



# Segregation and persistence of form in the lateral occipital complex

Susanne Ferber<sup>a,\*</sup>, G. Keith Humphrey<sup>b</sup>, Tutis Vilis<sup>b,c</sup>

<sup>a</sup> Department of Psychology, University of Toronto, 100 St. George Street Toronto, Ont., Canada M5S 3G3

<sup>b</sup> Department of Psychology, CIHR Group for Action and Perception, University of Western Ontario, London, Ont., Canada

<sup>c</sup> Department of Physiology and Pharmacology, CIHR Group for Action and Perception, University of Western Ontario, London, Ont., Canada

Received 25 May 2004; received in revised form 21 June 2004; accepted 28 June 2004

## Abstract

While the lateral occipital complex (LOC) has been shown to be implicated in object recognition, it is unclear whether this brain area is responsive to low-level stimulus-driven features or high-level representational processes. We used scrambled shape-from-motion displays to disambiguate the presence of contours from figure-ground segregation and to measure the strength of the binding process for shapes without contours. We found persisting brain activation in the LOC for scrambled displays after the motion stopped indicating that this brain area subserves and maintains figure-ground segregation processes, a low-level function in the object processing hierarchy. In our second experiment, we found that the figure-ground segregation process has some form of spatial constancy indicating top-down influences. The persisting activation after the motion stops suggests an intermediate role in object recognition processes for this brain area and might provide further evidence for the idea that the lateral occipital complex subserves mnemonic functions mediating between iconic and short-term memory.

© 2004 Elsevier Ltd. All rights reserved.

**Keywords:** LOC; Shape-from-motion; Perceptual persistence; Binding; Human; fMRI

## 1. Introduction

It is well established that the lateral occipital complex (LOC) is critical for visual object recognition (Corbetta, Miezin, Dobmeyer, Shulman, & Petersen, 1990; Faillenot, Toni, Decety, Grégoire, & Jeannerod, 1997; Grill-Spector, Kushnir, Hendler, & Malach, 2000; Grill-Spector, Kourtzi, & Kanwisher, 2001; Halgren et al., 1999; James, Humphrey, Gati, Menon, Goodale, 2000; Kanwisher, Chun, Mc Dermott, & Ledden, 1996; Kraut, Hart, Soher, & Gordon, 1997; Malach et al., 1995; Sergent, Ohta, & MacDonald, 1992), however, its precise role remains a mystery. Does it subserves higher order representational processes or lower level functions such as segregating a candidate figure from the background noise?

Normally, in order for something to be identified and represented as a figure, its contours need to be identified. Therefore, figure-ground segregation is intimately associated with the presence or at least perception of contours. Accordingly, it has been shown that for humans, neurons in the LOC (Grill-Spector et al., 1998, 2001) and, in monkeys, neurons in the inferior temporal (IT) region (Vogels, 1999) decrease their firing rate as images become more scrambled. Moreover, the LOC is typically located by subtracting the activation in response to scrambled objects from the response to intact objects. In addition, it has been demonstrated that in the absence of real contours the LOC responds to illusory contours (Mendola, Dale, Fischl, Liu, & Tootell, 1999; Stanley & Rubin, 2003).

These findings are open to two interpretations. According to the more obvious one, these findings imply that real or illusory contours are necessary to activate the LOC. The alternative interpretation is that the LOC activation reflects the automatic response to figure-ground segregation processes

\* Corresponding author. Tel.: +1 416 978 1537; fax: +1 416 978 4811.  
E-mail address: ferber@psych.utoronto.ca (S. Ferber).

in this area regardless of the presence of contours. In the later case, the response to contours would simply represent an epiphenomenon. How can we disambiguate figure perception from contour perception? One approach is to scramble the image of an intact object such that the contours are broken up, and then initiate a binding process of these fragments that does not rely on object-sensitive information, for example, by using relative motion. The shape-from-motion paradigm typically employs images of real objects such that one set of lines – depicting a real object – is moving as a group relative to a set of background lines which is also moving but in counter phase (Regan, 2000, see Fig. 1A). In the case of scrambled object fragments, the fragments moving in phase would be perceptually grouped into a figure which can be segregated from the background based on the Gestalt law of common fate. The re-

sulting figure, however, would not have any known contours, in fact it would have no contours at all (Fig. 1B). Perceptually, the scrambled shape-from-motion paradigm clearly leads to figure-ground segregation. It is unclear, however, how the LOC will respond to such a stimulus given that all contours are eliminated and the binding of fragments into a figure is based on motion, a non object-related cue.

