09)	148.00 (5.15)
	DB	148.61 (11.90)	146.73 (5.53)	148.03 (11.09)
	JT	150.63 (9.07)	148.83 (3.68)	146.99 (3.24)
RMS contrast	BP	54.60 (4.97)	49.58 (4.04)	52.75 (2.89)
	DB	48.53 (4.65)	50.97 (4.54)	49.73 (4.64)
	JT	52.08 (9.28)	55.11 (5.47)	53.90 (5.47)

Prior to scanning we ensured that participants were familiar with our stimuli. Using different images from those employed in the main experiment, we found that participants could name all identities. After scanning, we also checked that participants could recognise the identities from the images used in our experiment. Participants performed an identity matching task on the images used during the scan. On each trial they were shown two images, sequentially, and asked whether they depicted the same identity or different identities. Performance was close to ceiling (mean accuracy: 94%, ±3.2%) implying that participants were familiar with the identities shown, as this type of matching task leads to relatively poor performance with unfamiliar faces (Burton, White, & McNeill, 2010; Hancock, Bruce, & Burton, 2001).

In the fMRI experiment, five different face images were used for each condition. The five images were presented using a blocked design. Each stimulus block contained 6 images (5 unique images, 1 repeated) in quasi-random order. During the scan, a one-back task was used to maintain attention, where participants responded with a button press every time an identical image was directly repeated (one target per block). Performance was at ceiling (mean accuracy: 98.0% ± .01%). Within each stimulus block, each image was presented for

800 ms followed by a 200 ms blank screen. So, each stimulus block lasted 6s. Blocks were separated by a 9s fixation screen (a white fixation cross on a black background). Experiment 1 consisted of two runs, each with 4 repetitions of each of the nine conditions (total of 36 blocks per run) presented in a counterbalanced order.

2.3. *Experiment 2*

Experiment 2 compared neural responses in core face regions to facial expression and viewpoint across images that varied in identity, using relatively controlled images of unfamiliar faces. Face images were taken from the Radboud Faces Database (Langner et al., 2010) and depicted different people posing three different expressions (happiness, disgust, and fear) taken from three different viewpoints. Fig. 1b and Supplementary Fig. 2 show images from the nine stimulus conditions. Each stimulus block contained 5 images that varied in face identity. Each image was presented for 1000 ms followed by a 200 ms blank screen, leading to a 6s block duration. Each condition was repeated 6 times (total of 54 blocks) presented in a counterbalanced order. To maintain attention throughout the scan, participants responded with a button press every time a red spot appeared (one target per block). The red dot could appear at any location on the face. Performance was at ceiling (mean accuracy: 99.0% ± 1.9%).

2.4. *fMRI analysis*

Data from both fMRI experiments were collected using a GE 3 Tesla HD Excite MRI scanner at the York Neuroimaging Centre at the University of York. Functional data were collected using a gradient-echo echo planar imaging (EPI) sequence with a radio-frequency coil tuned to 127.4 MHz was used to acquire 38 axial slices ($TR = 3s$, $TE = 33$ ms, flip angle $= 90°$, $FOV = 288$ mm, matrix size $= 128 \times 128$, slice thickness $= 3$ mm, voxel size: $2.25 \times 2.25 \times 3$ mm). A T1-weighted structural MRI ($TR = 7.96$ ms, $TE = 3.05$ ms, $FOV = 290 \times 290$ mm, matrix size $= 256 \times 256$, voxel dimensions $= 1.13 \times 1.13$ mm, slice thickness $= 1$ mm, flip angle $= 20°$) was also collected for each participant.

First, we examined the magnitude of brain response to each condition in voxels within the region of interest (ROI). Statistical analysis of the fMRI data was carried out using FEAT version 5.0 in the FMRIB software library (FSL) toolbox (http://www.fmrib.ox.ac.uk/fsl). The first 3 volumes (9s) of each scan were removed to minimize the effects of magnetic saturation, and slice-timing correction was applied. Motion correction was followed by temporal high-pass filtering (cut-off, .01 Hz). Spatial smoothing (Gaussian) was applied at 6 mm full width at half maximum (FWHM), in line with previous studies employing smoothing in conjunction with multivariate pattern analysis (MVPA) (Op de Beeck, 2010; Watson, Hartley, & Andrews, 2014). Separately for each run, parameter estimates for each condition in the general linear model (GLM) were generated by regressing the haemodynamic response of each voxel against a box-car regressor convolved with a single-gamma haemodynamic response function. In Experiment 1, a fixed-effects analysis was used to determine the average parameter estimate for each condition across the two runs in each individual. Parameter estimates were normalised by subtracting the mean response per voxel across all experimental conditions.

The analysis was restricted to the core face-selective regions (FFA, OFA and STS). There were three important principles underlying the way in which we defined the face-selective ROIs. The first principle was that ROIs should be based on independent data. Given that we were investigating the reliability of patterns of response across individuals, it was desirable that this came from independent participants. The second principle was that the same ROIs could be used in both experiments, so that any differences in the pattern of findings across the experiments could not be due to subtle differences in the ROIs themselves. The third principle was that ROIs must be of the same size (number of voxels), to allow the MVPA analyses to have comparable potential power to detect underlying patterns of response in each region.

Face-selective masks (Supplementary Fig. 3) were therefore based on an independent localizer scan using different participants ($n = 83$), in which the response to faces (varying in identity, viewpoint and expression) was compared to the response to scrambled faces (Sormaz et al., 2016). The advantage of using a large group is that it allows us to define genuine population-level ROIs with high statistical validity. Masks in each region comprised the most significant 500 voxels (Montreal Neurological Institute – MNI space). This allowed us to compare patterns of response across the two experiments using the same ROIs based on face-selectivity determined from independent data.

The reliability of response patterns in each ROI was tested using a leave-one-participant-out (LOPO) cross-validation paradigm. First, parameter estimates were determined for each condition using a group analysis of all participants except one. This LOPO process was repeated such that every participant was left out of a group analysis once. These data were then submitted to correlation-based pattern analyses implemented using the PyMVPA toolbox (http://www.pymvpa.org/; Hanke et al., 2009). For each iteration of the LOPO cross-validation, the normalised patterns of response to each stimulus condition were correlated between the group and the left-out participant. The final correlation matrix provides a measure of the mean similarity in the pattern of response across different combinations of conditions. Prior to statistical analysis, a Fisher's Z-transform was applied.

