Multiple scales of organization for object selectivity in ventral visual cortex

Marijke Brants a,b, Annelies Baeck a,b, Johan Wagemans b, Hans P. Op de Beeck a,*

a Laboratory of Biological Psychology, University of Leuven (K.U. Leuven), Tiensestraat 102, B-3000 Leuven, Belgium
b Laboratory of Experimental Psychology, University of Leuven (K.U. Leuven), Tiensestraat 102 B-3000 Leuven, Belgium

INTRODUCTION

Object knowledge is hierarchical. Several hypotheses have proposed that this property might be reflected in the spatial organization of ventral visual cortex. For example, all exemplars of a category might activate the same patches of cortex, but with a slightly different position of the peak of activation in each patch. According to this view, category selectivity would be organized at a larger spatial scale compared to exemplar selectivity. No empirical evidence for such proposals is available from experiments with human subjects. Here, we compare the relative scale of organization for category and exemplar selectivity in two datasets with two methods: (i) by investigating the previously reported beneficial effect of spatial smoothing of the fMRI data on the reliability of multi-voxel selectivity patterns; and (ii) by comparing the relative weight of lower and higher spatial frequencies in the spatial frequency spectrum of these selectivity patterns. The findings are consistent with the proposal that selectivity for stimulus properties that underlie finer distinctions between objects is organized at a finer scale than selectivity for stimulus properties that differentiate categories. This finding confirms the existence of multiple scales of organization in the ventral visual pathway.

ARTICLE INFO

Article history:
Received 1 July 2010
Revised 7 February 2011
Accepted 27 February 2011
Available online 3 March 2011

Keywords:
fMRI
Multi-voxel pattern analyses
Spatial scales
Spatial smoothing
Ventral visual cortex

INTRODUCTION

Human category knowledge is organized hierarchically. For example, a Labrador belongs to the category of dogs, all dogs are mammals, and all mammals are animals. The most inclusive category within which attributes are common to most category members is what Rosch et al. (1976) called the basic-level category, e.g., dog. Classifications more general than the basic level are called super-ordinate categories (e.g., mammal) while more specific distinctions are referred to as subordinate (e.g., Labrador). Here, we examine whether the hierarchical property of object representations might be reflected in the spatial organization of the human object vision pathway, so that the spatial scale at which selectivity for objects is organized depends on the hierarchical level at which these objects are differentiated.

This question is motivated by earlier research in monkeys suggesting differences in the scale of organization for basic-level and subordinate selectivity. Originally, it was argued by Tanaka and colleagues (for a review see Tanaka, 1996, 1997) that all exemplars of a basic-level category might activate the same local cortical patches (organized as feature columns) in which all neurons share a preference for a particular feature of intermediate complexity with some differences between these neurons in their exact tuning properties (Fig. 1A). This implies that a differentiation between specific exemplars within a basic category is possible only by comparing the responses of neurons within these columns or patches (Logothetis and Sheinberg, 1996). According to this view, category selectivity would be organized at a particular spatial scale and no (or at most a very small-scale) organization would exist for selectivity for finer distinctions.

However, limited evidence from later studies suggested that neighboring columns might sometimes have a related stimulus preference so that particular parametric variations of a basic shape, for example caused by viewpoint changes, result in a gradual shift of activation across the local ensemble of columns (Tanaka, 1997, 2003; Wang et al., 1998). These findings resulted in a modification of the model in the sense that some changes to objects that do not change category membership might nevertheless change the pattern of activity across the cortical surface, albeit locally at a relatively small spatial scale (Fig. 1B). This modified model predicts that objects from different categories would be associated with a pattern of responses across feature columns that are widely separated in cortical space while objects from the same category, but undergoing different transformations, would only differentially activate nearby columns. Despite the potential importance of these hypotheses, the supporting evidence in monkeys is very limited (only a few studies) and scattered (evidence comes from the investigation of a few feature columns). No evidence exists from experiments with human subjects. The main problem in humans is that the noninvasive method with the highest spatial resolution, functional magnetic resonance imaging (fMRI), is still limited to the range of multiple millimeters. This implies that one voxel (the smallest unit of measurement in fMRI) still contains thousands or even tens of thousands of neurons and multiple feature columns—if such columns exist in humans. In order to infer the spatial
scale at which neural properties are organized from such fMRI-data, we combined multi-voxel pattern analyses with a manipulation of the degree of spatial smoothing applied to fMRI-images.

