Received: 21 June 2018 | Revised: 4 December 2018 | Accepted: 18 December 2018

DOI: 10.1111/ejn.14327

RESEARCH REPORT

WILEY EJN European Journal of Neuroscience FENS

Selectivity for mid-level properties of faces and places in the fusiform face area and parahippocampal place area

David D. Coggan | **Daniel H. Baker** | **Timothy J. Andrews**

Department of Psychology, University of York, York, UK

Correspondence
Timothy J. Andrews, Department of Psychology, York Neuroimaging Centre, University of York, York, UK.
Email: timothy.andrews@york.ac.uk

Abstract

Regions in the ventral visual pathway, such as the fusiform face area (FFA) and parahippocampal place area (PPA) are selective for images from specific object categories. Yet images from different object categories differ in their image properties. To investigate how these image properties are represented in the FFA and PPA, we compared neural responses to locally-SCRAMBLED images (in which mid-level, spatial properties are preserved) and globally-SCRAMBLED images (in which mid-level, spatial properties are not preserved). There was a greater response in the FFA and PPA to images from the preferred CATEGORY relative to their non-preferred category for the scrambled conditions. However, there was a greater selectivity for locally-scrambled compared to globally-scrambled images. Next, we compared the magnitude of fMR-adaptation to intact and scrambled images. fMR-adaptation was evident to locally-scrambled images from the preferred category. However, there was no adaptation to globally-scrambled images from the preferred category. These results show that the selectivity to faces and places in the FFA and PPA is dependent on mid-level properties of the image that are preserved by local-scrambling.

KEYWORDS

face processing, fMRI, object vision, visual cortex

1 | INTRODUCTION

Brain regions involved in object perception form a ventral visual pathway that projects from the occipital to the temporal lobe (Milner & Goodale, 1995; Ungerleider & Mishkin, 1982). Damage to these regions often results in deficits in the recognition and perception of specific categories of objects (McNeil & Warrington, 1993; Moscovitch, Winocur, & Behrmann, 1997). Consistent with these neuropsychological studies, neuroimaging studies have shown that discrete regions of this pathway respond selectively to different categories of objects Kanwisher, 2010. For example, the fusiform face area (FFA) has been shown to respond more to images of faces than to non-face images (Allison et al., 1994; Kanwisher, McDermott, & Chun, 1997; McCarthy, Puce, Gore, & Truett, 1997), whereas the parahippocampal place area (PPA) responds more to images of places than non-place images (Aguirre & D'Esposito, 1997; Epstein & Kanwisher, 1998).

Although these findings suggest specialization within the ventral temporal lobe, the extent to which this reflects an underlying selectivity to higher-level semantic properties or lower-level image properties remains unclear. This is because images from different object categories also have different image properties (Andrews, Watson, Rice, & Hartley, 2015; Rice, Watson, Hartley, & Andrews, 2014; Watson, Hartley, & Andrews, 2014; Watson, Young, & Andrews,

Abbreviations: ANOVA, analysis of variance; FDR, false discovery rate; FFA, fusiform face area; fMRI, functional magnetic resonance imaging; OFA, occipital face area; PPA, parahippocampal place area; ROI, region of interest; *SD*, standard deviation.

Edited by Guillaume Rousselet. Reviewed by Valerie Goffaux and Reza Rajimehr.

All peer review communications can be found with the online version of the article.

2016). Evidence that category-selective regions, such as the FFA and the PPA are sensitive to low-level image properties is shown by higher responses to low-level properties (such as orientation and spatial frequency) that are typical of the preferred category (Goffaux, Duecker, Hausfeld, Schiltz, & Goebel, 2016; Nasr & Tootell, 2012; Rajimehr, Devaney, Bilenko, Young, & Tootell, 2011). Other studies have used Fourier scrambled images to investigate selectivity to low-level properties in these regions (Andrews, Clarke, Pell, & Hartley, 2010; Rossion, Hanseeuw, & Dricot, 2012). The rationale for using scrambled images is that they contain many of the image properties found in intact images, but do not convey the same categorical or semantic information, thus providing a dissociation between higher-level and lower-level information. These studies have found mixed results. One study found selectivity to scrambled houses in the PPA, but not to scrambled faces in the FFA (Andrews et al., 2010). However, another study found selectivity to scrambled faces in the FFA and other face-selective regions (Rossion et al., 2012).

The aim of this study was to use different methods of image scrambling to understand which image properties are important in the neural representations found in category-selective regions. To address this question, we compared the neural response in the FFA and PPA to intact images of faces and places with locally-scrambled and globally-scrambled versions of these images (Figure 1). Globally-scrambled images were generated using a typical Fourier-scramble, i.e. keeping the global power of each two-dimensional frequency component constant while randomizing the phase

