A Way to Improve Functional Face Localiser: The Influence of Low-Level Visual Features

Olga Korolkova, Valentin Sinitsyn, & Ekaterina Pechenkova

Federal Center for Treatment and Rehabilitation, Moscow, Russia

Abstract

In functional MRI studies a very common procedure is to use one or more functional localisers, to delineate the regions in the brain that consistently show more activation to a particular type of events or stimuli. One example is faces versus places localiser, which helps to distinguish between areas that are active during visual perception of faces and other complex images like buildings. However, it is important to make sure that the low-level properties of the images do not differ significantly between the categories of stimuli, as it can lead to incorrect identification of the areas of interest. Here we discuss one possible way to improve the specificity of face localiser, achieved by eliminating the differences in colour and intensity distribution between images of faces and of buildings using the SHINE toolbox. A group of participants (N=8) was tested on both localisers with corrected and uncorrected images, and differences in visual cortex activation were found. Therefore we recommend to use this or similar procedure to enhance the quality and reliability of functional localisers.

Keywords: fMRI, FFA, PPA, localiser, stimulus properties, low-level visual features, SHINE

In functional MRI studies a very common procedure is to use one or more functional localisers, to delineate the regions in the brain that consistently show more activation to a particular type of events or stimuli. They can be used to identify regions-of-interest (ROI) for further analysis. One example is faces versus places localiser (Kanwisher, McDermott, & Chun, 1997), which helps to distinguish between areas that are active during visual perception of faces and other complex images like buildings. However, it is important to make sure that the low-level properties of the images do not differ significantly between the categories of stimuli, as it can lead to incorrect identification of the areas of interest. Here we discuss one possible way to improve the specificity of face localiser, achieved by eliminating the differences in colour and intensity distribution between images of faces and of buildings using the SHINE toolbox. A group of participants (N=8) was tested on both localisers with corrected and uncorrected images, and differences in visual cortex activation were found. Therefore we recommend to use this or similar procedure to enhance the quality and reliability of functional localisers.

Methods

Participants

Eight participants (all females, age 20–36, mean age 27.6) without history of neurological disorders volunteered to take part in the study. All had normal or corrected-to-normal vision. Seven of them were right-handed and one was ambidextrous. They gave their informed consent before scanning.

Stimuli

Images of non-famous faces and buildings were obtained through web search, as well as from published face databases and from own archives. Face images showed people with various

posture, expression, head orientation, gaze, and light conditions. Most images included other parts of the body and background; several faces were cropped out of background. Age, ethnicity and sex of the depicted persons also varied to provide a distribution similar to the natural diversity. In the first version of localiser (LOC1) 270 faces and 270 buildings were included. For the second version of localiser (LOC2) we excluded the images with purely white or gray background, taken from standardised databases, and replaced part of them with additional images from the web. In total, 252 images of faces and 252 images of buildings were used in LOC2. Number of male and female faces was equal.

To estimate the amount of pixels with different intensity on each of red, green and blue channels, as well as across the three channels, of each image in each category (faces or buildings), and average histograms within category, we used R language (R Core Team, 2016) and package EBIImage. The averaged histograms for faces and for buildings in LOC1 are shown on Fig. 1. The same procedure was applied to the images used for LOC2 (Fig. 2). Both figures show the difference between the two categories of stimuli.

To equalise the histograms between categories, we applied iterative procedure implemented in SHINE toolbox (Willenbockel et al., 2010). It allows to match the histograms of a number of greyscale images, with the function *histMatch*. However, the images we used were in colour, therefore previous to matching the histograms all the images were split into 3 colour channels using ImageMagick (<https://www.imagemagick.org>) function *convert*. After the histograms were matched by each channel separately, the resulting triplets of red- green- and blue-channel images were again combined with ImageMagick to provide full-colour images. The averaged histograms of the latter are shown on Fig. 3.

Figure 1. Averaged histograms for faces and for buildings used as stimuli in LOC1. Bins – level of luminance (0...255); counts – number of pixels with particular level of luminance for each channel separately (red, green and blue lines correspond to red, green and blue channels).

Figure 2. Averaged histograms for faces and for buildings selected for use in LOC2 before equalisation of the histograms.

Figure 3. Averaged histograms for faces and for buildings used as stimuli in LOC2 after equalisation of the histograms.

Procedure

Participants were scanned using Siemens Avanto 1.5T. They performed two runs of LOC1 and two runs of LOC2. In LOC1, one of the runs lasted 8:38 minutes and included 8 blocks of faces, 8 blocks of buildings and 9 blocks of baseline (fixation at a black cross on a white background). Each block lasted 20.16 s. The order of blocks was interleaved: baseline, faces, buildings, baseline, buildings, faces, etc. The other run of LOC1 lasted 7:38 minutes and was the same as the first one except that the number of blocks in each category was seven and the first block after the initial baseline was buildings. The order of runs was counterbalanced across participants. Voxel size used for LOC1 was $3.6 \times 3.6 \times 3.8$ mm; TR = 2520 ms; TE = 50 ms; Flip angle = 90 deg. The participants' task was to press the button each time an image of a face or a building is repeated twice in a row (1-back task). There was one repeated image per block.

In the same scanning session, all participants performed two runs of LOC2. Both runs were of the same length, 7:17 minutes, and had similar structure. Each run included 9 blocks of faces, 9 blocks of buildings and 10 blocks of fixation. Each block lasted for 15.12 s. The order of blocks was interleaved as in the LOC1. The order of runs was counterbalanced across participants. Voxel size for LOC2 was $4.0 \times 4.0 \times 4.0$ mm; TR = 2520 ms; TE = 50 ms; Flip angle = 90 deg. The participants' task was similar to the LOC1.