Multi-voxel pattern analyses focus on the pattern of responses across voxels (Haxby et al., 2001). The most straightforward multi-voxel pattern analysis correlates the spatial activity pattern for one condition in a subset of the data with the activity pattern for the same or a different condition in another subset of the data (Op de Beeck et al., 2006). If this correlation is reliably higher for correlations between data from the same condition than for correlations between data from different conditions then the activity pattern is said to be a reliable indicator of differences between conditions (Haxby et al., 2001). The suggestion has been made that these multi-voxel pattern analyses are able to differentiate patterns of activity at a finer scale than conventional univariate analyses (Downing et al., 2007; Haynes and Rees, 2005; Kamitani and Tong, 2005). The sensitivity of this method to functional organization at a fine spatial scale is difficult to prove conclusively (Kamitani and Sawahata, 2010; Op de Beeck, 2010a, 2010b; Shmuel et al., 2009), but some insights into the scale of organization can be obtained by investigating the effect of spatial smoothing on multi-voxel pattern analyses (Op de Beeck, 2010a, 2010b). Indeed, the expected effect of smoothing is different for a small-scale (sub-voxel) and a large-scale organization. A small-scale organization would be expected to benefit from enhancing higher spatial frequencies in the signal compared to the lower spatial frequencies while the opposite is true for a large-scale organization. Spatial smoothing reduces these high spatial frequencies and thus is expected to have less of a beneficial effect on the decoding of the fine-scale spatial signals. This prediction was quantified through simulations in earlier papers and held under a range of parameter settings and assumptions (Kamitani and Sawahata, 2010; Op de Beeck, 2010a, 2010b). Thus, a difference in the effect of smoothing on two selectivity patterns indicates a difference in the scale of organization.

Here, we applied this method on two datasets to find out whether multiple scales of organization exist in the human visual object pathway. With the first dataset, we tested and confirmed the hypothesis of Tanaka (1996, 2003) and Logothetis and Sheinberg (1996) by showing that category selectivity is organized at a larger scale than subordinate-level selectivity across the large region of cortex occupied by regions selective for faces, bodies and houses. The second dataset tested and confirmed the hypothesis based on Wang et al. (1998) and Tanaka (2003) that category selectivity is organized at a larger scale than selectivity for viewpoint changes across the large region of cortex with object-selective responses, known as the lateral occipital complex (LOC; Grill-Spector et al., 2001; Kourtzi and Kanwisher, 2000; Malach et al., 1995). The same conclusions are reached from a second method which focuses upon the amplitude of lower and higher spatial frequencies in the three-dimensional selectivity patterns related to category and subordinate-level selectivity. These findings confirm the general hypothesis that selectivity for stimulus properties that underlie finer distinctions between objects is organized at a finer scale than selectivity for stimulus properties that differentiate basic-level categories. We conclude that multiple scales of organization co-exist in human ventral visual cortex.

Materials and methods

Subjects

The data from Experiment 1 come from nine normal and naive subjects and were described in detail in a previous report by Op de Beeck et al. (2010). Eight normal and naive subjects participated in Experiment 2 (all participants were right-handed; 6 participants were female and all participants were between 18 and 25 years old). The experiments were approved by the relevant ethical boards, that is, the ethical committee of the Faculty of Psychology and Educational Sciences (K.U.Leuven) and the committee for medical ethics of the K.U.Leuven. All participants were paid for their participation and informed consent was obtained from all subjects.

Stimuli and fMRI task

In Experiment 1 the categories consisted of faces, buildings, and body parts and the subordinate-level distinctions were old and baby-faces, skyscrapers and rural buildings, and hands and torsos, respectively (Fig. 2A). In the scanner, participants performed a one-back task. Functional runs (N=6–12) consisted of 15 blocks of 15 s, including 3 fixation blocks (at the start, middle and end of the run) and 2 blocks of each condition. The order of the conditions was counterbalanced across subjects and across runs. A detailed description of Experiment 1 can be found in Op de Beeck et al. (2010).

Experiment 2 was very similar to the first scan sessions in a publication by Gauthier et al. (1999). Forty greyscale faces (obtained from Niko Troje and Heinrich Bülthoff, Max Planck Institute, Tübingen, Germany) and forty grayscale greebles (photorealistically rendered 3D objects that share similar parts in a common spatial configuration Gauthier and Tarr, 1997) were used as categories. In addition, the orientation of the stimuli was manipulated by presenting upright and inverted images of these faces and greebles (Fig. 2B). Participants performed a sequential matching task in the scanner. Functional runs (N=6) included 4 repetitions of a CACB cycle where A and B were sequential matching epochs of 24 s showing upright or inverted stimuli (order counterbalanced in different runs and across subjects) and C was 12 s of a fixation cross. Each epoch included eight trials and showed pairs of stimuli all upright or inverted. In one trial (3 s) a fixation cross was presented (800 ms), followed by two stimuli, separated by a brief (200 ms) pattern mask to prevent matching from retinal persistence. Two of the eight trials in a 24-second block were immediate repetitions. Subjects performed same/different judgments by pressing one of two buttons. Subjects made no more than 15% errors in this sequential matching task for any condition or at any point in the experiment. A localizer was obtained by passive viewing of faces, everyday objects and scrambled pictures. Localizer runs were 4 CABCD cycles with A, B and D passive viewing epochs of 20 s showing faces, everyday objects or scrambled pictures (order counterbalanced across different runs) and C was 6 s of a fixation cross. Each epoch contained 20 stimuli.