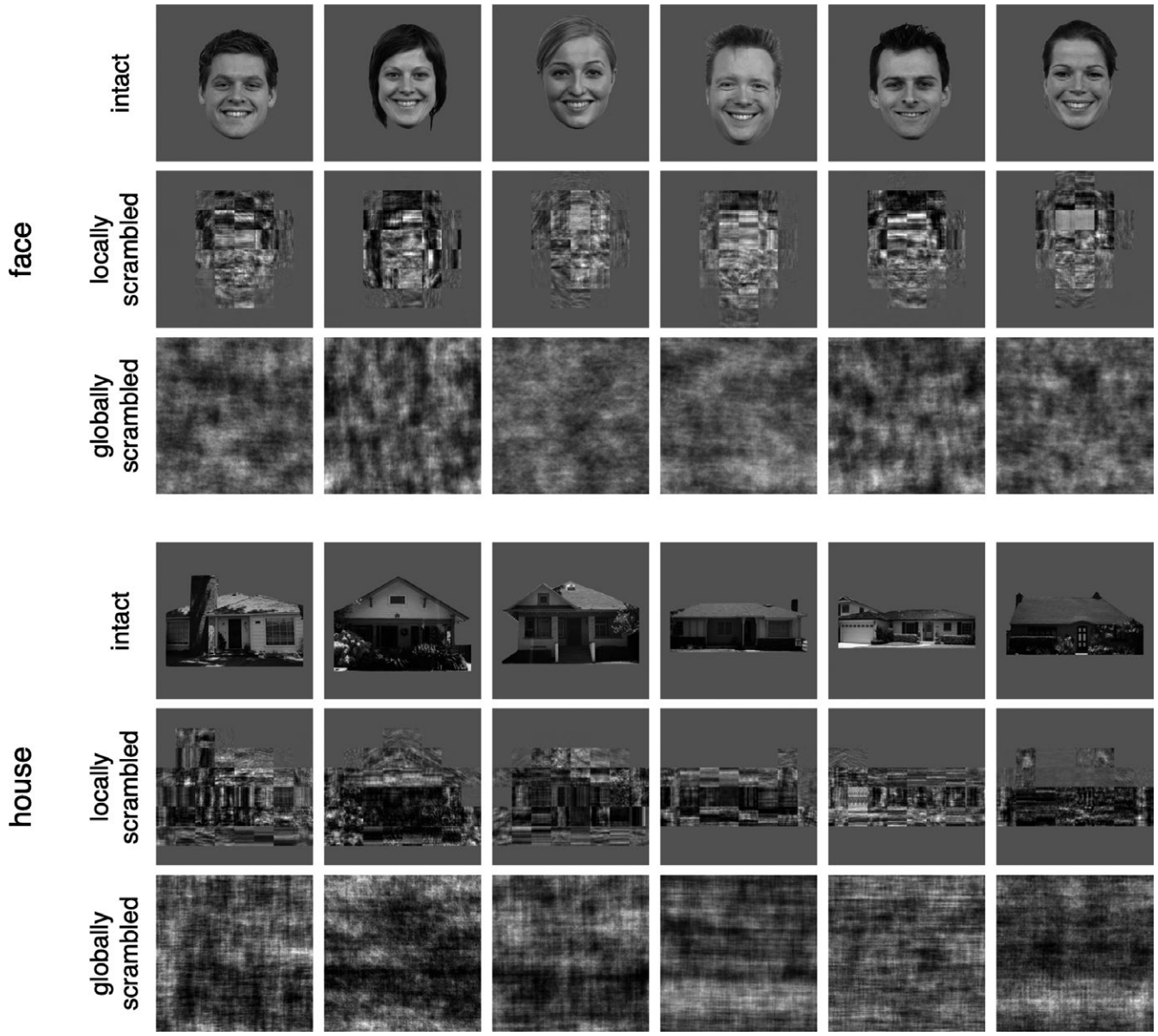

FIGURE 1 Stimulus set containing intact, locally-scrambled and globally-scrambled versions of six face and six house images

of the components. Locally-scrambled images were generated by windowing the original image into an 8 × 8 grid in image space and applying a phase-scramble to each window independently. A key difference between these methods of scrambling is that local-scrambling preserves the low-level properties in their approximate original spatial location. This preserves some of the mid-level properties (such as the spatial envelope – the region of the image taken up by the object), which may play an important role in the representation of objects. In a previous study, we found that both global- and local-SCRAMBLING rendered the images unrecognizable (Coggan, Liu, Baker, & Andrews, 2016). Despite the fact that images were unrecognizable, locally-scrambled images, but not globally-scrambled images, were found to elicit similar category-selective patterns of response across the ventral visual pathway (Coggan, Baker, & Andrews, 2016; Coggan, Liu et al., 2016). More recently, Long, Yu, and Konkle (2018) showed that images that preserve mid-level properties of objects, but were not recognizable elicited patterns of neural response to variation in animacy and real-world size that were comparable to intact objects.

Here, we ask whether there is a difference in the magnitude of response to locally-scrambled and globally-scrambled faces and places in the FFA and PPA. If selectivity is more evident to faces and places when they have been locally-scrambled compared with global-scrambled, then this shows that the magnitude of response in these regions can be explained in part by a sensitivity to the mid-level properties preserved by locally-scrambled images. If there is no difference between locally-scrambled and globally-scrambled images then this suggests that the selectivity found in scrambled images is due to the amplitude spectrum of the image. We also compared adaptation to faces and places with locally-scrambled and globally-scrambled images. The basis of fMRI adaptation is that repetition of a stimulus causes a reduction in the neural response, which leads to a lower fMRI signal (Andrews, Baseler, Jenkins, Burton, & Young, 2016; Andrews & Ewbank, 2004; Andrews et al., 2010; Avidan, Hasson, Hendler, Zohary, & Malach, 2002; Epstein, Graham, & Downing, 2003; Ewbank, Schluppeck, & Andrews, 2005; Grill-Spector & Malach, 2001; Psalta, Young, Thompson, & Andrews, 2014). Brain regions selective for a particular stimulus property will show greater adaptation (i.e. signal reduction) for a repeated stimulus than for a sequence in which the stimuli vary, whereas non-selective regions will show similar responses regardless of the sequence. The sensitivity of the neural representation can therefore be compared for different manipulations of the stimulus. If the underlying neural representation is insensitive to a particular type of manipulation in the stimulus (i.e. local- or global-scrambling), the adaptation of the fMRI signal will be similar to that produced by unchanged (in this case, intact) images.

2 | MATERIALS AND METHODS

2.1 | Participants

Twenty participants were recruited for the fMRI experiment (12 female, mean age = 29.0 years, median = 23, min = 16, max = 66, *SD* = 12.7). Participants were constituted by graduate students and staff of the Department of Psychology at the University of York, as well as members of the public responding to a participant mailing list held by York Neuroimaging Centre (YNiC). The study was approved by the YNiC Ethics Committee and adhered to the original wording of the Declaration of Helsinki. All participants reported that they had normal or corrected-to-normal vision and gave their informed, written consent.