In addition to each localiser, a fieldmap was acquired to control for inhomogeneity of the magnetic field. Also T1-weighted structural images were acquired with voxel size $1.0 \times 1.0 \times 1.0$ mm, TR = 1900 ms and TE = 2.91 ms.

Analysis

The data were analysed using SPM8 (Wellcome Trust Centre for Neuroimaging; <http://www.fil.ion.ucl.ac.uk/spm/software/spm8/>). The DICOM images were imported to NiFTI format, rotated to align the AC–PC line with horizontal, and preprocessed. Preprocessing of both LOC1 and LOC2 included calculation voxel displacement map from fieldmap data; realigning and unwarping functional data; coregistration of structural data to the mean functional; segmentation into white and grey matter and CSF; normalisation to MNI space and smoothing with FWHM with $8 \times 8 \times 8$ Gaussian kernel. For each localiser and each participant, first-level models were fitted, and linear contrasts for faces versus buildings and vice versa were estimated. At the second-level model we estimated the differences between faces versus buildings contrasts in LOC2 compared to LOC1 using paired *t*-tests ($p < .001$, uncorrected).

Clusters and peak coordinates for significant contrasts between faces and buildings in each localiser, and contrast between the two localisers, were extracted using peak_nii toolbox.

Results

Using second-level models, we first analysed contrast between faces and buildings in each of the localisers separately. In LOC1, four of these clusters were larger than 5 voxels. In LOC2, there were seven clusters larger than 5 voxels. The peak coordinates, cluster sizes and names of anatomical regions for these clusters are shown in Tables 1 and 2.

The second-level model revealed several clusters in which activation for faces versus buildings in LOC2 was significantly higher than in LOC1 ($p < .001$, uncorrected). Five of these clusters were larger than 5 voxels. The peak coordinates, cluster sizes and names of anatomical regions for these clusters are shown in Table 3. There were no significant clusters larger or equal to 5 voxels with activation higher in LOC1 than in LOC2.

Table 1

Peak coordinates in clusters of significant activation for faces versus buildings in LOC1

Cluster number	Cluster size	T	X	Y	Z	Region Name
1	61	7.4833	46	-56	18	r. Middle Temporal Gyrus
		7.3029	62	-48	10	
		6.5269	54	-72	2	
		6.325	54	-60	14	
		5.9994	46	-60	6	
2	29	12.2937	46	-48	-22	r. Inferior Temporal Gyrus
3	16	7.1506	30	-8	-14	r. Hippocampus
4	12	7.9752	38	20	18	r. Inferior Frontal Gyrus (pars triangularis)

Table 2

Peak coordinates in clusters of significant activation for faces versus buildings in LOC2

Cluster number	Cluster size	T	X	Y	Z	Region Name
1	17	8.3894	42	-48	-18	r. Fusiform Gyrus
2	17	6.9228	64	-46	8	r. Middle Temporal Gyrus
		4.8095	62	-56	6	
3	11	6.9589	62	-40	-2	r. Middle Temporal Gyrus
		6.5819	50	-40	2	
4	7	11.8296	30	-8	-34	r. Fusiform Gyrus
5	7	6.2604	-26	-4	-14	l. Amygdala
		9.9744	-14	-8	-10	l. Hippocampus
6	6	7.0705	-38	4	-18	l. Superior Temporal Pole
7	5	5.6861	42	20	26	r. Inferior Frontal Gyrus (pars triangulars)

Table 3

Peak coordinates in clusters of significant activation for LOC2 versus LOC1

Cluster number	Cluster size	T	X	Y	Z	Region Name
1	15	10.6748	26	-44	-14	r. Temporal Occipital Fusiform Cortex
2	13	7.9788	-10	-88	2	l. Calcarine Cortex
		5.9363	-6	-80	-2	l. Temporal Occipital Fusiform Cortex
3	9	6.9035	12	-90	0	r. Calcarine Cortex
4	7	7.7002	-22	-44	-10	l. Lingual Gyrus
5	5	5.8707	-26	-72	-6	l. Occipital Fusiform Gyrus

Discussion

Our main aim was to improve a functional localiser of brain areas that show higher activation to faces than to other visual stimuli (buildings). To achieve this, we equalised the low-level stimuli properties, namely, the luminance in each of the three colour channels, red, green and blue, across images of the two categories. Then we tested participants on the both localisers with luminance-adjusted and luminance-unadjusted stimuli, and the functional data analysis revealed significant differences in several brain areas, in which activation difference between faces and buildings was higher with the luminance-controlled stimuli. Namely, these areas include bilateral region in temporal and occipital fusiform cortex, which is also known as Fusiform Face Area (FFA). Also, bilateral region in occipital cortex was also higher activated with the luminance-controlled stimuli. Most additional activation was revealed in the left hemisphere, which is subdominant for face perception. Therefore, after controlling for low-level features, the functional localiser allowed us to reveal more face-selective regions.

References

- 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. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/9151747>
- R Core Team. (2016). R: A Language and Environment for Statistical Computing. *R Foundation for Statistical Computing, Vienna, Austria*. Retrieved from <http://www.r-project.org/>
- Willenbockel, V., Sadr, J., Fiset, D., Horne, G. O., Gosselin, F., & Tanaka, J. W. (2010). Controlling low-level image properties: the SHINE toolbox. *Behavior Research Methods*, 42(3), 671–84. <http://doi.org/10.3758/BRM.42.3.671>