To assess the relative contributions of identity, viewpoint, and expression to the neural response patterns a binary regressor was generated for each dimension. A value of one (yellow) was given to those elements where the relevant factor was shared and zero (red) on all other elements of the correlation matrix (Figs. 3a and 4a). Within-condition elements (white) were not included. A multiple regression was then applied to the fMRI data across participants (see Watson et al., 2016). This yielded a beta value for each regressor, which would be expected to differ significantly from zero if that regressor were able to explain a significant amount of the variance in the neural correlations. A t-contrast was used to assess the significance of the differences between the regression coefficients for each model.

To investigate the effects of low-level image properties on patterns of neural response in face-selective regions, the

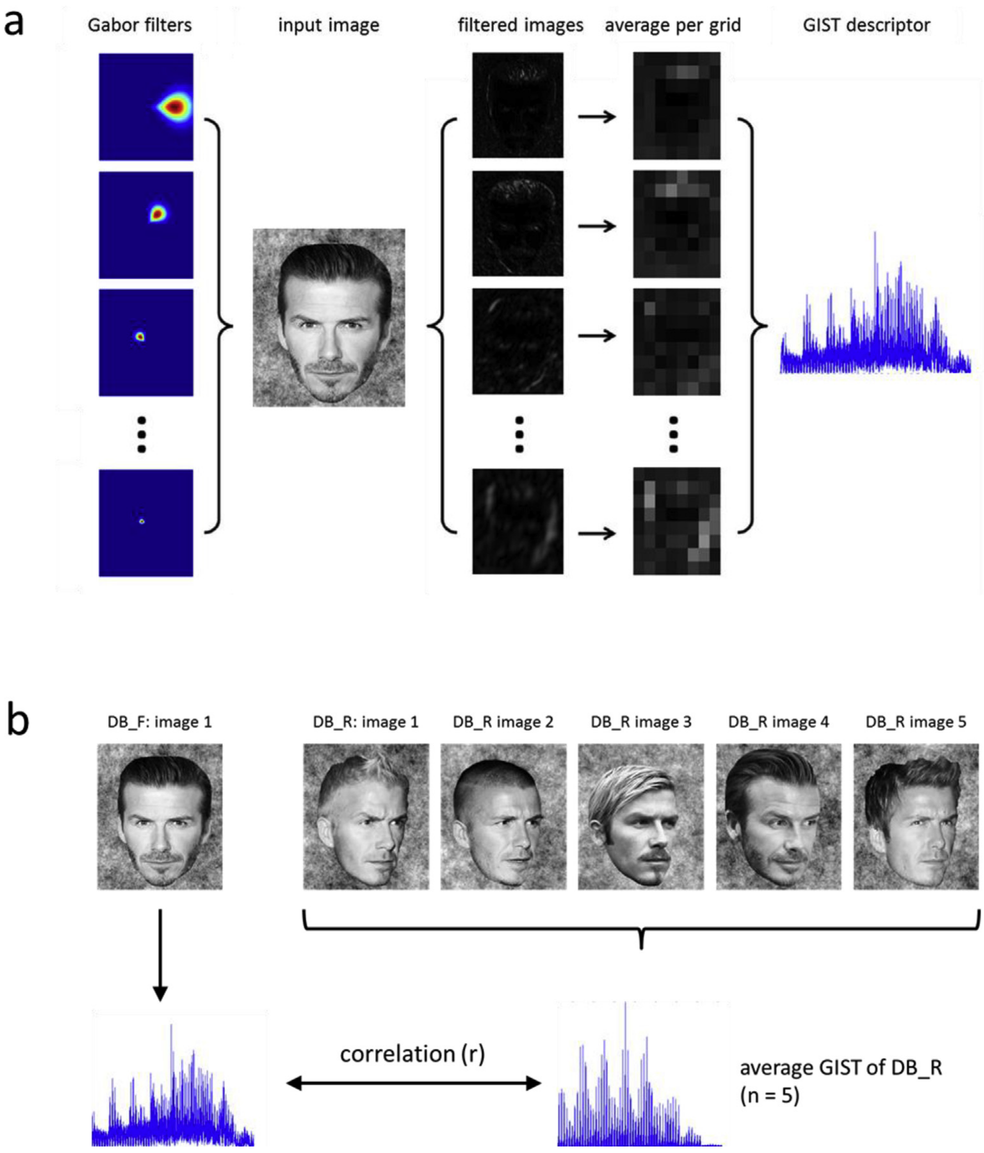

Fig. 2 – Image analysis using the GIST descriptor. (a) Schematic illustration of the GIST descriptor for an example image. A series of Gabor filters across eight orientations and four spatial frequencies are applied to the image. Each of the resulting 32 filtered images is then windowed along an 8 × 8 grid to give a final GIST descriptor of 2048 values (right). (b) To assess image similarities between conditions, the GIST descriptor of an image from one condition was correlated with the average GIST descriptor from a different condition.

image statistics of each object were computed using the GIST descriptor (http://people.csail.mit.edu/torralba/code/spatialenvelope/) (Oliva & Torralba, 2001). For each image, a vector of 2048 values was obtained by passing the image through a series of 32 Gabor filters (eight orientations at four spatial frequencies), and windowing the filtered images along a 8 × 8 grid or 64 spatial locations (Fig. 2). Each vector represents the image in terms of the output of each gabor filter at each position across the image (Rice et al., 2014; Watson et al., 2014; Watson et al., 2016). Image similarities between conditions were measured by correlating the GIST-descriptors for all 25 combinations of images. The similarity matrix of the correlation values for the GIST descriptor across all pairwise combinations of conditions was then used as a regressor in a regression analysis with the fMRI data. Again, only between-condition elements were included in the regression analysis.