2.2 | Stimuli

Twenty four face and 24 house images were taken from a database of objects (Rice et al., 2014). Images were grey-scale, superimposed on a mid-grey background and had a resolution of 720 × 720 pixels. Face images originated from the Radboud face database (Langner et al., 2010). Six face and six house images were selected for adaptation scans, with the remaining images used in a localizer scan. For experimental scans, two different phase-scrambled versions of each image were generated. Global-scrambling involved a typical Fourier-scramble, i.e. keeping the global power of each two-dimensional frequency component constant while randomizing the phase of the components. Local-scrambling involved windowing the original image into an 8 × 8 grid in image space and applying a phase-scramble to each window independently. Images subtended a maximum retinal angle of approximately 15° and were viewed on a screen at the rear of the scanner via a mirror placed immediately above the participant's head. Examples of the images are shown in Figure 1. The images used in this study have been validated by a behavioural study (Coggan, Liu et al., 2016) in which participants were asked to name each image. The results of the naming task show that accuracy was at ceiling for intact images. However, local- and global-scrambling renders the images unrecognizable.

2.3 | Design and procedure

There were 12 conditions in the experimental scan: two categories (face, house) × three image types (intact, locally-scrambled, globally-scrambled) × two adaptation sequences (same image, different images). The experiment was divided into three scan runs each lasting 8 min, with globally-scrambled images presented in the first run, locally-scrambled images presented in the second run and intact images presented in the third run. Scrambled images were presented before the intact images to

prevent any potential category priming effects. Images were presented in 6 s blocks. In each stimulus block, six images from a condition were presented for 800 ms with a 200 ms inter-stimulus-interval. In the "same image" condition, six identical images were shown. A fixation cross was presented for 9 s after each block. There were eight repetitions of each condition in each scan.

To maintain attention, participants were instructed to press a button when a red dot appeared on any of the images. Subjects responded with a mean response latency of 423 ms ($SEM = 10$ ms). The number of correct responses was at ceiling for intact (mean = 99.0%, $SEM = 0.4\%$), locally-scrambled (mean = 99.5%, $SEM = 0.2\%$) and globally-scrambled (mean = 100%, $SEM = 0\%$) conditions. Response latencies were entered into a one-way analysis of variance (ANOVA), which showed no effect of level of scrambling ($F_{2,34} = 0.62$, $\eta^2 = 0.03$, $p = 0.5460$).

2.4 | Data acquisition and analysis

fMRI data were acquired using a GE 3T HD Excite MRI scanner at YNiC at the University of York, fitted with an eight-channel, phased-array, head-dedicated gradient insert coil tuned to 127.4 MHz. A gradient-echo echo-planar imaging (EPI) sequence was used to collect data from 38 contiguous axial slices (TR = 3,000 ms, TE = 32.7 ms, FOV = 288 × 288 mm, matrix size = 128 × 128, slice thickness = 3 mm). The fMRI data from the localizer and experimental scans were initially analysed with FEAT v5.98 (http://www.fmrib.ox.ac.uk/fsl). In all scans, the initial 9 s of data were removed to reduce the effects of magnetic saturation. Motion correction (MCFLIRT, FSL) and slice-timing correction were applied followed by temporal high-pass filtering (Gaussian-weighted least-squares straight line fitting, $\sigma = 50$ s). Gaussian spatial smoothing was applied at 6 mm FWHM.

A localizer scan was performed after the experimental scan to localize the FFA and PPA in each individual. This involved a block-design paradigm with the same temporal parameters as the adaptation scans. Intact faces and houses were presented in alternate blocks, with six repetitions of each category. Images were different to those used in the adaptation experiment. Face- and place-selective voxels were identified using face > house and house > face contrasts respectively. The resulting statistical maps were thresholded at $z > 2.3$. Within the anatomical location of the FFA and PPA, a flood-filling algorithm was used to define 100 spatially contiguous voxels in each hemisphere. The voxel with the highest z-score for each contrast was located. Then, voxels contiguous to that voxel with the highest z-score were iteratively added to generate a progressively larger mask. This process continued until 100 voxels had been reached or there were no more significant contiguous voxels. It was not possible to identify the FFA and PPA in three participants (three males aged 66, 33 and 24), so they were removed from further analyses. The final sample consisted of 17 subjects (12 female, mean age = 26.9 years, median = 23, min = 16, max = 56, $SD = 9.9$). The average location of the FFA and PPA across participants is shown in Table 1.

2.5 | Experimental scan

To compare the magnitude of response to each condition, parameter estimates were generated by regressing the haemodynamic response of each voxel against a boxcar function convolved with a single-gamma HRF. Responses from each voxel were averaged within each region of interest (ROI) and converted to percent signal change. A repeated-measures analysis of variance (ANOVA) was then used to determine the effect of ROI (FFA, PPA), Image Type (intact, locally-scrambled, globally-scrambled), Adaptation (same, different) and Preferred Category (FFA: preferred = face, non-preferred = house; PPA: preferred = house, non-preferred = face). An FDR correction for multiple comparisons Benjamini & Hochberg, 1995 was applied to all post-hoc, pairwise comparisons. All comparisons were two-tailed. FSL's featquery was used to obtain signal change estimates in each ROI. From there, the ANOVA, post-hoc tests and plotting were all performed using R (https://www.r-project.org). The R code and signal change estimates are available at https://github.com/ddcoggan/p004. Statistical analyses were performed on the mean values across participants.