Interestingly, the percept in shape-from-motion paradigms persists for a second or two after the motion stops, at least for real objects. This means that once the binding of fragments into a shape is completed, the perceptual system can briefly hold onto the perceived shape. Accordingly, measuring perceptual persistence allows us to assess the strength and disintegration of the binding process. Previously, we demonstrated that perceptual persistence in shape-from-motion

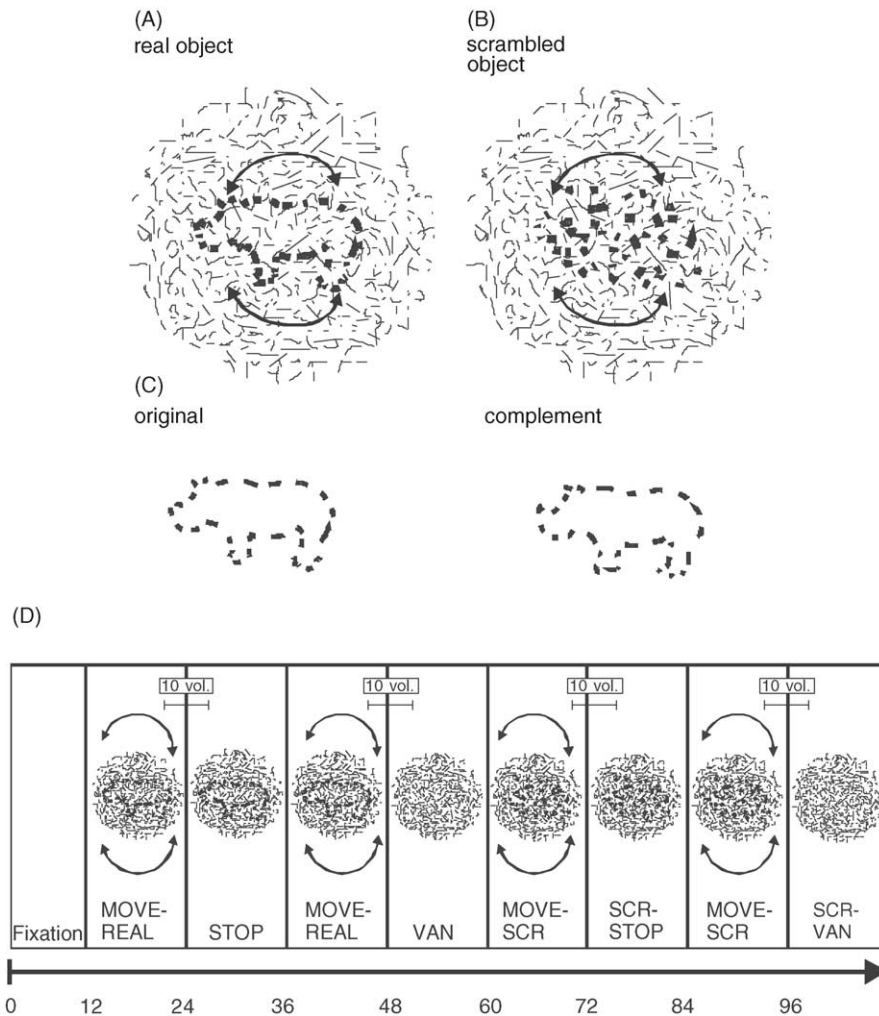


Fig. 1. Examples of stimuli, paradigm. (A) Real object: a fragmented 2D line-drawing of an animal (depicted in black) was superimposed on a background of pseudo-randomly oriented lines. The fragments covered 70% of the outline of the object. In the actual experiment, all lines – background and object – had the same thickness. Thus, the objects were essentially camouflaged in the stationary displays, and subjects could not identify a shape without motion. The arrows indicate that the fragmented line-drawing of the object rotated clockwise and counterclockwise  $\pm 15^\circ$  with a period of 2.5 s. The background also rotated but in counter phase. (B) Scrambled object: in the scrambled condition, the same objects were used as for the real objects but now the fragments were pseudo-randomly displaced. (C) Original: The line fragments covered 50% of the outline of the object. Complement: Objects representing the same shape were created by a set of complementary lines (i.e. lines over what had previously been gaps and vice versa). (D) Paradigm: In all conditions, the stimulus (real object or scrambled object) moved relative to the background for 12 s; then the motion stopped and the stimulus was either removed (vanish conditions), or it remained in the display (stop conditions).

displays depicting real objects is reflected in persisting brain activation in the LOC (Ferber, Humphrey, & Vilis, 2003). Here we asked whether the strength of the binding process is the same regardless of the presence or absence of contours. Given that the perceptual system does not have any prior representation of the shape it is unclear whether scrambled shape-from-motion displays will lead to a perceptual or neural persistence effect. We used functional magnetic resonance imaging (fMRI) to test whether the LOC can maintain the grouping of initially unbound, non-collinear surface fragments when the binding was informed by motion (see <http://psych.utoronto.ca/~ferber/flash-demo.html> for a demo). If the LOC responds to those displays and maintains its activation level after the binding-inducing cue is removed this would indicate that this area subserves figure-ground segregation processes, a relative low-level, bottom-up function in the object recognition hierarchy. Also, we collected perceptual persistence data to test whether the perceptual system discriminates between intact and scrambled objects when the binding does not rely on object-related features.