3. Results

3.1. *Experiment 1: identity versus viewpoint*

To determine the relative role of identity and viewpoint on patterns of response in face-selective regions, we compared patterns of response to images of the three familiar faces shown at three different viewpoints. For each ROI, a correlation based MVPA was used to measure the similarity in the pattern of response. Fig. 3b shows the similarity in response across all combinations of the 9 conditions.

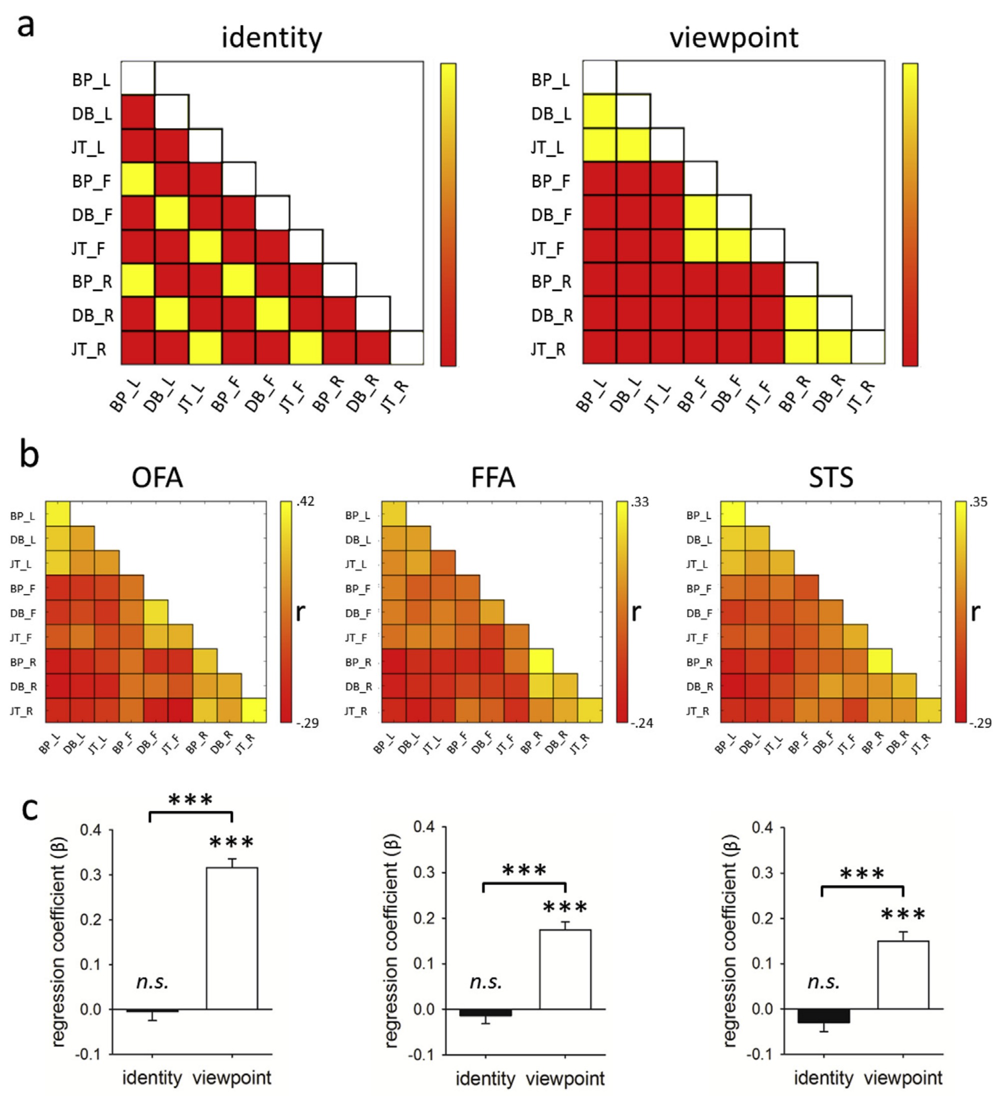

Fig. 3 – Experiment 1 – MVPA analysis of the fMRI response patterns to identity and viewpoint. (a) Idealised identity and viewpoint models used for the regression analysis of the fMRI response patterns. Binary models were defined representing the cases where the patterns of response are entirely predicted by either identity (BP = Brad Pitt, DB = David Beckham, and JT = Justin Timberlake) or viewpoint (left, frontal, or right). (b) Correlation matrix showing the similarity of fMRI response patterns within face-selective regions. (c) These matrices were compared against binary regressors for identity and viewpoint using a multiple regression analysis. The resulting beta coefficients for each regressor show that the patterns of response were explained by viewpoint, but not identity. Error bars represent 1 SE (*$p < .001$).**

First, we asked whether there were distinct patterns of response to each condition by comparing the within-condition and between-condition correlations. There was a significant effect for each region (OFA: $t(43) = 4.85$, $p < .001$, $d = 1.81$; FFA: $t(43) = 4.13$, $p < .001$, $d = 1.54$; STS: $t(43) = 4.53$, $p < .001$, $d = 1.69$). Although this implies that there are distinct patterns of response to each combination of identity and viewpoint, a stronger test is to compare within-identity correlations with corresponding between-identity correlations at the same viewpoint. We found significantly greater within-identity correlations in STS ($t(8) = 3.15$, $p = .014$, $d = .658$) and marginally different in the OFA ($t(8) = 2.29$, $p = .051$, $d = .625$) and FFA ($t(8) = 2.03$, $p = .077$, $d = .658$).

