# Methods

## Subjects

Eighteen healthy subjects (age 25.1 ± 3.28, 5 males) completed three scanning sessions. Two subjects were excluded from data analysis: one due to inattentiveness and motion in the scanner, the second due to wrong scanner settings. Before scanning, we tested the acuity of the subjects with the Freiburg Visual Acuity and Contrast Test (FrACT), which revealed all subjects had normal or corrected-to-normal vision (LogMar mean = -0.01 ± 0.11, decVA = 1.08 ± 0.39). All subjects reported being right-handed as measured with the Edinburgh handedness Questionnaire (82.9±17.3; Oldfield, 1971).

## Stimuli

Each subject performed the main experiment as well as a functional localiser to delineate the face-specialised voxels in the ventral pathway at the individual level. For both the localiser and main experiment, the stimuli were presented with PsychoPy v3.2.4 on an NNL LCD Monitor (32-inch, 1920 x 1080 pixels, 698.40 x 392.85 mm, refresh rate = 60 Hz) situated at the end of the scanner bore. Subjects viewed the stimuli via a mirror attached to the head-coil, at a viewing distance of 175cm. Across experiments, all images were attributed a mean (i.e., the global luminance) and RMS contrast of 0.45±0.1. Both the main experiment and the functional localiser had a block design.

### Functional localiser

Stimuli for the localiser experiment consisted of three image categories (faces, hands, instruments) each containing 20 greyscale images (See *supplementary*; Stigliani, Weiner, & Grill-Spector, 2015) along with their phase-scrambled counterparts. The stimuli were superimposed onto the Fourier phase-scrambled backgrounds (procedure as described above) and consisted of unfamiliar faces (10 males) of various viewpoints (including hair), isolated hands in various poses, and stringed instruments (e.g. guitar, cello, lute, etc) positioned in different orientations.

## Procedure

All subjects participated in 20 experimental runs divided over 3 scanning sessions. The first session included a T1-weighted anatomical scan and the second session a functional face localising run. All functional runs had a block design, of 10-second blocks, alternated by 10 seconds of fixation with a 12-second fixation period at the beginning and end of the run.

To ensure subjects paid attention during functional runs, we instructed them to detect a rare and brief colour change of the stimulus by pressing a button with the right index finger. In each block, there were two targets: HSV profile 1.0, 1.0, 0.8. One colour change occurred per block half, but never during the first stimulus of a block. A fixation cross, made of two thin black lines that connected opposite corners of the square stimuli, was visible throughout all runs. Subjects were instructed to fixate on its central intersection at all times.

### Functional localiser

The functional localiser was used to find voxels responding more to faces than non-face objects. Subjects viewed grayscale images of either intact or scrambled faces, hands or instruments. Images from all categories were presented in 36 blocks in total (one image category per block), with 6 blocks per condition. A condition was never repeated twice in a row and the order of conditions was counterbalanced across subjects. A block consisted of 10 images in a random order, all images were repeated three times during the whole localiser run. Each image appeared for 500ms, followed by a 500ms interstimulus interval.

## MRI acquisition

Subjects were scanned in a 3Tesla GE Signa Premier MRI scanner with a 48ch head coil, based at Cliniques Universitaires UCL Saint-Luc in Brussels. As anatomical references, whole-brain T1-weighted images were obtained during the first sessions (3D MP-RAGE, 1 x 1 x 1 mm, FOV = 256 mm, TI = 900ms, FA = 8°). Functional T2\*-weighted GE echo-planar imaging was used to obtain the blood oxygen level-dependent (BOLD) signal as an indirect measurement of neural activity. Thirty-two 2.4-mm axial slices were acquired (2.4mm isotropic, FOV = 240mm, TR = 2000ms, TE = 30ms, FA = 90°).