Other findings are indicating that the LOC subserves higher-order, top-down functions. For example, several studies have demonstrated that the LOC shows perceptual constancy despite changes in object size, location, viewpoint or contour, indicating that representations of shapes in the LOC are abstract (Grill-Spector et al., 1999; Kourtzi & Kanwisher, 2001; Vuilleumier, Henson, Driver, & Dolan, 2002). Perceptual constancy means that once elements have been grouped into a shape, the form can be maintained despite changes in the particulars of the display. Here, we asked to what extent perceptual constancy relies on object-sensitive information. In shape-from-motion paradigms, the grouping process is informed by motion, not by object-related features. In our previous study, the persistence of the grouping required the visual presence of the constituent fragments; in our present study, we investigate whether the grouping process can be generalized to newly presented fragments that maintain the perceived shape. Experiment 2 examined neurophysiological and perceptual persistence when the fragmented line-drawing of an object was replaced, once the motion had stopped, by a set of complementary lines representing the same stimulus shape (see Fig. 1C). Retaining the object shape but altering the composition of the elements allowed us to determine whether the motion-induced binding process has some form of spatial constancy. Neural and perceptual persistence after the motion stopped would indicate that perceptual constancy of objects can be achieved even when the binding was informed by non object-related cues.

## 2. Materials and methods

### 2.1. Participants and imaging

All subjects were scanned on a 4.0T Varian Siemens whole-body imaging system obtaining a structural T1

weighted scan for each experiment (96 slices; TE = 5.5 ms; in-plane pixel size, 0.75 mm × 0.75 mm; slice thickness, 1.25 mm) and a series of blood-oxygenation-sensitive T2\* weighted echoplanar image volumes (experiments 1 and 2: volume acquisition time: 500 ms, two shots, TE = 15 ms, FA = 22°, FOV = 19.2 cm × 19.2 cm; in-plane pixel size = 3 mm × 3 mm; slice thickness = 5 mm, five slices). We used a 15.5 cm × 11.5 cm quadrature radio frequency surface coil placed at the occipital pole to improve the signal-to-noise ratio.

Seven healthy volunteers (four men and three women) participated in experiment 1 and six healthy volunteers (five men and one woman) participated in experiment 2. All subjects (age range: 22–34 years) gave written consent and all procedures were approved by the University of Western Ontario Ethics Review Board.

### 2.2. Experimental paradigm

In all experiments, subjects viewed, through a mirror, images that were back-projected onto a screen. Subjects were instructed to maintain fixation on a centrally presented red dot. The LOC is a large area comprising several sub-regions subserving slightly different functions. We are focusing here on LO, lateral occipital (Grill-Spector, 2003). MT+ and LO were localized with independent localizer runs (for details see Ferber et al., 2003). The LO region-of-interest (ROI) was identified as a region in the lateral occipito-temporal cortex (inferior and posterior to MT+) where voxels produced a significantly higher fMRI response to intact than to scrambled 2D line-drawings of animals. Twelve line-drawings of animals were presented in each 12 s epoch at 1 s intervals. Subjects performed a one-back matching task in which they pressed a response key whenever they saw two identical images in a row. It has been shown that images of animals preferentially activate an area which is located just posterior and inferior to MT+ (Grill-Spector, 2003). This is important in the current context because the binding-inducing cue in our paradigm is motion. The MT+ ROI was identified as a region where voxels produced a significantly higher fMRI response to moving than to stationary randomly oriented lines (Tootell et al., 1995; Zeki, Watson, & Frackowiak, 1993). The display alternated between 12 s epochs of stationary or moving lines. As a group, the lines were either rotating, translating, contracting or expanding.

The localizer scans were analyzed while our subjects were still in the magnet. From those localizer scans (volume acquisition time 2 s; two shots) we selected five contiguous slices for experiments 1 and 2. All displays subtended 7.5° × 7.5° of visual angle. Each shape-from-motion experiment consisted of four scans.

In general, one set of lines (depicting the figure) was rotating clockwise and counterclockwise ± 15° with a period of 2.5 s relative to a different set of lines (background, see Fig. 1) which also rotated but in counter phase (MOVE epochs). The epochs were 12 s long and were always followed by an epoch of equal duration with stationary displays. Perceptual persis-

tence and persistence of brain activation was measured during the stationary display epochs in which the set of lines representing the figure during the MOVE epoch either remained physically present (STOP condition) or disappeared (vanish condition: VAN). Subjects indicated with a button press when, in their subjective experience, the percept of a coherent shape disappeared. Each scan began with an initial fixation period of 12 s.

In experiment 1, two different types of MOVE epochs (real objects and scrambled objects) and four different types of stationary epochs were used (see Fig. 1D). In the MOVE-REAL condition, a fragmented line-drawing depicting a real object rotated relative to the background. The line fragments covered 70% of the outline of the object. This condition was followed by either a STOP condition (stationary display in which the lines depicting the object remained in the display) or a VAN condition (stationary display in which the lines depicting the object were removed from the display the moment the motion stopped). In a different MOVE epoch, the MOVE-SCR condition, a set of pseudo-randomly oriented lines representing scrambled versions of the objects used in the MOVE-REAL conditions rotated relative to the background. This condition was followed by either a SCR-STOP condition (stationary display in which the lines depicting the same scrambled object remained in the display) or a SCR-VAN condition (stationary display in which the lines depicting the scrambled object were removed from the display the moment the motion stopped). All four stationary conditions were pseudo-randomly repeated four times within a scan.