To compare the roles of identity and viewpoint in the brain responses, we used a representational similarity analysis (Kriegeskorte, Mur, & Bandettini, 2008). Model correlation matrices were generated where the patterns of brain response would be entirely predicted by identity or by viewpoint (Fig. 3a). These models were then used in a multiple regression analysis of the fMRI data (Fig. 3c). This analysis was restricted to the off-diagonal elements in the matrix. Identity did not predict patterns of response in any face-selective region (OFA: $\beta = -.01$, $p = .791$; FFA: $\beta = -.03$, $p = .119$; and STS: $\beta = -.01$, $p = .425$). In contrast, viewpoint predicted patterns of response in all face-selective regions (OFA: $\beta = .32$, $p < .001$; FFA: $\beta = .15$, $p < .001$; and STS: $\beta = .17$, $p < .001$). Moreover, the effect of

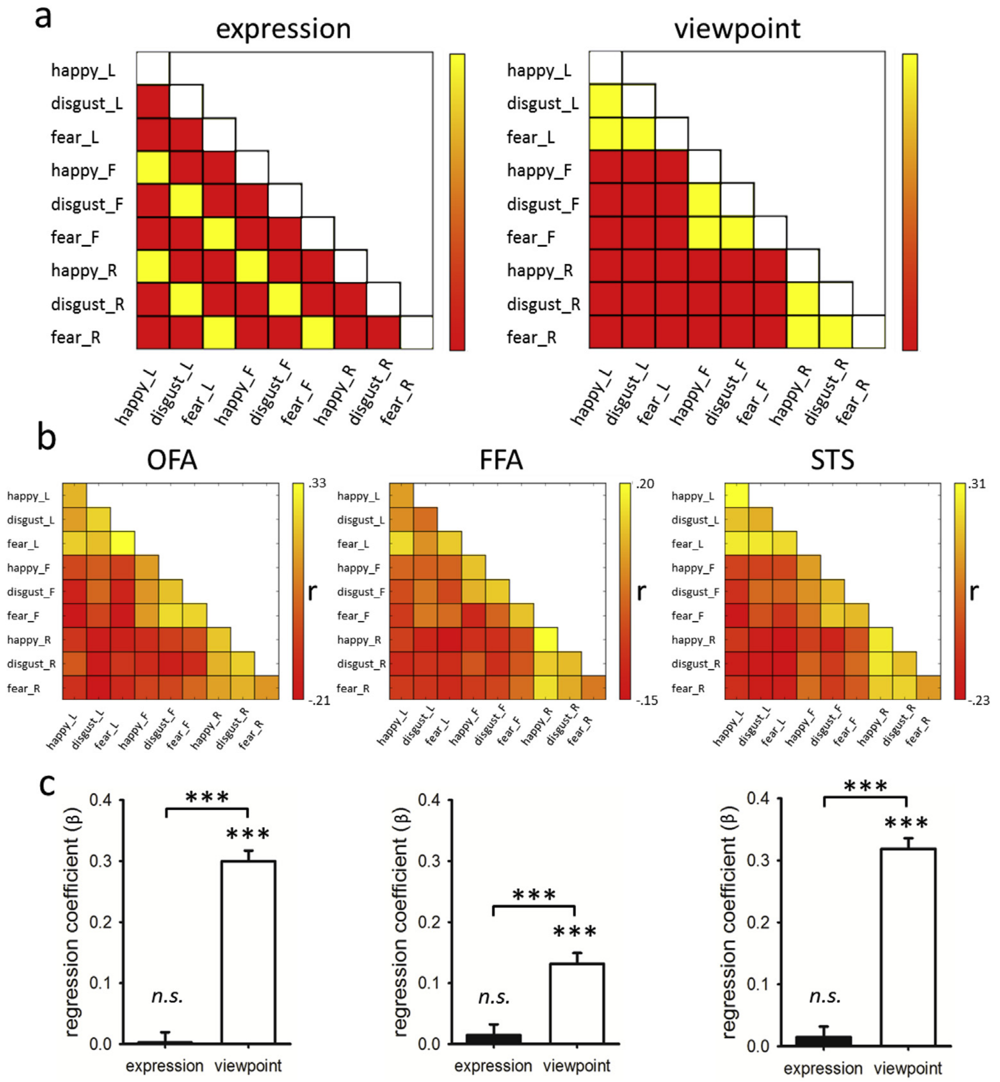

Fig. 4 – Experiment 2 – MVPA analysis of the fMRI response patterns to expression and viewpoint. (a) Idealised expression and viewpoint models used for the regression analysis of the fMRI response patterns. Binary models were defined representing the cases where the patterns of response are entirely predicted by either expression (happy, disgust, fear) or viewpoint (left, frontal, or right). (b) Correlation matrix showing the similarity of fMRI response patterns within face-selective regions. (c) These matrices were compared against binary regressors for expression and viewpoint using a multiple regression analysis. The resulting beta coefficients for each regressor show that the patterns of response were explained by viewpoint, but not expression. Error bars represent 1 SE (*$p < .001$).**

viewpoint was significantly greater than identity for all face-selective regions (OFA: t (18) = 13.98, $p < .001$; FFA: t (18) = 8.02, $p < .001$; and STS: t (18) = 9.35, $p < .001$).

3.2. *Experiment 2: expression versus viewpoint*

To determine the relative role of expression and viewpoint on patterns of response in face-selective regions, we compared patterns of response to faces with three different expressions shown at three different viewpoints. For each ROI, a correlation based MVPA was used to measure the similarity in the pattern of response. Fig. 4b shows the similarity in response to each condition.

First, we asked whether there were distinct patterns of response to each condition by comparing the within-condition and between-condition correlations. There was a significant effect for each region [OFA: t(43) = 7.95, $p < .001$, $d = -1.90$; FFA: t(43) = 4.44, $p < .001$, $d = -1.66$; STS: t(43) = 6.90, $p < .001$, $d = -1.63$]. Although this implies that there are distinct patterns of response to each combination of expression and viewpoint, a stronger test is to compare within-expression correlations with corresponding between-expression correlations at the same viewpoint. We found significantly greater within-expression correlations in OFA [t(8) = 2.53, $p = .035$, $d = .497$], but no differences in the FFA [t(8) = 1.15, $p = .282$, $d = .531$] and STS [t(8) = .00, $p = 1.00$, $d = .00$].