3 | RESULTS

First, the selectivity for the preferred object category (faces for FFA, houses for PPA) was measured with intact,

TABLE 1 Group means ($N = 17$) and standard deviations (in parentheses) for peak coordinates, number of voxels and average Z-score for each ROI in the localization contrast

		Peak coordinates				
ROI	**Hemisphere**	***x***	***y***	***z***	**Voxels**	***Z***
FFA	Left	−42 (4.3)	−58 (1.0)	−23 (5.4)	88 (22)	4.0 (1.0)
	Right	41 (4.2)	−53 (7.3)	−25 (4.8)	87 (15)	4.4 (1.2)
PPA	Left	−27 (3.5)	−54 (5.7)	−14 (2.7)	95 (21)	4.8 (1.2)
	Right	28 (3.2)	−55 (9.0)	−14 (5.1)	99 (5)	5.2 (1.3)

Note. ROIs were transformed from individual-space into MNI152 2 mm space for the purpose of this table.

locally-scrambled and globally-scrambled images. The magnitude of response to intact and scrambled faces and houses to the preferred and non-preferred categories is shown in Figure 2. There were main effects of Preferred Category ($F_{1,16} = 384.41$, partial $\eta^2 = 0.96$, $p < 0.0001$) and Image Type ($F_{2,32} = 42.48$, partial $\eta^2 = 0.73$, $p < 0.0001$). The effect of Preferred Category was due to a higher response to the preferred compared to the non-preferred stimulus with intact (t(16) = 18.94, mean difference = 0.78 [95% CI = 0.69, 0.86], Cohen's $d = 4.59$, $p < 0.0001$), locally-scrambled (t(16) = 7.55, mean difference = 0.19 [0.14, 0.25], Cohen's $d = 1.83$, $p < 0.0001$) and globally-scrambled (t(16) = 2.72, mean difference = 0.03 [0.01, 0.06], Cohen's $d = 0.66$, $p = 0.0229$) images.

There was a two-way interaction between Preferred Category and Image Type ($F_{2,32} = 189.74$, partial $\eta^2 = 0.92$, $p < 0.0001$). The interaction suggests that selectivity for the preferred category varies for different image types. To test this, we compared the difference between the preferred and non-preferred category for each image type. The difference between the preferred and non-preferred category was greater for intact images compared to both locally-scrambled (t(16) = 12.27, mean difference = 0.58 [0.48, 0.68], Cohen's $d = 2.97$, $p < 0.0001$) and globally-scrambled (t(16) = 18.14, mean difference = 0.74 [0.66, 0.83], Cohen's $d = 4.40$, $p < 0.0001$) images. However, there was also a bigger difference between the preferred and non-preferred stimulus for locally-scrambled compared to globally-scrambled images (t(16) = 5.38, mean difference = 0.16 [0.10, 0.23], Cohen's $d = 1.30$, $p < 0.0001$). There was no significant interaction between Preferred Category, Image Type and ROI ($F_{2,32} = 1.38$, partial $\eta^2 = 0.08$, $p = 0.27$). Consistent with the ROI analysis, Figure 3 shows a whole-brain analysis of the relative response to faces and houses which is similar for intact and locally-scrambled images. However, a different pattern of response is evident to the globally-scrambled images (Coggan, Baker et al., 2016; Coggan, Liu et al., 2016).

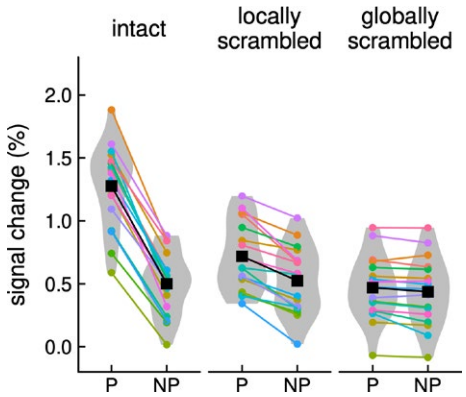

FIGURE 2 Percent signal change in category-selective regions (collapsed across fusiform face area and parahippocampal place area) to intact, locally-scrambled and globally-scrambled images of their preferred (P) and non-preferred (NP) categories. Colours reflect different subjects, with the group mean shown in black. Grey surrounds are symmetrical kernel density estimates reflecting distribution of values

Next, we asked whether the FFA and PPA would show fMR-adaptation to intact, locally-scrambled and globally-scrambled images of preferred and non-preferred categories (Figure 4). There was a main effect of Adaptation ($F_{1,16} = 36.98$, partial $\eta^2 = 0.70$, $p < 0.0001$) and a three-way interaction between Preferred Category, Image Type and Adaptation ($F_{2,32} = 11.49$, partial $\eta^2 = 0.42$, $p = 0.0002$). This indicates that the level of adaptation varied with the preferred stimulus and level of scrambling (Figure 5).

Pairwise comparisons revealed significant adaptation (different > same) to the preferred category for intact (t(16) = 8.38, mean difference = 0.28 [0.21, 0.35], Cohen's $d = 2.03$, $p < 0.0001$) and locally-scrambled images (t(16) = 5.09, mean difference = 0.12 [0.07, 0.16], Cohen's $d = 1.24$, $p = 0.0002$), but not to globally-scrambled images (t(16) = 0.44, mean difference = 0.01 [−0.03, 0.06], Cohen's $d = 0.11$, $p = 0.7199$). The magnitude of the adaptation to the preferred category was bigger for intact images compared to locally-scrambled (t(16) = 4.39, mean difference = 0.16 [0.08, 0.24], Cohen's $d = 1.06$, $p = 0.0015$) and globally-scrambled (t(16) = 6.63, mean difference = 0.27 [0.18, 0.36], Cohen's $d = 1.61$, $p < 0.0001$) images. The magnitude of the adaptation to the preferred category was bigger for locally-scrambled images compared to globally-scrambled images (t(16) = 3.02, mean difference = 0.11 [0.03, 0.18], Cohen's $d = 0.73$, $p = 0.0182$). In contrast, there was only an effect of adaptation for the non-preferred category with intact images (t(16) = 2.43, mean difference = 0.10 [0.01, 0.19], Cohen's $d = 0.59$, $p = 0.0465$) and no significant effect for locally-scrambled (t(16) = 1.60, mean difference = 0.05 [−0.02, 0.11], Cohen's $d = 0.39$, $p = 0.1673$) or globally-scrambled (t(16) = 0.20, mean difference = 0.01 [−0.05, 0.07], Cohen's $d = 0.02$, $p = 0.8622$) images. Finally, there was no significant interaction between Preferred Category, Image Type, Adaptation and ROI ($F_{2,32} = 0.90$, partial $\eta^2 = 0.05$, $p = 0.42$), again demonstrating that these effects generalize across regions.