Scanning

For Experiment 1, functional images were acquired in a 3 T Philips Intera magnet (Department of Radiology of K.U.Leuven) with an 8-channel SENSE head coil with an echo-planar imaging sequence (75 time points per time series or “run” repetition time, 3 s; echo time, 30 ms; acquisition matrix 104×104, resulting in a 2.1×2.1 in plane...
In the left panel, the stimuli used in Experiment 1 are presented. The categories are faces and houses and these categories are subdivided in subordinate-level distinctions: old faces and baby faces and rural buildings and skyscrapers, respectively. The experiment also included two non-face body parts conditions, torsos and hands, but this distinction is not used here because we found previously that it might not be at a subordinate level, both perceptually and in terms of neural selectivity (Op de Beeck et al., 2010). In the right panel, the stimuli used in Experiment 2 are shown. The basic-level categories are faces and greebles. By inverting these stimuli, we created subordinate-level distinctions: upright and inverted faces or greebles.

### Analysis of imaging data

For both experiments, data were analyzed using the Statistical Parametric Map Software package (SPM5, Wellcome Department of Cognitive Neurology, London, UK) as well as custom Matlab code (Mathworks, Inc.). Preprocessing involved realignment to correct for motion, co-registration of functional and anatomical images, segmentation and spatial normalization. During spatial normalization, functional images were resampled to a voxel size of 2×2×2 mm in Experiment 1 and to a voxel size of 2.5×2.5×2.5 mm in Experiment 2. All our data are equally affected by spatial normalization and resampling given that conditions were counterbalanced in terms of presentation order—therefore these steps will have no differential effect on the different conditions. We analyzed the data before and after spatial smoothing in the full volume (3D kernel) in two levels, by means of Gaussian kernels with a full-width-at-half-maximum (FWHM) of twice and four times the original voxel size (FWHM of 4 and 8 mm in Experiment 1, FWHM of 5 and 10 mm in Experiment 2).

Statistical modeling of the signal in each voxel in each subject was accomplished this by calculating the correlations for the across-category condition on single runs, instead of all runs. This made sure that conditions were counterbalanced in terms of presentation order—therefore these steps will have no differential effect on the different conditions. We analyzed the data before and after spatial smoothing in the full volume (3D kernel) in two levels, by means of Gaussian kernels with a full-width-at-half-maximum (FWHM) of twice and four times the original voxel size (FWHM of 4 and 8 mm in Experiment 1, FWHM of 5 and 10 mm in Experiment 2).

### Multi-voxel pattern analyses

In both experiments, we computed correlations following previously introduced methods for multi-voxel pattern analyses (Haxby et al., 2001; Op de Beeck et al., 2008b). The data were divided in two random equally sized subsets of runs (this procedure was repeated 12 times in Experiment 1, and 4 times in Experiment 2 to make sure that results were not restricted to a specific split). We constructed lists as long as the number of voxels in an ROI, and each list contained the parameter estimates for all voxels for a particular condition in one subset. These responses were converted into a measure of selectivity per voxel by subtracting conditions pair-wise (e.g., rural buildings minus skyscrapers). Then, the values in each list of the first subset were correlated with the values in each list of the second subset. Because the correlations for the across-category condition were always much higher than the correlations for the subordinate condition, we had to bring the correlations to the same level prior to smoothing to be able to compare the conditions reliably. We accomplished this by calculating the correlations for the across category condition on single runs, instead of all runs. This made sure

---

**Fig. 2.** In the left panel, the stimuli used in Experiment 1 are presented. The categories are faces and houses and these categories are subdivided in subordinate-level distinctions: old faces and baby faces and rural buildings and skyscrapers, respectively. The experiment also included two non-face body parts conditions, torsos and hands, but this distinction is not used here because we found previously that it might not be at a subordinate level, both perceptually and in terms of neural selectivity (Op de Beeck et al., 2010). In the right panel, the stimuli used in Experiment 2 are shown. The basic-level categories are faces and greebles. By inverting these stimuli, we created subordinate-level distinctions: upright and inverted faces or greebles.
that the correlations for both conditions were similar when the data were not smoothed (FWHM = 0; see Fig. 7), making it possible to compare the effects of smoothing in both conditions.