To compare the roles of expression and viewpoint in the brain responses, we used a representational similarity analysis. Model correlation matrices were generated based on patterns of response that could be entirely predicted by expression or by viewpoint (Fig. 4a). These models were then used in a multiple regression analysis of the fMRI data (Fig. 4c). Expression did not predict patterns of response in any face-selective region (OFA: $\beta = .00$, $p = .897$; FFA: $\beta = .01$, $p = .825$; and STS: $\beta = .01$, $p = .853$). However, just as in Experiment 1, viewpoint was a significant predictor of patterns of response in all face-selective regions (OFA: $\beta = .30$, $p < .001$; FFA: $\beta = .13$, $p < .001$; STS: $\beta = .32$, $p < .001$). The effect of viewpoint was also significantly greater than expression in each core region [OFA: $t(23) = 14.93$, $p < .001$; FFA: $t(23) = 5.71$, $p < .001$; STS: $t(23) = 15.35$, $p < .001$].

3.3. *Can low-level image properties explain patterns of response?*

Next, we asked whether the brain responses within face-selective regions could be explained by the lower-level properties of the images. To address this question, we measured the lower-level image properties of each image using the GIST descriptor (Oliva & Torralba, 2001). We compared the images within and between conditions by correlating their GIST using a Leave-One-Image-Out (LOIO) analysis separately for each experiment. This yielded two average GIST similarity matrices (one for the stimuli from each experiment; see Fig. 5a and c). Next, we used this GIST similarity matrix as a regressor to predict brain response similarity within face-selective regions. The GIST analysis predicted brain response similarity within all face-selective regions in Experiment 1 (OFA: $\beta = 1.46$, $p < .001$; FFA: $\beta = .89$, $p < .001$; STS: $\beta = 1.00$, $p < .001$) and in Experiment 2 (OFA: $\beta = 1.19$, $p < .001$; FFA: $\beta = .53$, $p < .001$; $\beta = 1.29$, $p < .001$) (Fig. 5b and d). The observed beta values were about five times higher for GIST compared to viewpoint.

4. Discussion

The aim of this study was to investigate the organizing principles underlying the topography of face-selective regions in

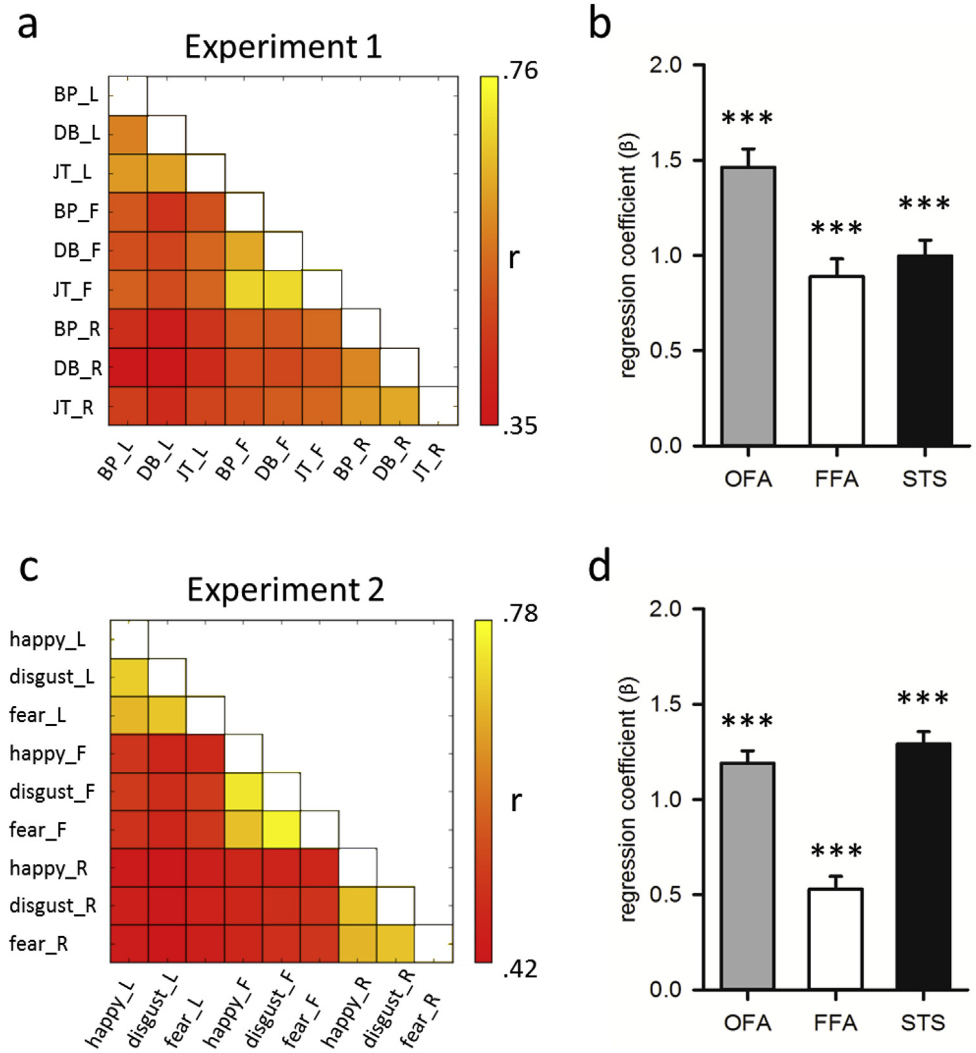

Fig. 5 – Patterns of response in face-selective regions can be predicted by the lower-level properties of the stimulus. (a), (c) Correlation of the GIST descriptors (see Fig. 2) across all conditions in Experiment 1 and 2. (b), (d) Regression analysis using the GIST similarity matrix to predict brain response in each face-selective region in Experiment 1 (Fig. 3) and Experiment 2 (Fig. 4). The resulting beta coefficients show that image properties were a strong predictor of the patterns of fMRI response. Error bars represent 1 SEM (*$p < .001$).**

the human brain. We compared the relative contributions of identity, expression and viewpoint to spatial patterns of neural response that were consistent across participants in the core face-selective regions. Despite the fact that information about the identity or expression of the face was clearly evident in each stimulus condition, we found that only viewpoint significantly predicted the spatial patterns of response.