Finally, we investigated whether the results described above were inherited from responses to the images in early visual cortex (Figure 6). To address this, we registered individual-level data into a standard space (MNI152) and restricted our analysis to a V1 mask taken from a probabilistic atlas of retinotopic regions Wang, Mruczek, Arcaro, & Kastner, 2015. A three-way repeated-measures ANOVA revealed the main effects of Image Type ($F_{2,32} = 27.6$, partial $\eta^2 = 0.63$, $p < 0.001$), Category ($F_{1,16} = 28.7$, partial $\eta^2 = 0.64$, $p < 0.001$) and Adaptation ($F_{1,16} = 25.2$, partial $\eta^2 = 0.61$, $p < 0.001$). In contrast to the responses in higher-level regions, the effect of Image Type in V1 was due to

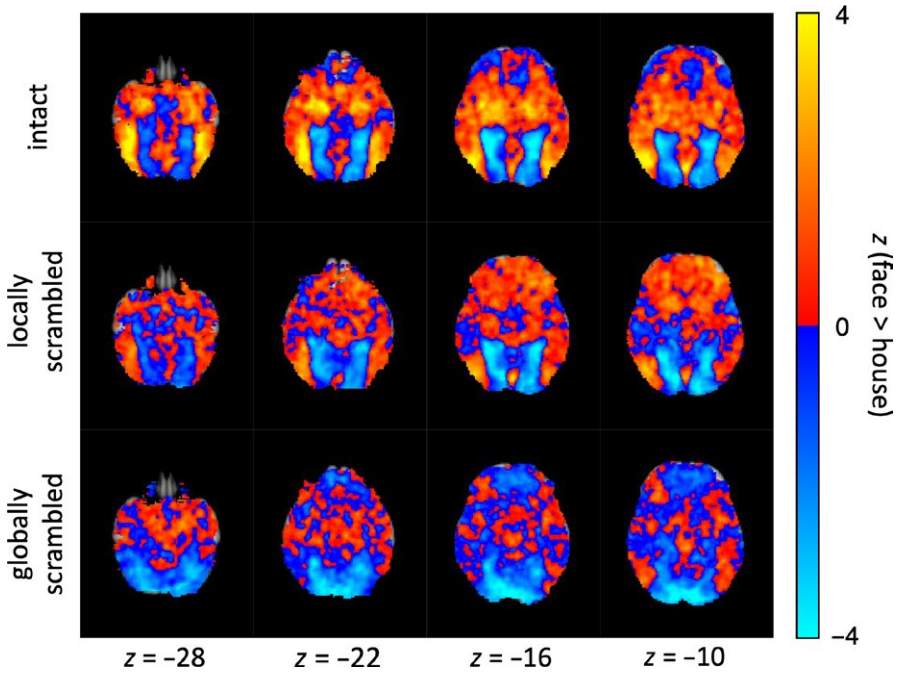

FIGURE 3 Axial slices showing group-level z statistics for a face > house contrast for each Image Type in 2 mm MNI space. For each image type, data were collapsed across "different" and "same" sequence types to form one parameter estimate per category. Red/yellow regions were more responsive to faces; blue regions were more responsive to houses

higher responses to globally-scrambled compared to both locally-scrambled ($t(16) = 2.94$, mean difference = 0.27 [0.07, 0.46], Cohen's $d = 0.71$, $p = 0.0096$) and intact ($t(16) = 6.80$, mean difference = 0.71 [0.49, 0.94], Cohen's $d = 1.65$, $p < 0.0001$) images. There was a higher response to locally-scrambled compared to intact images ($t(16) = 4.71$, mean difference = 0.45 [0.25, 0.65], Cohen's $d = 1.14$, $p = 0.0004$). The effect of category was due a higher response to houses compared to faces ($t(16) = 5.36$, mean difference = 0.17 [0.10–0.24], Cohen's $d = 1.30$, $p < 0.0001$). The effect of Adaptation was due to sequences of different images eliciting greater response than sequences of the same image. There were no significant interactions between Image Type and Adaptation ($F_{2,32} = 2.09$, partial $\eta^2 = 0.12$, $p = 0.139$), Category and Adaptation ($F_{1,16} = 3.29$, partial $\eta^2 = 0.17$, $p = 0.088$) or between Image Type, Category and Adaptation ($F_{2,32} = 0.13$, partial $\eta^2 = 0.01$, $p = 0.878$). Taken together, this shows that adaptation in V1 was not significantly different across Image Type or Category.

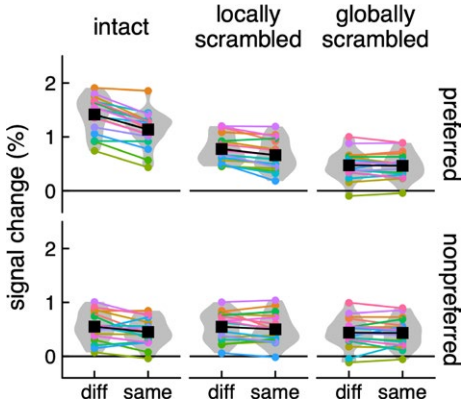

FIGURE 4 Percent signal change in category-selective regions (collapsed across fusiform face area and parahippocampal place area) in response to intact, locally-scrambled and globally-scrambled images of their preferred and non-preferred categories, presented as a sequence of different images or the same image repeated. Colours reflect different subjects, with the group mean shown in black

4 | DISCUSSION

The aim of this study was to explore the sensitivity of category-selective regions in the ventral stream to low-level image properties. To test this, neural responses to intact, locally-scrambled and globally-scrambled images of faces and houses were compared in the face-selective FFA and place-selective PPA. The rationale for using scrambled images is that they contain many of the image properties found in intact images, but do not convey the same categorical or semantic information, thus providing a dissociation between higher-level and lower-level information. The major difference between the locally-scrambled and globally-scrambled images is that some of the mid-level, spatial properties of the image are preserved in the locally-scrambled images, but not in the globally-scrambled images. The key finding from this study is that selectivity and