Analyses by 3D Fourier transformation

To back up our correlation analyses we included a second, more direct measure of the scale of organization. We computed the power spectrum of the data in both experiments by means of 3D Fourier transformation (Fast Fourier Transform or FFT). We defined a rectangular volume of the size needed to encompass all voxels in the ROI in one hemisphere, and took the 3D transform of the differential beta maps within this volume. This results in an amplitude spectrum of the signal and we divided this spectrum into 5 intervals and each interval represents another level of frequency (interval one: the lowest frequency range, 1–4 cycles per large cortical cube; interval two: 5–8 cycles per large cortical cube; interval three: 9–12 cycles per large cortical cube; interval four: 13–16 cycles per large cortical cube; interval five: the highest frequency range, more than 16 cycles per large cortical cube). Within each frequency interval and across all orientations, we calculated the mean amplitude of the signal (per subject), for the category condition as well as for the subordinate condition, for both the smoothed (highest level of smoothing) and the unsmoothed data (see Fig. 3 for the expected pattern of means). From these means we calculated two ratios to investigate whether there is proportionally more or less power in the subordinate condition: (the mean amplitude for the highest frequencies in the subordinate condition/the mean amplitude of the highest frequencies in category condition; in Fig. 3 this corresponds with A/B) and (the mean amplitude for the lowest frequencies in the subordinate condition/the mean amplitude for the lowest frequencies in the category condition; in Fig. 3 this corresponds with C/D). We worked with these ratios because these ratios are invariant for overall differences in the strength/power of selectivity maps.

2.7. Multidimensional scaling

To investigate whether category selectivity is indeed much stronger than subordinate-level selectivity, which is an almost trivial assumption that we make throughout this text, we performed multidimensional scaling on the data obtained with the middle level of smoothing (FWHM = 4–5 mm). For this analysis, correlations were computed for the response to individual conditions after subtracting the mean response across all conditions (instead of the selectivity between pairs of conditions). These correlational analyses provide a full matrix with in each cell i, j—the spatial correlation between stimulus conditions i and j. We performed metric multidimensional scaling (metric MDS; using the Matlab function ‘mdscale’) to visualize the representational space of the four relevant conditions in each experiment according to the neural similarity data. MDS places the conditions in a low-dimensional space in a way that optimizes the correspondence between the relative proximity among the stimuli and the similarity in the correlation matrix (Op de Beeck et al., 2001).

Simulation of the expected effect of smoothing

Simulated high-resolution functional maps were created. They were implemented as two-dimensional matrices, each with 299 × 299 measurement units. The total matrix size was defined to be 6 cm in width, so each unit had a size of 0.2 × 0.2 mm. Three pairs of two functional maps were created, shown in the first two columns of Fig. 4. These maps simulate the pattern of responses in two conditions, Stim1 and Stim2. The difference between the two conditions gives the high-resolution selectivity map in the third column of Fig. 4.

We modelled how the selectivity in these cortical surfaces would be measured through fMRI. We computed the response in voxels of 2 × 2 mm by averaging the response across 10 × 10 cells in each matrix, and adding noise to this average to obtain different measurements or ‘runs.’ For each run the matrix was slightly shifted by a random amount in a uniform range of 2 mm to simulate the effect of subject motion. The noise was correlated across neighboring pixels (a white noise field was smoothed with a Gaussian kernel of FWHM of 4 mm), and its power was chosen to result in correlations within the range obtained in the empirical data. Subject motion and noise correlations are important for the absolute effect of smoothing, but the values of these parameters are not crucial for the main predictions about the difference in the effect of smoothing between selectivity patterns of a different spatial scale. We created a maximum of 30 ‘runs’ for each matrix. These runs were divided in two subsets of runs and subjected to different levels of smoothing with a Gaussian kernel (no smoothing and smoothing of 4, 8, 12, 20, and 28 mm). A measured selectivity map (a non-smoothed and a smoothed version) was created for each subset of runs by taking the difference between the two conditions Stim1 and Stim2. These selectivity maps were correlated across voxels between the two subsets of runs.

The overall level of performance in these simulations, i.e. the obtained correlations, depended strongly on signal-to-noise ratio. We adjusted the signal and noise properties to obtain non-smoothed correlations in the range found with the empirical data. In addition, the size of the difference between the two conditions Stim1 and Stim2 was adjusted so that the correlations obtained for a weak large-scale organization (Fig. 4B) and for a strong small-scale organization (Fig. 4C) were comparable.