The dominant influence of viewpoint was found across contrasting paradigms involving ambient images of highly familiar faces (Experiment 1) and relatively controlled images of unfamiliar faces (Experiment 2). The use of ambient or controlled images offers complementary perspectives (Sutherland, Rhodes, & Young, 2017; Sutherland, Young, & Rhodes, 2017) that allow us to conclude that this dominance of viewpoint is a general phenomenon that is not limited to a particular combination of stimulus conditions.

To account for this, we noted that changes in viewpoint typically result in larger changes in the image than changes in identity or expression. So, it is possible that the effect of viewpoint might largely be explained by view-contingent changes in the images themselves. Our results are consistent with this suggestion, since they show that low-level image properties were able to predict the spatial patterns of responses in all face-selective regions.

An important feature of our findings is that the spatial patterns of response to viewpoint generalized across participants. Neuroimaging studies have shown that the locations of face-selective regions in the ventral visual pathway are broadly consistent across individuals (Davies-Thompson & Andrews, 2012; Kanwisher, McDermott, & Chun, 1997). This implies that common principles may well underpin the organization of these regions. In our analysis, we compared the pattern of response in individual participants with the pattern from a group analysis in which that participant was left out (Coggan et al., 2016; Poldrack, Halchenko, & Hanson, 2009; Rice et al., 2014; Watson et al., 2014). The success of this approach shows that much of the topographic pattern of response to faces is consistent across individuals. These observations are significant in that they suggest that our findings reflect the operation of large-scale organizing principles that are consistent across different individuals.

The inability to detect patterns of response specific to identity or expression may reflect the possibility that the neural representation of these facial attributes is based on more idiosyncratic representations. Previous studies have reported distinct spatial patterns of response for faces from different identities in face-selective regions such as the FFA (Anzellotti et al., 2013; Axelrod & Yovel, 2015; Guntupalli et al., 2017; Nestor, Plaut, & Behrmann, 2011; Verosky et al., 2013; Zhang et al., 2016). However, other studies have not been able to demonstrate consistent patterns of response to identity in the FFA, but have found identity-specific patterns in more anterior regions of the temporal lobe (Kriegeskorte, Formisano, Sorger, & Goebel, 2007; Natu et al., 2010). Spatial patterns of response in the STS have also been related to particular facial expressions (Said et al., 2010; Wegrzyn et al., 2015; Zhang et al., 2016). However a more recent study did not find patterns of response for particular expressions, but did find spatial patterns that correspond to the action units from which facial expressions are derived (Srinivasan, Golomb, & Martinez, 2016). Viewpoint-dependent spatial patterns of response have been found in the core face regions (Axelrod & Yovel, 2012; Carlin et al., 2011; Guntupalli et al., 2017, Kietzmann et al., 2012; Ramirez et al., 2014).

In all these previous studies, the analysis was performed at the individual participant level. This approach is often grounded in an assumption of substantial differences between individual brains and contrasts with the across participant analysis used in the current study. Our results show that patterns of response to viewpoint, but not identity and expression are consistent across participants. While this demonstrates the important point that there is some consistent organization across the brains of different individuals, it doesn't of course rule out the likely presence of more subtle differences. Interestingly, we did find that distinct patterns of response to identity were evident when we restricted our analysis to specific viewpoints. However, a consistent pattern of response to different identities was not evident across viewpoint. These different findings can be reconciled by suggesting that, while global organizing principles underlie the spatial representation of viewpoint (as we have demonstrated here), more idiosyncratic and perhaps finer-grained representations may underly the spatial representation of identity and expression.

These results have important implications for understanding how face regions are organized. A dominant perspective on the organization of face-selective regions is that they might represent higher-level properties of the face, such as identity or expression (Duchaine & Yovel, 2015; Haxby et al., 2000; Kanwisher, 2010). However, this hypothesized organization contrasts markedly with the continuous, topographic maps found in early stages of visual processing, which are tightly linked to low-level properties of the image. Until now, it has proved difficult to explain how selectivity for these higher-level properties of faces suddenly emerges from these low-level representations. The strong linear relationship we observed between low-level image properties and the spatial patterns of response suggests instead that the topographic organization of face regions can also be in large part explained by low-level properties of the image.

Our results are consistent with a number of recent studies that have shown that lower-level image properties can be used to predict patterns of response to different object categories in the ventral visual pathway (Andrews et al., 2015; Bracci & de Beeck, 2016; Coggan et al., 2016; O'Toole, Jiang, Abdi, & Haxby, 2005; Rice et al., 2014; Watson et al., 2014, 2016, 2017). For example, the spatial pattern of response of different object categories can be predicted by the image properties of exemplars from that category (Rice et al., 2014; Watson et al., 2014, 2016, 2017). The importance of low-level properties is further supported by studies in which the pattern of neural response is minimally affected by scrambling manipulations that impair semantic properties, but preserve many of the low-level image properties of objects (Coggan et al., 2016; Watson et al., 2017). Conversely, spatial patterns of response to objects can be significantly affected by manipulations (such as changes in size) that affect the low-level image properties, but do not affect the semantic properties (Watson et al., 2016, 2017).

Our findings do not, however, imply that the representation of image properties in face-selective regions is identical to the way information is represented in early visual regions. We used a measure of visual properties (GIST – Oliva & Torralba, 2001) that was designed to capture the low-level properties of the image. Similar, image-based models such as Gabor-jet models have also been successful in explaining the perception and neural representation of faces (Lades et al., 1993; Xu, Yue, Lescroart, Biederman, & Kim, 2009; Yue, Biederman, Mangini, von der Malsburg, & Amir, 2012). However, an important property of natural images is that they contain strong statistical dependencies, such as location-specific combinations of orientation and spatial frequency corresponding to image features such as edges. Indeed, the character and extent of these statistical dependencies are likely to be diagnostic for different classes of objects (Coggan et al., 2017). Recent studies have shown that patterns of response to features of the face (mouth, eyes) are strongly dependent on the typical visual field position (de Haas et al., 2016; Henriksson, Mur, & Kriegeskorte, 2015). So, it seems likely that core face regions will represent combinations of low-level properties that are more commonly found in face images. Evidence for this is apparent in studies that show biased responses in face regions to low-level properties associated with faces, such as curvature (Wilkinson et al., 2000), horizontal information (Goffaux, Duecker, Hausfeld, Schiltz, & Goebel, 2016), and visual field position (Levy, Hasson, Avidan, Hendler, & Malach, 2001). In this study, we show that patterns of response are linked to the viewpoint of the face, which will necessarily contain combinations of low-level features that are diagnostic of different facial viewpoints. Variation in these low-level properties can be captured in Principal Components Analysis (PCA) of the shape and surface properties of faces (Burton, Kramer, Ritchie, & Jenkins, 2016; Calder, Burton, Miller, Young, & Akamatsu, 2001). Recent studies suggest that the neural representation of faces may reflect an underlying sensitivity to these shape and surface properties (Andrews, Baseler, Jenkins, Burton, & Young, 2016; Chang & Tsao, 2017; Harris, Young, & Andrews, 2014; Sormaz et al., 2016).