In experiment 2, subjects were again presented with two different types of MOVE epochs (12 s long each) and four different types of stationary displays (12 s long each). In the MOVE-ORIG condition, a fragmented line-drawing depicting a 2D object rotated relative to the background. The line fragments covered 50% of the outline of the object. This condition was followed either by an O-STOP condition (stationary display in which the lines depicting the object remained in the display) or an O-VAN condition (stationary display in which the lines were removed from the display the moment the motion stopped). In an additional condition, REPLACE, which could also follow a MOVE-ORIG condition, the originally rotating line-drawing of an object was replaced by a set of complementary lines (i.e. lines over what were once gaps and vice versa, see Fig. 1C) the moment the motion stopped. This manipulation allowed us to maintain the object's perceived shape while we changed the locally available information. Furthermore, we included a MOVE-COMP condition to test whether or not the complementary line-drawings per se resulted in different patterns of perceptual persistence and/or brain activation. Thus, in the MOVE-COMP condition a complementary fragmented line-drawing of an object rotated relative to the background. The MOVE-COMP condition was always followed by a complement stop (C-STOP) condition in which the complementary fragmented line-drawing of the object presented during the

MOVE-COMP condition remained in the display after the motion stopped (see Fig. 4A).

Owing to the high temporal resolution of our shape-from-motion scans fMRI data collection was limited to five slices covering the lower half of the occipital lobes. In both experiments, we measured the perceptual persistence the observers experienced after the motion stopped in all stationary stimulus conditions by asking subjects to indicate with a button press when the percept of a coherent shape had disappeared after the motion stopped.

In a separate session, we identified V1 in four subjects participating in experiment 2 using established retinotopic horizontal and vertical visual field meridian mapping techniques (Serenio et al., 1995).

### 2.3. Behavioural data

For experiment 1, we collected behavioural pilot data from six undergraduate volunteers outside the scanner. Subjects viewed stationary displays with the same background lines as described above. Superimposed on the background was either a fragmented line-drawing depicting a real 2D animal or a scrambled version of the same stimulus. The task for the subjects was to tell whether they were able to see an object and, if so, which one. None of the six subjects could reliably differentiate between real animals and the scrambled versions when presented with stationary displays. Furthermore, none of the subjects identified a single animal correctly in stationary displays.

For experiment 2, we collected behavioural pilot data from seven undergraduate volunteers outside the scanner to show that fragmented line-drawings of objects which covered only 50% of the outline of the object still lead to significant persistence effects. Subjects viewed the same displays as described above and indicated with a button press when the percept had disappeared.

### 2.4. Imaging data analysis

We used BrainVoyager 4.9 for data analysis. No spatial or temporal smoothing procedures were applied.

We identified LO with three or four independent localizer scans for each subject. We contrasted the activation for intact objects with the activation for scrambled objects. For all voxels surviving the threshold ( $P < 10^{-4}$ ), we then identified LO as an area in the lateral occipital-temporal region. Within this area the point of peak activation was located, and a cuboid region of interest ( $9 \text{ mm} \times 9 \text{ mm} \times 10 \text{ mm}$ ) was centered around this centroid. Using this independently defined ROI, fMRI responses of experiments 1 and 2 were extracted by averaging the data from all activated voxels in this ROI ( $P < 10^{-4}$ ). The same procedure was applied to define the MT+ ROI (moving lines versus stationary lines). Thus, our ROIs were always defined according to an individual subjects' functional neuroanatomy yielding functional correspondency across subjects.



Typically, 10 volumes from the initial fixation period were used as a baseline. In case the fMRI signal did not saturate at a low level within the initial volumes and to account for drift effects, the baseline was calculated by taking the lowest fMRI signal from 10 volumes from the initial and final fixation periods.

To analyze the persistence of brain activation, we normalized the data to the activation of 10 volumes of the transition period between the MOVE epochs and the following stationary epochs (the average of the last five volumes of a MOVE epoch and the first five volumes of the following condition). Our previous study (Ferber et al., 2003) demonstrated that a robust measure of the point in time when the activation started dropping was the time when the signal had decreased by 25%. Normalizing the data individually for each subject allowed us to determine this point for each subject.

Owing to the high temporal resolution of the fMRI scans required for both experiments, we were able to collect data from only five slices centred around LO and covering MT+. In experiment 1, the five slices covered LO only in the right hemisphere in two subjects. In experiment 2, we could not include the fMRI data of one subject in the analysis because of magnet problems.

### 2.5. Area between curves

To confirm that the 25% drop measure is reliable we used a second measure for persistence, namely the area between the fMRI signal curves for VAN, STOP and SCR-STOP. All data points belonging to those epochs were included. We defined a polygon by using the fMRI signal time course data from two conditions as the vertices with time as the  $x$  coordinate and fMRI signal change as the  $y$  coordinates. The area of the polygon whose  $n$  vertices have the coordinates  $(x_i, y_i)$  for  $1 \leq i \leq n$  is defined as

$$A = \frac{1}{2} \sum_{i=1}^n (x_i y_{i+1} - x_{i+1} y_i)$$

with  $x_{n+1} = x_1$  and  $y_{n+1} = y_1$ .

These values were then compared to zero.

## 3. Results

In a pilot study we established that subjects could neither identify the line-drawings of real animals nor reliably differentiate between line-drawings of real animals and the scrambled versions of those stimuli prior to the motion-induced binding process in stationary displays.