Results

Formal predictions about effects of smoothing as derived from a simulation study

We will first present the results of the simulation study that provides the formal background for the predictions tested in the two
empirical experiments. Fig. 4 illustrates the situations that we want to differentiate based on the effect of smoothing. Assume a large-scale response pattern for two very different conditions Stim1 and Stim2 that activate disparate regions in the brain (e.g., Stim1 = faces and Stim2 = buildings). A contrast of the two response patterns results in a very clear selectivity map (Fig. 4A). Here, we question the properties of selectivity maps for more similar conditions, e.g., Stim1 = rural buildings and Stim2 = skyscrapers. While traditional voxel-by-voxel analyses are not able to differentiate between these two conditions, multi-voxel pattern analyses are able to pick up such fine differences. Two opposing hypotheses can explain such a pattern of results. First, the overall response of the different cortical regions can be slightly different for Stim1 and Stim2 (Fig. 4B). The difference in activity is too small to pick up the difference with voxel-by-voxel analyses, but multi-voxel analyses can detect the difference because they combine the information from multiple regions. If this hypothesis is true, then the selectivity map for Stim1 versus Stim2 in Fig. 3B would have the same spatial scale as the selectivity map in Fig. 4A. Therefore, we will refer to this hypothesis as the “one-scale” hypothesis. The second hypothesis, illustrated in Fig. 4C, states that Stim1 and Stim2 activate slightly different regions which implies that the peak of activity has shifted slightly, as was shown for face orientation in optical imaging data (Wang et al., 1998). Again, this shift might be too small to be detected with voxel-by-voxel analyses. If this hypothesis is true, then the resulting selectivity map (Fig. 4C, right column) has a smaller spatial scale than the selectivity map shown in Fig. 4A. We will refer to this hypothesis as the “multiple-scales” hypothesis. Here, we present the two hypotheses as extremes, though it is more likely they represent two points on a continuum.

Previous simulations (Kamitani and Sawahata, 2010; Op de Beeck, 2010a, 2010b) suggest that the two hypotheses can be differentiated by investigating the effects of smoothing on the outcome of multi-voxel pattern analyses: larger-scale maps benefit more from smoothing.

Here, we will illustrate the general procedure for the maps shown in Fig. 4. We simulated an experiment with the two conditions shown in panel 4A, with an experiment being composed of multiple fMRI-runs in which the two conditions were presented. To simulate the limited resolution of fMRI, each map was sub-sampled by voxels of 2 x 2 mm, and multiple runs were generated by adding noise (see Multidimensional scaling for more details). For each run, we contrasted the two conditions resulting in a measured selectivity map for each run. The selectivity maps were averaged for two groups of runs to obtain two estimates of the selectivity map. These two estimated maps were correlated across voxels. This is the same procedure as followed in many previous fMRI-studies (e.g., Haxby et al., 2001; Op de Beeck et al., 2008b, 2010). Fig. 4 displays the resulting correlations for the conditions shown in Fig. 4A–C. We focus first on the comparison of the two weak selectivity maps associated with the two hypotheses that we want to contrast. The across-voxel correlations were comparable between the two hypotheses when no smoothing was applied. However, consistent with previous simulations, smoothing has a much stronger beneficial effect on the weak selectivity pattern in the case of the one-scale hypothesis than for the multiple-scales hypothesis. Furthermore, we see that this difference between the one-scale hypothesis and the multiple-scales hypothesis increases with higher levels of smoothing.

In this simulation we know the underlying organization and we know which hypothesis is true. However, this is not known for real data. Given that the hypotheses in the Introduction were formulated in terms of different levels of stimulus similarity being associated with a different scale of organization, we will have to compare the spatial scale of the selectivity maps associated with different levels of

**Fig. 4.** Different hypotheses about the type of organization that can underlie weak selectivity maps related to the difference between two stimulus conditions. We show simulated maps of the activity elicited by two stimuli (two left columns), and the difference between these activity maps, that is, a selectivity map obtained by subtracting the two activity maps. (A) The top row illustrates a strong selectivity map as would be obtained by comparing two very different conditions that activate very different cortical regions (e.g., a face as Stim 1 and a house as Stim 2). (B) The two other rows illustrate the two possible types of organization that might underlie weak selectivity maps associated with relatively small stimulus differences (e.g., for two different types of faces). These stimulus conditions might activate the same cortical regions, but with slight differences in the level of activity at each spot. In this case, the selectivity map would reflect an organization at the same large scale as the original activity map (one-scale hypothesis). (C) Here, a weak selectivity map is based on small shifts of the peak of activation for the different stimulus condition, exactly as was shown in previous optical imaging experiments (Tanaka, 2003). The resulting selectivity map is organized at a finer scale (multi-scales hypothesis).
stimulus similarity. This comes down to comparing the spatial scale of the map in Fig. 4A with the spatial scale of the weaker selectivity maps (panels 4B and C). The problem is that the average correlations will be much higher in the case of stronger selectivity maps, even if the scale is the same and irrespective of the degree of spatial smoothing. This is not visible in Fig. 5 because we used a trick: for the weak selectivity patterns in Fig. 4B and C many runs were averaged to obtain the correlations while for the strong pattern in Fig. 4A we used a smaller number of runs to compensate for the stronger signal. With all runs, the correlations for the strong selectivity map went up to 0.90 without smoothing and 0.93 with smoothing, suggesting a ceiling effect in this condition.