Although facial identity and expression are important for social interactions, faces convey many other cues to a range of characteristics that are equally as important in guiding behaviour, such as gender, race, trustworthiness and attractiveness (Bruce & Young, 2012; Todorov, 2017; Young, 2018). The interpretation of such characteristics is critically dependent on particular and often task-specific combinations of different cues (Santos & Young, 2011; Sutherland, Young, et al., 2017; Todorov, 2017; Vernon, Sutherland, Young, & Hartley, 2014; Young, 2018). In consequence, an advantage of a relatively image-based representation in core face regions may be that it can be used more flexibly in the perception of faces (Chang & Tsao, 2017; Quiroga, 2017). For example, a recent neuroimaging study has found that it is possible to read out many different attributes of the face in a task-dependent manner (Harel, Kravitz, & Baker, 2014). In this respect it is worth noting that our data do not rule out the idea that FFA contributes to the perception of invariant and STS to the perception of changeable facial properties (Haxby et al., 2000); rather, they suggest that the underlying mechanism is based on representations of the important low-level properties that underpin the perception of these attributes.

In conclusion, these results show that consistent patterns of response in core face-selective regions are dominated by changes in the viewpoint of the face rather than changes in identity or expression, and that these patterns of response are readily predicted by lower-level image properties. These findings suggest that topographic response patterns within face-selective regions involve image properties that typically co-occur with different facial cues, rather than higher level properties per se. We suggest that this relatively image-based representation in core face regions may have the advantage that it can support flexibility in the perception and recognition of different facial characteristics.

Conflict of interest

None.

Acknowledgements

KW and TRF were supported by studentships from the University of York (M0221177).

Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.cortex.2018.03.009.

REFERENCES

Andrews, T. J., Baseler, H., Jenkins, R., Burton, A. M., & Young, A. W. (2016). Contributions of feature shapes and surface cues to the recognition and neural representation of facial identity. *Cortex, 83*, 280–291.

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., 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. *Journal of Vision, 15*(7), 3–3.

Anzellotti, S., Fairhall, S. L., & Caramazza, A. (2013). Decoding representations of face identity that are tolerant to rotation. *Cerebral Cortex*, bht046.

Axelrod, V., & Yovel, G. (2012). Hierarchical processing of face viewpoint in human visual cortex. *The Journal of Neuroscience, 32*(7), 2442–2452.

Axelrod, V., & Yovel, G. (2015). Successful decoding of famous faces in the fusiform face area. *PLoS One, 10*(2), e0117126.

Bracci, S., & de Beeck, H. O. (2016). Dissociations and associations between shape and category representations in the two visual pathways. *Journal of Neuroscience, 36*(2), 432–444.

Bruce, V., & Young, A. W. (2012). *Face perception*. 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., Jenkins, R., & Schweinberger, S. R. (2011). Mental representations of familiar faces. *British Journal of Psychology, 102*, 943–958.

Burton, A. M., Kramer, R. 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., White, D., & McNeill, A. (2010). The Glasgow face matching test. *Behavior Research Methods, 42*, 286–291.

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

Carlin, J. D., Calder, A. J., Kriegeskorte, N., Nili, H., & Rowe, J. B. (2011). A head view-invariant representation of gaze direction in anterior superior temporal sulcus. *Current Biology, 21*(21), 1817–1821.

Chang, L., & Tsao, D. Y. (2017). The code for facial identity in the primate brain. *Cell, 169*(6), 1013–1028.

Coggan, D. D., Allen, L. A., Farrar, O. R., Gouws, A. D., Morland, A. B., Baker, D. H., et al. (2017). Differences in selectivity to natural images in early visual areas (V1–V3). *Scientific Reports, 7*.

Coggan, D. D., Liu, W., Baker, D. H., & Andrews, T. J. (2016). Category-selective patterns of neural response in the ventral visual pathway in the absence of categorical information. *NeuroImage, 135*, 107–114.

Davies-Thompson, J., & Andrews, T. J. (2012). Intra-and interhemispheric connectivity between face-selective regions in the human brain. *Journal of Neurophysiology, 108*(11), 3087–3095.

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*(2), 370–377.

Dubois, J., de Berker, A. O., & Tsao, D. Y. (2015). Single-unit recordings in the macaque face patch system reveal limitations of fMRI MVPA. *Journal of Neuroscience, 35*(6), 2791–2802.

Duchaine, B., & Yovel, G. (2015). A revised neural framework for face processing. *Annual Review of Vision Science, 1*, 393–416.

Freiwald, W. A., & Tsao, D. Y. (2010). Functional compartmentalization and viewpoint generalization within the macaque face-processing system. *Science, 330*(6005), 845–851.

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

Goffaux, V., Duecker, F., Hausfeld, L., Schiltz, C., & Goebel, R. (2016). Horizontal tuning for faces originates in high-level Fusiform Face Area. *Neuropsychologia, 81*, 1–11.

Guntupalli, J. S., Wheeler, K. G., & Gobbini, M. I. (2017). Disentangling the representation of identity from head view along the human face processing pathway. *Cerebral Cortex, 27*(1), 46–53.

de Haas, B., Schwarzkopf, D. S., Alvarez, I., Lawson, R. P., Henriksson, L., Kriegeskorte, N., et al. (2016). Perception and processing of faces in the human brain is tuned to typical feature locations. *Journal of Neuroscience, 36*(36), 9289–9302.