### 3.1. Experiment 1: behavioural analyses

Experiment 1 tested the perceptual and physiological persistence of scrambled shape-from-motion displays. A repeated measures ANOVA with all four stimulus conditions

as the within subject factor and perceptual persistence as the dependent variable revealed significant differences between conditions ( $F_{[3,18]} = 22.03, P < 0.001$ ). Fig. 2B shows the perceptual persistence averaged across all seven subjects. Post hoc comparisons ( $t$ -tests for paired samples) showed that subjects perceived longer persistence in the stop condition (3449 ms) than in the vanish condition (786 ms,  $t = -4.868, P = 0.003$ ), confirming the results from our previous study (Ferber et al., 2003). The same was true for the two scrambled conditions: subjects perceived longer persistence for the scrambled stop condition (2118 ms) than for the scrambled vanish condition (785 ms,  $t = 4.313, P = 0.005$ ). However, subjects perceived longer persistence for object stop than for scrambled stop ( $t = 4.551, P = 0.004$ ). The net persistence was 2663 ms for fragmented line-drawings of real objects (STOP minus VAN) and 1333 ms for scrambled shapes (SCR-STOP minus SCR-VAN,  $P = 0.003$ ). In summary, we found that the perceptual system is capable of holding onto a nonsense figure even though this figure consists of a random arrangement of a substantial number of fragmented elements without clear contours. However, the binding process seems to be weaker for scrambled than for real objects as indicated by the significantly shorter perceptual persistence.

### 3.2. Experiment 1: fMRI analyses

The fMRI experiment tested whether or not LO supports the perceptual persistence of different stimuli, regardless of the absence of contours (i.e. an identifiable object versus a group of non-connected scrambled lines). LO responded strongly to scrambled shape-from-motion displays (averaged peak percent signal change for MOVE-SCR combined: 1.34%). However, the averaged peak percent signal change for line-drawings of real animals (MOVE-REAL combined) was even higher (1.58%,  $P = 0.006$ ), reflecting a slight preference for recognizable shapes. Fig. 3 shows the time courses of activation for all conditions (see Fig. 3A and B: LO; Fig. 3C: MT+). To compare the persistence time, we normalized the brain activation to the average of the last five volumes (volume acquisition time: 500 ms) of each condition while the object still moved (last five volumes of MOVE-REAL or MOVE-SCR) and to the first five volumes of the following epoch (see Fig. 1D), and then we determined when the signal had decreased by 25% of the normalized peak activation. The absolute time points of the fMRI signal include the haemodynamic lag and thus are longer overall than the behavioural measures of persistence. A repeated measures ANOVA with our stimulus conditions as the within subject factor and time points of the 25% signal drop as the dependent variable showed significant differences between conditions ( $F_{[3,18]} = 10.15, P < 0.001$ ). Fig. 2C shows the persistence of brain activation after the motion stopped averaged across all seven subjects. Planned post hoc comparisons revealed that the brain activation in both stop conditions persisted longer than in both vanish conditions (VAN versus STOP:  $t = -3.85, P = 0.008$ ; VAN versus SCR-STOP:  $t = -4.26, P = 0.005$ ;