In the empirical experiments, we will adopt the approach suggested by these simulations and compare the effect of smoothing on the correlations obtained for the relatively strong selectivity maps for category selectivity as measured with a subset of the runs with the effect of smoothing on the correlations obtained for the relatively weak selectivity maps for subordinate-level distinctions as measured with all runs. This comes down to a comparison of Fig. 4A and B (if the one-scale hypothesis is true) or a comparison of Fig. 4A and C (if the multiple-scales hypothesis is true).

Experiment 1: category selectivity versus subordinate-level distinctions

As quantified before (Op de Beeck et al., 2010), category selectivity for faces versus buildings is much stronger than subordinate-level selectivity among faces or among buildings. This is illustrated in Fig. 6 which displays the spatial configuration obtained from a multidimensional scaling analysis of the neural correlation matrix. The proximity of the subordinate-level distinctions reflects their high similarity in terms of elicited response patterns. In Fig. 7, we represent an illustrative figure of the differential activation patterns for the cross- and within-category comparisons. This figure illustrates for one subject how the differential activations for the within category comparison (subordinate condition) can be represented on a smaller spatial scale than the selectivity for the cross category comparison. Furthermore, we can observe that by smoothing the data, the activations become less differentiated.

To investigate whether the spatial response patterns for subordinate-level distinctions are organized at a finer spatial scale in the brain, we first correlated the pattern of selectivity in the category-selective region associated with the difference between two conditions in half of the runs with the pattern of selectivity associated with the same two conditions in the other half of the runs. We did this for no-smoothing (FWHM = 0) and smoothing of twice the voxel size (FWHM = 4). Correlations were significantly higher than zero (t-test across runs, p < 0.001) for each subject for each level of smoothing. In addition, smoothing resulted in higher correlations: in each subject the correlations were significantly higher with FWHM = 4 than with FWHM = 0 (paired t-test across runs, p < 0.001). Then, we compared the effect of smoothing on the strong (contrast between faces and houses) and the weak (contrast between baby-faces and elderly faces, and rural buildings and skyscrapers) selectivity map. When all runs were used to calculate the correlations (for all contrasts), we found higher correlations for the strong selectivity map in comparison to the weak selectivity maps (0.77 versus 0.36 without smoothing and 0.86 versus 0.41 with smoothing). When calculating correlations between single runs for the strong selectivity map with no smoothing, which was the suggested method based on our simulations, we obtain correlations similar to the ones for the weak selectivity map with no smoothing (see Fig. 8). Most importantly, after smoothing there is a boost to all correlations, but the correlations increase significantly less for the weak selectivity map (contrast between the subordinate-level categories) in comparison to the strong selectivity map (contrast between the category distinctions), t(8) = 4.19, p < 0.005. Additional analyses with smoothing of four times the voxel size (FWHM = 8) showed the same effect, indicating that it is not restricted to a particular level of smoothing. When a repeated measures ANOVA is applied on these data, we find a significant interaction (smoothing x condition) effect (F(2,16) = 14.85; p < 0.0001) which further emphasizes the differential effect of smoothing on the weak and the strong selectivity map.

The results of the Fourier transformation analyses agree with these of the correlation analyses (see Fig. 9). For the unsmoothed data (FWHM = 0) we find that there is significantly more power in the low spatial frequencies in the category condition compared to the within-category condition: The ratio obtained from comparing conditions in the highest frequency interval differs significantly from the ratio obtained from comparing the conditions in the low frequency band (t(8) = 5.4142, p < 0.0001). This indicates that the differential activations for the category condition are indeed situated at a larger spatial scale than the differential activations for the subordinate condition. However, when we look at the smoothed data (FWHM = 8), this difference between conditions disappears (t(8) = 1.7101, ns).