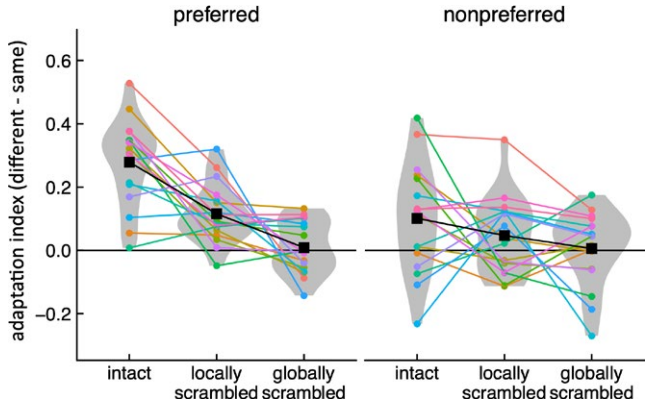

FIGURE 5 Adaptation index (different–same) in category-selective regions (collapsed across fusiform face area and parahippocampal place area) in response to intact, locally-scrambled and globally-scrambled images of their preferred and non-preferred categories. Colours reflect different subjects, with the group mean shown in black

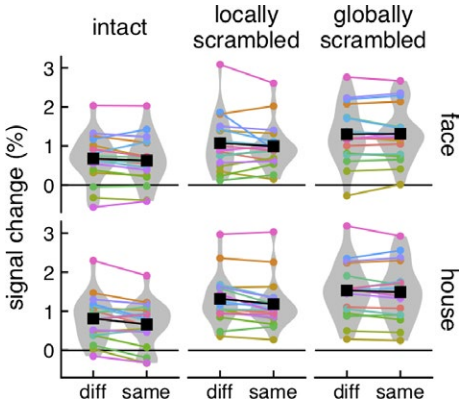

FIGURE 6 Signal change in V1 in response to sequences of different and same images for each Image Type and Category. Adaptation (different–same) changed very little across categories or image types. Colours reflect different subjects, with the group mean shown in black

adaptation to the preferred category in the FFA and PPA were greater for the locally-scrambled images compared to the globally-scrambled images.

The selectivity to locally-scrambled objects in the FFA and PPA complement recent multivariate studies in which the pattern of neural response in the ventral visual pathway to intact images is shown to be more similar to locally-scrambled compared to globally-scrambled images Coggan, Baker et al., 2016; Coggan, Liu et al., 2016. These results show that the selectivity to faces and places in the FFA and PPA is to some extent determined by the image properties that are preserved by local-scrambling, such as the spatial envelope of the image. This fits with previous studies that have demonstrated selectivity in higher-level regions of the ventral stream to spatial properties of the image (Bracci & Op de Beeck, 2016; Cichy et al., 2013; Golomb & Kanwisher, 2012; Levy, Hasson, Avidan, Hendler, & Malach, 2001; Ponce, Sturmfels, & Trager, 2017; Watson et al., 2016). More generally, these results are consistent with previous studies that have shown that patterns of response in high-level visual regions are sensitive to image properties (Rice et al., 2014; Watson et al., 2016; Xu, Yue, Lescroart, Biederman, & Kim, 2009; Yue, Tjan, & Biederman, 2006). 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 occipital face area (OFA) and FFA (Xu et al., 2009; Yue et al., 2006).

One potential issue with the interpretation of these results is that local-scrambling affects the amplitude spectrum of the image. Applying a grid to the image prior to scrambling introduces horizontal and vertical edges to the image. These are known to affect the processing of faces and scenes (Dakin & Watt, 2009; Goffaux & Dakin, 2010; Hansen, Essock, Zheng, & Deford, 2003; Pachai, Sekuler, & Bennett, 2013). However, the key statistics we report are the contrasts between locally-scrambled conditions (e.g. preferred vs. non-preferred or different vs. same). If the artefacts created by local-scrambling do have some effect, this effect should cancel in these comparisons. On the other hand, if local-scrambling has a disproportionate effect on face or scene processing then this should be reflected in an interaction between preferred Category, Image Type and Region and this is not evident in our analysis. This suggests that the selectivity to the preferred category in locally-scrambled images is not an artefact of the image generation.

The selectivity and adaptation for the preferred category was greater in intact images compared to locally-scrambled images. One possible explanation for this difference is that the neural representation is selective for higher-level, semantic information about the image that is only available from the intact images (Kanwisher, 2010). However, an alternative possibility is that unexplained variance might reflect image properties disrupted by the scrambling process. An important feature of intact images is the strong statistical dependencies between features, such as specific combinations of spatial frequency and orientation at particular locations in the image. Indeed, the behavioural sensitivity to these regularities in intact objects suggests that they play an important role in differentiating between different classes of images (Loschky & Larson, 2010; Loschky et al., 2007). It is possible that these mid-level properties also contribute to the patterns of response in category-selective regions. Recently, images that preserve mid-level properties of objects, but were not recognizable, elicited similar patterns of neural response

to animacy and real-world size found for intact objects (Long et al., 2018). When evaluating these possibilities, it is important to recognize that high-level, mid-level and low-level contributions to the observed representational structure are not mutually exclusive. The extraction of any high-level features depends on the availability of relevant mid-level features and the extraction of any mid-level features depends on the availability of relevant low-level features.

Although adaptation was most evident to images from the preferred category, we also found significant adaptation to intact images from the non-preferred category. This finding is relevant to recent accounts that have attempted to explain the organization of the occipital–temporal cortex (Behrmann & Plaut, 2013). The domain-specific approach suggests that discrete cortical regions are selective for the processing of specific categories of objects (Kanwisher, 2010). In contrast, the domain-general approach suggests a distributed and overlapping representation of visual information along the ventral visual pathway (Haxby et al., 2001). Neuropsychological studies are often used as evidence for a domain-specific representation (McNeil & Warrington, 1993; Moscovitch et al., 1997). Our finding of adaptation to the non-preferred object category is consistent with previous studies that have found adaptation to non-preferred stimuli in category-selective regions (Ewbank et al., 2005). This suggests that the representation of objects and places is not restricted to those regions that respond maximally, but is distributed across the ventral visual pathway. However, it is also important to reiterate that the magnitude of the adaptation was much greater for the preferred compared to the non-preferred category.