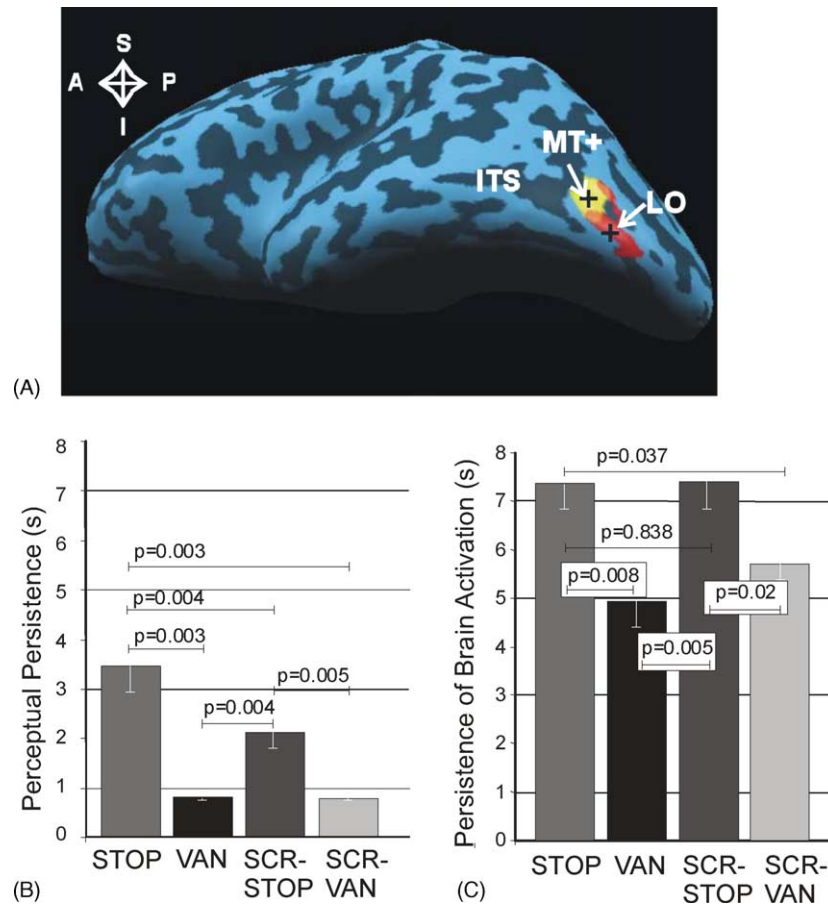


Fig. 2. ROIs, results, experiment 1. (A) We identified the LO ROI by determining the point of peak activation for 2D line-drawings depicting real objects. Then we centred a  $9 \text{ mm} \times 9 \text{ mm} \times 10 \text{ mm}$  ROI around this point (indicated by black cross). Correspondingly, we identified the MT+ complex by determining the point of peak activation for moving versus stationary lines and centred a  $9 \text{ mm} \times 9 \text{ mm} \times 10 \text{ mm}$  ROI around this point (indicated by black cross). This figure shows the inflated cortex of one subject showing MT+ and LO. A: anterior; P: posterior; I: inferior; S: superior. Functional data collected with a surface coil were aligned to high-resolution 3D anatomical images collected with a head coil. (B) Averaged perceptual persistence after the motion stopped in seconds. Subjects were asked to indicate with a button press when the percept of the object was gone after the motion stopped. STOP: the fragmented line-drawing of a real object remained in the display after the motion stopped. VAN: the line fragments of the object were removed the moment the motion stopped. SCR-STOP: the fragmented line-drawing of a scrambled object remained in the display after the motion stopped. SCR-VAN: The lines were removed the moment the motion stopped. In both stop conditions (real objects and scrambled objects) the percept persisted longer than in both vanish conditions. But even though the scrambled shape-from-motion displays resulted in significant persistence, the persistence was still shorter than persistence for real objects. Vertical bars indicate standard errors. (C) Averaged persistence of brain activation after the motion stopped. The time in seconds the signal needed to drop by 25% is plotted. Stimulus conditions are the same as (B).

SCR-STOP versus SCR-VAN:  $t = 3.15$ ,  $P = 0.02$ ; STOP versus SCR-VAN:  $t = 2.67$ ,  $P = 0.037$ ). These results suggest that LO can not only maintain the grouping of line fragments that form the outline of real objects, but can also maintain the grouping of randomly distributed line fragments that form unrecognizable novel shapes without contours. Interestingly, we found that the brain activation in the object stop condition does not persist longer than the brain activation in the scrambled object stop condition ( $t = -0.213$ ,  $P = 0.838$ ). Thus, the longer perceptual persistence in the object stop condition was not accompanied by longer persistence of brain activation.

Furthermore, we used a second measure for persistence to demonstrate that the reported findings do not simply depend on this particular measure of the 25% signal decline. We analyzed the area between the curves of the fMRI signal after the motion stopped (see Section 2). In particular, we tested

whether the areas between the VAN and STOP condition, between SCR-VAN and SCR-STOP, and between STOP and SCR-STOP differed significantly from zero. Confirming our previous observation, we found that the areas between the curves for VAN and STOP ( $t = 4.883$ ,  $P = 0.003$ ) and between the curves for SCR-VAN and SCR-STOP ( $t = 3.87$ ,  $P = 0.008$ ) differed significantly from zero whereas the area between the curves for STOP and SCR-STOP did not differ significantly from zero ( $t = -0.273$ ,  $P = 0.794$ ).

In order to address a possible confound with attention, we also analyzed the persistence of brain activation in the MT+ complex. It has been shown that the motion aftereffect in the MT+ complex disappears when attention is controlled for (Huk, Ress, & Heeger, 2001). According to this logic, the observed persistence effect in LO could simply reflect the subjects' mental effort such that in both vanish conditions