Experiment 2: category selectivity versus orientation selectivity

Fig. 5 (right panel) shows the MDS-derived spatial configuration for the 4 conditions in Experiment 2. Here, the strong category-level selectivity map is associated with contrasting greebles and faces while the weak selectivity map is generated by contrasting upright versus inverted faces or greebles (subordinate-level distinctions). We first correlated the pattern of selectivity in the LOC in half of the runs with the pattern of selectivity associated with the same pair of conditions in the other half of the runs. We did this for no smoothing (FWHM = 0) and smoothing of twice the voxel size (FWHM = 5). Again, correlations were significantly higher than zero (t-test across runs, p < 0.001) for each subject for each level of smoothing and smoothing increased the size of the correlations (paired t-test across runs, p < 0.001). Regarding the effect of smoothing, we observed similar results as in Experiment 1: all correlations increase after smoothing, but the weak selectivity maps benefit less from smoothing than the strong selectivity map (t(7) = 2.63, p < 0.05); correlations for the strong selectivity map were again calculated on the basis of restricted data (in this case 1 instead of 3 runs; see Fig. 7). Additional analyses with smoothing of four times the voxel size (FWHM = 10) showed that this effect is also present when higher levels of smoothing are applied and compared to the no-smoothing condition. When a repeated measures ANOVA (with two within subject factors) is applied, we also find a significant interaction effect (F(2,14) = 4.63, 0.01).
which further emphasizes the differential effect of smoothing on the weak and the strong selectivity map.

The results of the Fourier transformation analyses are again in agreement with these of the correlation analyses (see Fig. 9). When data are not smoothed (FWHM=0), we find that there is significantly more power in the low spatial frequencies in the category condition compared to the within-category condition: The ratio obtained from comparing conditions in the highest frequency interval differs significantly from the ratio obtained from comparing the conditions in the low frequency band ($t(7)=3.258, p<0.01$). This indicates that the differential activations for the category condition are indeed situated at a larger spatial scale than the differential activations for the subordinate condition. However, when we look at the smoothed data (FWHM=10), this difference between conditions disappears ($t(7)=1.3569, \text{ns}$).

These results are in line with the existence of multiple scales of organization. The larger effect of smoothing for the categorical distinctions indicates that the underlying selectivity maps are organized at a larger spatial scale than the selectivity maps underlying more subordinate distinctions.

**Discussion**

To summarize, we obtained similar findings in both experiments. In agreement with previous studies (Op de Beeck et al., 2008a,b), correlations in the selectivity pattern between runs increased after smoothing (for all conditions). Most importantly, smoothing had a larger beneficial effect on category-level distinctions than on subordinate-level distinctions. Based on our a priori predictions and

![Fig. 6](image)

**Fig. 6.** The MDS-derived two dimensional representation space of the four conditions in Experiment 1 and Experiment 2. In Experiment 1, the MDS is based on the correlation of the selectivity pattern between conditions across all category-selective voxels. In Experiment 2, the MDS is based on the correlations of the selectivity pattern between conditions in the LOC. If the correlation is high between two conditions, then their points in this spatial representation are close together.

![Fig. 7](image)

**Fig. 7.** An illustrative figure (for Experiment 1) showing the differential activation patterns for the cross and within category comparisons in the ventral part of the category-selective ROI of one subject. This subject strongly illustrates the overall phenomenon that the within category comparison is represented on a smaller spatial scale than the between category comparison. Furthermore, we can observe that by smoothing the small-scale differences between conditions from the same category mostly disappear.
a simulation, this differential effect of smoothing argues in favor of ‘multi-scales’ hypotheses. This implies that finer distinctions between objects in conceptual space are also organized at a smaller spatial scale in the brain than larger distinctions between categories. To back up this correlation analyses we included a second, more direct measure of the scale of organization by comparing the relative weight of lower and higher spatial frequencies in the spatial frequency spectrum of the multi-voxel selectivity patterns. When we analyzed the unsmoothed data, we found proportionately more power in the low spatial frequencies in the category case than in the within-category comparison. However, when we looked at the smoothed data, this difference between conditions disappeared. This implies that the differential activation for category comparisons is indeed at a larger spatial scale than for the within-category comparison.

Here, we presented the ‘one-scale’ hypothesis and the ‘multi-scales’ hypothesis as two extremes and find evidence for one of them. However, in reality, there is probably any number of ways in which selectivity might be differentially distributed across spatial scales. In particular, in our simulations the smaller scale organization was always organized at a finer scale, and always at the same finer scale for each spot in the map. This is unrealistic. It would be as unrealistic as generalizing the data from Wang et al. (1998) about a gradual shift for some visual dimensions in a few patches to the whole ventral visual pathway by saying that the same gradual shift occurs over the entire region. It would be more appropriate to interpret our findings as the result of the ‘average’ spatial scale at which particular object distinctions are organized. So we conclude from the effects of spatial smoothing that, on average, this scale is smaller for subordinate-level distinctions than for category-level distinctions. Not all sub-regions within ventral visual cortex will show this difference to the same degree, so the average might hide a substantial variability. The large field of view of fMRI enables us to investigate such overall differences, in contrast to the more examplatory approach of optical imaging.