Finally, we asked whether the pattern of results found in higher-level regions reflected responses at early stages of processing. The pattern of response in V1 was quite different to that found in the category-selective regions. We found the highest response to globally-scrambled images, which presumably reflects differences in the amount of the visual field that was stimulated (Grill-Spector, Kushnir, Edelman, Itzchak, & Malach, 1998). Although there was adaptation to repetitions of the same object, this was not significantly different for intact or scrambled images. This demonstrates that the responses in the FFA and PPA are emergent properties of the visual system.

In conclusion, we have shown that the selectivity to objects in category-selective regions is also evident to locally-scrambled objects in which the spatial properties of the image are preserved, but is less evident to globally-scrambled objects in which spatial properties are disrupted. This suggests that the neural representation in high-level visual cortex is particularly sensitive to the spatial properties of the stimulus. Nevertheless, it is clear that the selectivity and adaptation demonstrated by scrambled images does not explain all of the variance in the intact images. Further studies will be needed to understand the relative role of image properties not preserved by scrambling and higher-level semantic properties in the neural representation of category-selective regions.

ACKNOWLEDGEMENTS

D.D.C was supported by a studentship from the University of York.

CONFLICT OF INTEREST

There are no conflicts of interest.

DATA ACCESSIBILITY

The R code and secondary level analyses are available at https://github.com/ddcoggan/p004. Full MRI data is available on request.

AUTHOR CONTRIBUTIONS

D.D.C., D.H.B. and T.J.A. designed and performed research; D.D.C., D.H.B. and T.J.A. wrote the paper.

ORCID

Daniel H. Baker https://orcid.org/0000-0002-0161-443X
Timothy J. Andrews https://orcid.org/0000-0001-8255-9120

REFERENCES

Aguirre, G. K., & D'Esposito, M. (1997). Environmental knowledge is subserved by separable dorsal/ventral neural areas. *Journal of Neuroscience*, *17*, 2512–2518. https://doi.org/10.1523/JNEUROSCI.17-07-02512.1997

Allison, T., Ginter, H., McCarthy, G., Nobre, A., Puce, A., Luby, M., & Spencer, D. (1994). Face recognition in human extrastriate cortex. *Journal of Neurophysiology*, *71*, 821–825. https://doi.org/10.1152/jn.1994.71.2.821

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

Andrews, T. J., Clarke, A., Pell, P., & Hartley, T. (2010). Selectivity for low-level features of objects in the human ventral stream. *NeuroImage*, *49*, 703–711. https://doi.org/10.1016/j.neuroimage.2009.08.046

Andrews, T. J., & Ewbank, M. P. (2004). Distinct representations for facial identity and changeable aspects of faces in the human temporal lobe. *NeuroImage*, *23*, 905–913. https://doi.org/10.1016/j.neuroimage.2004.07.060

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*, 3. https://doi.org/10.1167/15.7.3

Avidan, G., Hasson, U., Hendler, T., Zohary, E., & Malach, R. (2002). Analysis of the neuronal selectivity underlying low fMRI signals. *Current Biology*, *12*, 964–972. https://doi.org/10.1016/S0960-9822(02)00872-2

Behrmann, M., & Plaut, D. C. (2013). Distributed circuits, not circumscribed centers, mediate visual recognition. *Trends in Cognitive Sciences*, *17*, 210–219. https://doi.org/10.1016/j.tics.2013.03.007

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B: Methodological*, *57*, 289–300.

Bracci, S., & Op de Beeck, H. (2016). Dissociations and associations between shape and category representations in the two visual pathways. *Journal of Neuroscience*, *36*, 432–444. https://doi.org/10.1523/JNEUROSCI.2314-15.2016

Cichy, R. M., Sterzer, P., Heinzle, J., Elliott, L. T., Ramirez, F., & Haynes, J. D. (2013). Probing principles of large-scale object representation: category preference and location encoding. *Human Brain Mapping*, *34*, 1636–1651. https://doi.org/10.1002/hbm.22020

Coggan, D. D., Baker, D. H., & Andrews, T. J. (2016). The role of visual and semantic properties in the emergence of category-specific patterns of neural response in the human brain. *eNeuro*, *3*, 1–10. ENEURO.0158-16.2016.

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. https://doi.org/10.1016/j.neuroimage.2016.04.060

Dakin, S. C., & Watt, R. J. (2009). Biological "bar codes" in human faces Steven. *Journal of Vision*, *9*, 1–10.

Epstein, R., Graham, K. S., & Downing, P. E. (2003). Viewpoint-specific scene representations in human parahippocampal cortex. *Neuron*, *37*, 865–876. https://doi.org/10.1016/S0896-6273(03)00117-X

Epstein, R., & Kanwisher, N. (1998). A cortical representation of the local visual environment. *Nature*, *392*, 598–601. https://doi.org/10.1038/33402

Ewbank, M. P., Schluppeck, D., & Andrews, T. J. (2005). fMR-adaptation reveals a distributed representation of inanimate objects and places in human visual cortex. *NeuroImage*, *28*, 268–279. https://doi.org/10.1016/j.neuroimage.2005.06.036

Goffaux, V., & Dakin, S. C. (2010). Horizontal information drives the behavioral signatures of face processing. *Frontiers in Psychology*, *1*, 1–14.