## Preprocessing

Functional and anatomical data were organised into the Brain Imaging Data Structure (K. J. Gorgolewski et al., 2016). Pre-processing of the data was carried out with fMRIPrep 20.1.1 (Esteban et al., 2018; Esteban et al., 2019), which is based on Nipype 1.5.0 (K. Gorgolewski et al., 2011; K. J. Gorgolewski, Nichols, Kennedy, Poline, & Poldrack, 2018). To ensure reproducibility using the specific software versions for fMRIPrep and all its dependencies, it was executed via its Docker container (Merkel, 2014).

### Anatomical data pre-processing

Each T1-weighted (T1w) volume was corrected for intensity non-uniformity (Tustison et al., 2010; N4BiasFieldCorrection), and skull-stripping was executed (*antsBrainExtraction*.shv, OASIS30ANTs template). Next, brain tissue segmentation of cerebrospinal fluid (CSF), white-matter (WM) and grey-matter (GM) was performed on the brain-extracted T1w (using *fast* from FSL 5.0.9; Zhang, Brady, & Smith, 2001). Finally, brain surfaces were reconstructed (using recon-all from FreeSurfer 6.0.1; Dale, Fischl, & Sereno, 1999), and the brain mask estimated previously was refined with a custom variation of the method to reconcile ANTs-derived and FreeSurfer-derived segmentations of the cortical grey-matter of *Mindboggle* (Klein et al., 2017).

### Functional data pre-processing

For each of the 24 functional runs per subject (across all tasks and sessions), the following pre-processing was performed. First, to generate a functional reference, volumes with substantial T1w contrast derived from nonsteady states of the scanner (volumes at the beginning of EPI sequence) were identified, realigned, and averaged. After skull-stripping of the functional reference volume, head motion parameters with respect to the functional reference (transformation matrices, and six corresponding rotation and translation parameters) were estimated (Jenkinson, Bannister, Brady, & Smith, 2002; *mcflirt* *- FSL 5.0.9*). On average, the maximum movement was 1.7±0.32mm.

After correcting for slice timing (Cox & Hyde, 1997; *3dTshift* from *AFNI 20160207)*, the functional reference was co-registered to the T1w reference using boundary-based registration (*bbregister, FreeSurfer*; Greve & Fischl, 2009). Since the volumes were within the same subject, co-registration was configured with six degrees of freedom (i.e. 3 rotations and 3 translations).

Next, all functional data were resampled onto their original, native space by applying the transforms to correct for head-motion. Several confounding time series were calculated based on the functional data: framewise displacement (FD), DVARS and three region-wise global signals. FD was computed with the absolute sum of relative motions (*Power,* Power et al., 2014) and the relative root mean square displacement between affines (*Jenkinson,* Jenkinson et al., 2002). FD and DVARS were calculated for each functional run, both using their implementations in *Nipype* (Power et al., 2014). Frames that exceeded a threshold of 0.5 mm FD or 1.5 standardised DVARS were annotated as motion outliers. The three global signals were extracted within the CSF, the WM, and the whole-brain masks.

Using FEAT (FMRI Expert Analysis Tool; Version 6.0; FMRIB's Software Library, www.fmrib.ox.ac.uk/fsl) the functional data were smoothed in the space domain using a Gaussian kernel of FWHM 5mm. And high-pass temporal filtering was carried out using a Gaussian-weighted least-squares straight-line fitting (sigma=50.0s).

## Regions of interest

The regions of interest (ROIs) in **EVC** were delineated based on retinotopy!!! 

The FFA, responding preferentially to **faces**, was identified independently for each subject, based on the localiser scan. First, to select the brain regions more responsive to faces than non-face objects we computed the conjunction between [intact faces - intact hands] and [intact faces - intact instruments] contrasts. Next, to exclude activity to low-level image properties we selected the voxels that showed a significantly larger response to intact than scrambled faces [intact face - scrambled face]. Significant voxel clusters on t maps were selected at a *q*[false discovery rate, *FDR*] < 0.01. After visual inspection, the threshold for some subjects was set increased to delineate between face preferring clusters. Overall, there were more face-selective voxels in the right FFA (143±104) compared to the left (108±93).