It is probably also unrealistic to narrow the number of spatial scales down to two. As an example, orientation selectivity in primary visual cortex (V1), arguably the most intensively studied property in the visual system, is organized at a sub-millimeter scale in orientation columns, at a millimeter scale in pinwheels, linear orientation bands, and relative to ocular dominance domains, and at a multi-millimeter scale in terms of polar angle relative to the fovea (e.g., radial bias). At an absolute level, it might be possible to differentiate whether fMRI selectivity comes from the lower end of the scale (sub-millimeter or...
Thus, our conclusions are formulated in relative terms: the organizational scale for subordinate selectivity is smaller than for category-level selectivity. Previous simulation studies (Kamitani and Sawahata, 2010; Op de Beeck, 2010a,b) and experimental findings (Swisher et al., 2010) have shown how the exact effect of spatial smoothing depends on various factors independent of the scale of organization: subject motion, the properties of noise, magnetic field strength and scanning resolution. For example, spatial smoothing (expressed in FWHM) with twice the voxel size typically increases the reliability of the data, resulting in higher correlations between different datasets, and does not decrease decoding performance of pattern classifiers (e.g., Op de Beeck et al., 2008a, 2010; Swisher et al., 2010). This is true, whether the scanning resolution is 0.3, 1, 2 or 3 mm (spatial smoothing with 0.6, 2, 4 and 6 mm FWHM, respectively), even though all these studies investigated the same organization: orientation maps in primary visual cortex. However, the same set of studies confirms the validity of relative conclusions: in a situation where subject motion, noise and scanning parameters are controlled by testing the different conditions in the same experiment, same subject and counterbalanced in the same scan runs, differences in the effect of spatial smoothing between conditions indicate differences in the spatial scale of organization.

Another reason to refrain from attempts to deduce the exact scale of organization from our fMRI measurements is that the relationship between the two depends on many parameters. For example, the spatial spread of the fMRI signal is affected by the BOLD point spread (Parkes et al., 2005), which would not only lead to further correlations in the noise, but also correlate signal among neighboring voxels. Another factor is that voxels do not act as simple averaging boxes. Instead, the imaging point spread function (as distinct from the BOLD point spread) can be better described as a sinc function (Haakcke et al., 1999). Both of these effects would serve to reduce SNR at high spatial frequencies.

What are the implications of our findings for the main models about spatial organization of ventral visual cortex? First and foremost, we have to think about this cortical region as containing multiple spatial scales. A central debate is centered on the distinction between modular and distributed representations. At the one extreme of this debate, one can account for specialized modules dedicated to the processing of specific object categories such as faces (FFA; Haxby et al., 2000; Kanwisher et al., 1997; Kanwisher and Yovel, 2006), buildings (PPA; Aguirre et al., 1998; Epstein and Kanwisher, 1998) and human body parts (Downing et al., 2001). At the other extreme, there are accounts that argue for widely distributed representations of all object categories, in which various objects are coded through the unique patterns of neural activity within the entire representation (Edelman, 1998; Haxby et al., 2001). Other authors integrate both ideas and claim that a distributed and orderly map of eccentricity (the anatomical segregation of central versus peripheral visual field bias which has been previously demonstrated for early retinotopic areas; Sereno et al., 1995; Tootell et al., 1997) actually extends into ventral high-order objects areas (Goesaert and Op de Beeck, 2010; Levy et al., 2001; Malach et al., 2002) and determines the location of the category-selective modules. Such an integrative hypothesis is already suggesting the existence of multiple scales, a retinotopic map at a very large scale and a map of category-level distinctions at a (potentially) finer scale. Our findings suggest the existence of even finer scales of organization. By itself, multiple scales are perfectly consistent with a modular account, which is illustrated by the aforementioned example of the multiple-scale organization of orientation selectivity in V1 (assuming orientation columns are ‘modules’). However, at the very least, we have to acknowledge that there is more to the organization of ventral visual pathway than a loose set of modules. There might be a larger-scale organization (Malach et al., 2002) and our current data suggest there is also an organization at a finer spatial scale, such as, a shift of the spatial pattern of selectivity within the modules as a function of subordinate-level distinctions.

Acknowledgments

This work was supported by the Research Council of K.U.Leuven (“IMPH/06/GHW, CREA/07/004”), the Fund for Scientific Research—Flanders (“G.0281.06, G.0562.10”), by a Methusalem grant (“METH/08/02”) from the Flemish Government, and by a federal research action (“IUAP P6/29”). We thank L de-Wit, J Kubilius and W Braet for relevant discussions, and R Peeters for technical assistance.

References