Hancock, P. J. B., Bruce, V., & Burton, A. M. (2001). Recognition of unfamiliar faces. *Trends in Cognitive Sciences, 4*, 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.

Harel, A., Kravitz, D. J., & Baker, C. I. (2014). Task context impacts visual object processing differentially across the cortex. *Proceedings of the National Academy of Sciences of the United States of America, 111*(10), E962–E971.

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.

Henriksson, L., Mur, M., & Kriegeskorte, N. (2015). Faciotopy—a face-feature map with face-like topology in the human occipital face area. *Cortex, 72*, 156–167.

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

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

Johnston, P., Overell, A., Kaufman, J., Robinson, J., & Young, A. W. (2016). Expectations about person identity modulate the face-sensitive N170. *Cortex, 85*, 54–64.

Kanwisher, N. (2010). Functional specificity in the human brain: A window into the functional architecture of the mind. *Proceedings of the National Academy of Sciences of the United States of America, 107*(25), 11163–11170.

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, 17*(11), 4302–4311.

Kietzmann, T. C., Swisher, J. D., König, P., & Tong, F. (2012). Prevalence of selectivity for mirror-symmetric views of faces in the ventral and dorsal visual pathways. *The Journal of Neuroscience, 32*(34), 11763–11772.

Kramer, R. S. S., Young, A. W., & Burton, A. M. (2018). Understanding face familiarity. *Cognition, 172*, 46–58.

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

Lades, M., Vorbruggen, J. C., Buhmann, J., Lange, J., von der Malsburg, C., Wurtz, R. P., et al. (1993). Distortion invariant object recognition in the dynamic link architecture. *IEEE Transactions on Computers, 42*(3), 300–311.

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

Levy, I., Hasson, U., Avidan, G., Hendler, T., & Malach, R. (2001). Center-periphery organization of human object areas. *Nature Neuroscience, 4*(5), 533.

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.

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.

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), 580–590.

Oliva, A., & Torralba, A. (2001). Modeling the shape of the scene: A holistic representation of the spatial envelope. *International Journal of Computer Vision, 42*(3), 145–175.

Op de Beeck, H. P. (2010). Against hyperacuity in brain reading: Spatial smoothing does not hurt multivariate fMRI analyses? *NeuroImage, 49*, 1943–1948.

Peirce, J. W. (2007). PsychoPy—psychophysics software in Python. *Journal of Neuroscience Methods, 162*(1), 8–13.

Perrett, D. I., Oram, M. W., Harries, M. H., Bevan, R., Hietanen, J. K., Benson, P. J., et al. (1991). Viewer-centred and object-centred coding of heads in the macaque temporal cortex. *Experimental Brain Research, 86*(1), 159–173.

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

Quiroga, R. Q. (2017). How do we recognize a face? *Cell, 169*(6), 975–977.

Ramírez, F. M., Cichy, R. M., Allefeld, C., & Haynes, J. D. (2014). The neural code for face orientation in the human fusiform face area. *Journal of Neuroscience, 34*(36), 12155–12167.

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, 34*(26), 8837–8844.

Said, C. P., Moore, C. D., Engell, A. D., Todorov, A., & Haxby, J. V. (2010). Distributed representations of dynamic facial expressions in the superior temporal sulcus. *Journal of Vision, 10*(5), 1–12, 11.

Santos, I. M., & Young, A. W. (2011). Inferring social attributes from different face regions: Evidence for holistic processing. *Quarterly Journal of Experimental Psychology, 64*, 751–766.

Schweinberger, S. R., Pickering, E. C., Jentzsch, I., Burton, A. M., & Kaufmann, J. M. (2002). Event-related brain potential evidence for a response of inferior temporal cortex to familiar face repetitions. *Cognitive Brain Research, 14*, 398–409.

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

Srinivasan, R., Golomb, J. D., & Martinez, A. M. (2016). A neural basis of facial action recognition in humans. *Journal of Neuroscience, 36*(16), 4434–4442.

Sutherland, C. A. M., Rhodes, G., & Young, A. W. (2017). Facial image manipulation: A tool for investigating social perception. *Social Psychological and Personality Science, 8*, 538–551.

Sutherland, C. A. M., Young, A. W., & Rhodes, G. (2017). Facial first impressions from another angle: How social judgments are influenced by changeable and invariant facial properties. *British Journal of Psychology, 108*, 397–415.

Todorov, A. (2017). *Face value: The irresistible influence of first impressions*. Princeton, NJ: Princeton University Press.

Vernon, R. J. W., Sutherland, C. A. M., Young, A. W., & Hartley, T. (2014). Modelling first impressions from highly variable facial images. *Proceedings of the National Academy of Sciences of the United States of America, 111*, E3353–E3361.

Verosky, S. C., Todorov, A., & Turk-Browne, N. B. (2013). Representations of individuals in ventral temporal cortex defined by faces and biographies. *Neuropsychologia, 51*(11), 2100–2108.

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.

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.

Watson, D. M., Hartley, T., & Andrews, T. J. (2017). Patterns of response to scrambled scenes reveal the importance of visual properties in the organization of scene-selective cortex. *Cortex, 92*, 162–174.

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

Wilkinson, F., James, T. W., Wilson, H. R., Gati, J. S., Menon, R. S., & Goodale, M. A. (2000). An fMRI study of the selective activation of human extrastriate form vision areas by radial and concentric gratings. *Current Biology, 10*(22), 1455–1458.

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. (2018). Faces, people and the brain: The 45th Sir Frederic Bartlett Lecture. *Quarterly Journal of Experimental Psychology*. https://doi.org/10.1177/1747021817740275.

Young, A. W., & Burton, A. M. (2017). Recognizing faces. *Current Directions in Psychological Science, 26*, 212–217.

Yue, X., Biederman, I., Mangini, M. C., von der Malsburg, C., & Amir, O. (2012). Predicting the psychophysical similarity of faces and non-face complex shapes by image-based measures. *Vision Research, 55*, 41–46.

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.