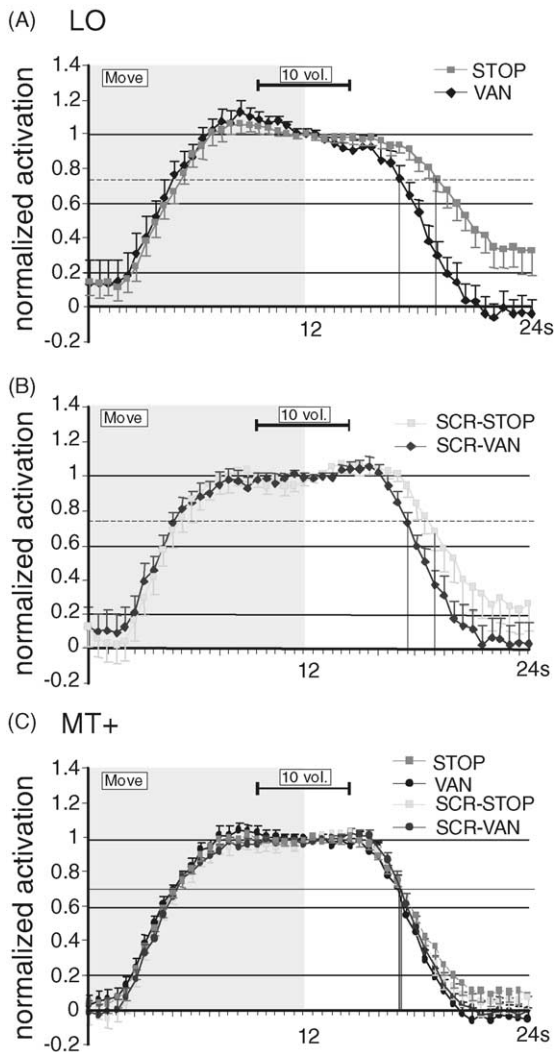


Fig. 3. Time course data, experiment 1. (A) Normalized average fMRI signal (y axis) plotted against time (x axis) in LO. The object rotated for 12 s (shaded epoch). The time course was normalized to the last five volumes (or 2.5 s) of the preceding MOVE epoch and the following five volumes after the motion stopped. Vertical bars indicate standard errors. The dashed line indicates a signal drop by 25%. (B) Normalized average time courses for scrambled objects. Axes same as above. (C) Normalized average fMRI signal (y axis) plotted against time (x axis) in MT+. Axes same as above.

subjects may stop attending immediately after the motion stopped while they briefly maintain attention for both stop conditions. The time courses for all four different conditions for the MT+ complex is given in Fig. 3C. A repeated measures ANOVA with our stimulus conditions as the within subject factor and time points of the 25% signal drop as the dependent variable did not reveal any significant differences between conditions ( $F_{[3,18]} = 1.40$ ,  $P = 0.275$ ).

Experiment 2 tested whether or not the perceptual persistence and the persistence of activation in LO relies on the very lines that were originally segregated by motion. In other words, can we find persistence after the original line fragments are removed and replaced by different fragments representing the same perceived shape?

### 3.3. Experiment 2: behavioural analyses

The response times of the six subjects who participated in the fMRI experiment can be seen in Fig. 4B. Not surprisingly, we found in an initial *t*-test for paired samples that the O-STOP condition (1962 ms, S.D. = 780.4) and the C-STOP condition (1979 ms, S.D. = 1058.9;  $t = -0.072$ ,  $P = 0.946$ ) resulted in virtually identical percepts. Thus, we combined the data from the two conditions for further analysis (O/C-STOP: 1971 ms, S.D. = 884.2). A repeated measures ANOVA with the remaining three stimulus conditions (O/C-STOP, REPLACE, O-VAN) as the within subjects factor revealed significant differences ( $F_{[2,10]} = 11.38$ ,  $P = 0.003$ ). O/C-STOP resulted in significantly longer perceptual persistence than O-VAN (701 ms, S.D. = 114.3;  $t = -3.651$ ,  $P = 0.015$ ), again confirming our initial observation. The comparison between REPLACE and O-VAN revealed that even when the original was replaced by its complement, the percept of a given object persisted (1525 ms, S.D. = 690.8;  $t = -2.973$ ,  $P = 0.031$ ). However, it was evident that replacing the rotating original by its complement shortened the persistence observers experienced when compared to O/C-STOP ( $t = -3.067$ ,  $P = 0.028$ ).

### 3.4. Experiment 2: fMRI analyses

Again, we used fMRI to determine how long the brain activation persisted in LO after the motion stopped by comparing the time points for the 25% signal drop (see Fig. 4C). A *t*-test for paired samples showed that the persistence of the fMRI signal in the O-STOP condition did not differ from the persistence in the C-STOP condition ( $t = -0.93$ ,  $P = 0.405$ ). Thus, the data points from both conditions were combined for further analysis (O/C-STOP). A repeated measures ANOVA showed significant differences between the three remaining conditions ( $F_{[2,8]} = 12.923$ ,  $P = 0.003$ ). Corresponding to the results from the behavioural study, we found that O/C-STOP resulted in longer persistence of brain activation than O-VAN ( $t = 4.108$ ,  $P = 0.015$ ). By the same token, the comparison between REPLACE and O-VAN revealed that even when the original was replaced by its complement, the fMRI signal in LO persisted ( $t = -3.72$ ,  $P = 0.02$ ). However, contrary to the behavioural results, we did not find any differences in the persistence of brain activation between O/C-STOP and REPLACE ( $t = 0.389$ ,  $P = 0.717$ ).

In addition, we analyzed the area between the curves of the normalized fMRI signal after the motion stopped (see Fig. 5). In particular, a *t*-test for paired samples revealed that the areas between the VAN and O-STOP and VAN and C-STOP did not differ significantly ( $t = -1.382$ ,  $P = 0.239$ ). Therefore and in line with the above analysis, O-STOP and C-STOP were combined for further analysis (O/C-STOP). Next, we found that the area between the curves for VAN and O/C-STOP did differ significantly from zero ( $t = 2.604$ ,  $P = 0.03$ , one-tailed) confirming our previous observation and indicating persistence. Also, the area between the curves