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. https://doi.org/10.1016/j.neuropsychologia.2015.12.004

Golomb, J. D., & Kanwisher, N. (2012). Higher level visual cortex represents retinotopic, not spatiotopic, object location. *Cerebral Cortex*, *22*, 2794–2810. https://doi.org/10.1093/cercor/bhr357

Grill-Spector, K., Kushnir, T., Edelman, S., Itzchak, Y., & Malach, R. (1998). Cue-invariant activation in object-related areas of the human occipital lobe. *Neuron*, *21*, 191–202. https://doi.org/10.1016/S0896-6273(00)80526-7

Grill-Spector, K., & Malach, R. (2001). fMR-adaptation: a tool for studying the functional properties of human cortical neurons. *Acta Psychologica*, *107*, 293–321. https://doi.org/10.1016/S0001-6918(01)00019-1

Hansen, B. C., Essock, E. A., Zheng, Y., & Deford, J. K. (2003). Perceptual anisotropies in visual processing and their relation to natural image statistics. *Network Computation in Neural Systems*, *14*, 501–526. https://doi.org/10.1088/0954-898X_14_3_307

Haxby, J. V., Gobbini, M., Furey, M., Ishai, A., Schouten, J., & Pietrini, P. (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science*, *293*, 2425–2430. https://doi.org/10.1126/science.1063736

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*, 11163–11170. https://doi.org/10.1073/pnas.1005062107

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*, 4302–4311. https://doi.org/10.1523/JNEUROSCI.17-11-04302.1997

Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D. H. J., Hawk, S. T., & van Knippenberg, A. (2010). Presentation and validation of the Radboud Faces Database. *Cognition and Emotion*, *24*, 1377–1388. https://doi.org/10.1080/02699930903485076

Levy, I., Hasson, U., Avidan, G., Hendler, T., & Malach, R. (2001). Center – periphery organization of human object areas. *Nature Neuroscience*, *4*, 533–539. https://doi.org/10.1038/87490

Long, B., Yu, C.-P., & Konkle, T. (2018). Mid-level visual features underlie the high-level categorical organization of the ventral stream. *Proceedings of the National Academy of Sciences of the United States of America*, *115*, E9015–E9024. https://doi.org/10.1073/pnas.1719616115

Loschky, L. C., & Larson, A. M. (2010). The natural/man-made distinction is made before basic-level distinctions in scene gist processing. *Visual Cognition*, *18*, 513–536. https://doi.org/10.1080/13506280902937606

Loschky, L. C., Sethi, A., Simons, D. J., Pydimarri, T. N., Ochs, D., & Corbeille, J. L. (2007). The importance of information localization in scene gist recognition. *Journal of Experimental Psychology: Human Perception and Performance*, *33*, 1431–1450.

McCarthy, G., Puce, A., Gore, J. C., & Truett, A. (1997). Face-specific processing in the human fusiform Gyrus. *Journal of Cognitive Neuroscience*, *9*, 605–610. https://doi.org/10.1162/jocn.1997.9.5.605

McNeil, J. E., & Warrington, E. K. (1993). Prosopagnosia: a face-specific disorder. *The Quarterly Journal of Experimental Psychology. A, Human Experimental Psychology*, *46*, 1–10. https://doi.org/10.1080/14640749308401064

Milner, A. D., & Goodale, M. A. (1995). *The Visual Brain in Action*. Oxford: Oxford University Press.

Moscovitch, M., Winocur, G., & Behrmann, M. (1997). What is special about face recognition? Nineteen experiments on a person with visual object agnosia and dyslexia but normal face recognition. *Journal of Cognitive Neuroscience*, *9*, 555–604. https://doi.org/10.1162/jocn.1997.9.5.555

Nasr, S., & Tootell, R. B. H. (2012). A cardinal orientation bias in scene-selective visual cortex. *Journal of Neuroscience*, *32*, 14921–14926. https://doi.org/10.1523/JNEUROSCI.2036-12.2012

Pachai, M. V., Sekuler, A. B., & Bennett, P. J. (2013). Sensitivity to information conveyed by horizontal contours is correlated with face identification accuracy. *Frontiers in Psychology*, *4*, 1–9.

Ponce, J., Sturmfels, B., & Trager, M. (2017). Congruences and concurrent lines in multi-view geometry. *Advances in Applied Mathematics*, *88*, 62–91. https://doi.org/10.1016/j.aam.2017.01.001

Psalta, L., Young, A. W., Thompson, P., & Andrews, T. J. (2014). The thatcher illusion reveals orientation dependence in brain regions

involved in processing facial expressions. *Psychological Science*, *25*, 128–136. https://doi.org/10.1177/0956797613501521

Rajimehr, R., Devaney, K. J., Bilenko, N. Y., Young, J. C., & Tootell, R. B. H. (2011). The "parahippocampal place area" responds preferentially to high spatial frequencies in humans and monkeys. *PLoS Biology*, *9*, e1000608.

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. https://doi.org/10.1523/JNEUROSCI.5265-13.2014

Rossion, B., Hanseeuw, B., & Dricot, L. (2012). Defining face perception areas in the human brain: a large-scale factorial fMRI face localizer analysis. *Brain and Cognition*, *79*, 138–157. https://doi.org/10.1016/j.bandc.2012.01.001

Ungerleider, L. G., & Mishkin, M. (1982). Two cortical visual systems. In D. J. Ingle, M. A. Goodale, & R. J. W. Mansfield (Eds.), *Analysis of Visual Behaviour* (pp. 549–586). Cambridge, MA: MIT Press.

Wang, L., Mruczek, R. E. B., Arcaro, M. J., & Kastner, S. (2015). Probabilistic maps of visual topography in human cortex. *Cerebral Cortex*, *25*, 3911–3931. https://doi.org/10.1093/cercor/bhu277

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. https://doi.org/10.1016/j.neuroimage.2014.05.045

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. https://doi.org/10.1016/j.neuroimage.2015.11.043

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*, 2800–2807. https://doi.org/10.1016/j.visres.2009.08.021

Yue, X., Tjan, B. S., & Biederman, I. (2006). What makes faces special ? *Vision Research*, *46*, 3802–3811. https://doi.org/10.1016/j.visres.2006.06.017

How to cite this article: Coggan DD, Baker DH, Andrews TJ. Selectivity for mid-level properties of faces and places in the fusiform face area and parahippocampal place area. *Eur J Neurosci*. 2019;00:1–10. https://doi.org/10.1111/ejn.14327