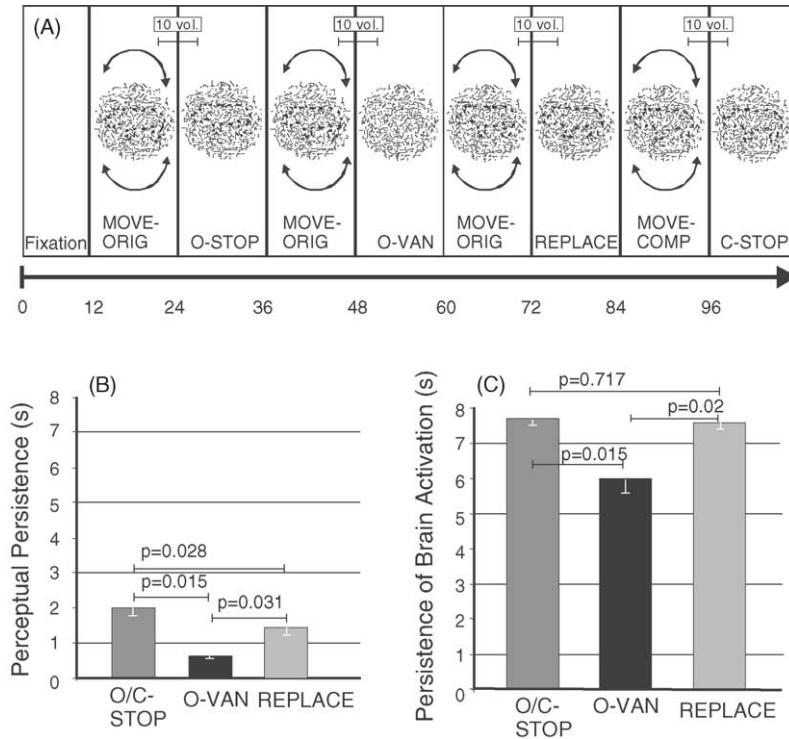


Fig. 4. Paradigm, results, experiment 2. (A) Paradigm: In all MOVE conditions, the stimulus (original object or its complement) was moving relative to the background for 12 s; then the motion stopped and the stimulus was either removed (O-VAN condition), or it remained in the display (O-STOP or C-STOP conditions), or it was replaced by its complement (REPLACE). (B) Averaged perceptual persistence after the motion stopped in seconds. O/C-STOP: the fragmented lines of the original object or the complement object remained in the display. O-VAN: the fragmented lines of the object were removed the moment the motion stopped. REPLACE: the moment the motion stopped, the fragmented line-drawing of an object was replaced by a set of complementary lines representing the same object. (C) Averaged persistence of brain activation after the motion stopped. The time in seconds the signal needed to drop by 25% is plotted. Stimulus conditions are the same as (B).

of VAN and REPLACE differed significantly from zero ( $t = 2.306, P = 0.041$ , one-tailed). Finally, a  $t$ -test for paired samples for the area between the curves of VAN-REPLACE and VAN-O/C-STOP revealed no significant differences

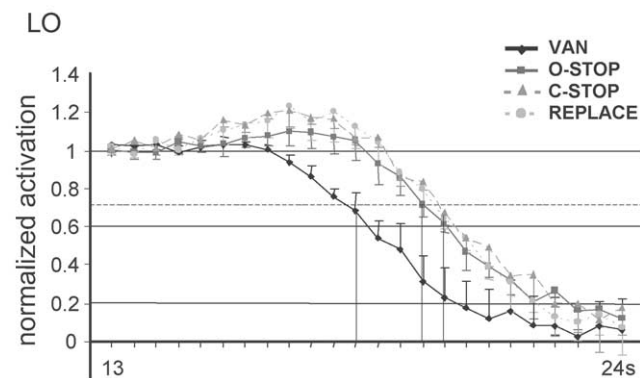


Fig. 5. Time course data, experiment 2. Normalized fMRI signal (y axis) in LO averaged across all subjects plotted against time (x axis). The object stopped rotating at point 12 s. Depicted is the time course of activation during stationary epochs when the fragmented line-drawing was either removed (VAN), still physically present (O-STOP, C-STOP) or replaced by a set of complementary lines representing the same object (REPLACE). Vertical bars indicate standard errors. The dashed line indicates a signal drop by 25%.

( $t = -0.042, P = 0.969$ ) indicating that the time courses of activation for REPLACE and O/C-STOP behave similarly after the motion stops.

Furthermore, we calculated the areas between curves for non-normalized time courses in V1. We found that the area between the curves for VAN and O-STOP did not differ significantly from zero ( $t = 0.794, P = 0.485$ ); the same was true for the area between the curves for VAN and C-STOP ( $t = 1.517, P = 0.226$ ). Also, the area between the curves of VAN and REPLACE did not differ significantly from zero ( $t = 1.351, P = 0.27$ ).

#### 4. Discussion

We showed that an object-sensitive area in the ventral stream, LO, can bind and maintain the grouping of fragmented elements into a figure when these fragments are segregated from the background by motion. Even more importantly, LO subserves this binding process even in the absence of contours or closed-loop shapes. Furthermore, this motion-induced binding process extends beyond the initially grouped features and applies to new fragments presented after the motion stopped.



The previously reported decrease in neuronal activation in response to scrambled displays (Grill-Spector et al., 1998; Vogels, 1999) may at first seem contrary to our finding of robust persisting activation in LO after the motion stopped in both experiments. This inconsistency vanishes if one assumes that the function of LO is to subserve the grouping process and the figure-ground segregation. The previously reported decrease in neuronal activation in LO that occurs when images are made more scrambled may in fact reflect a decline in the grouping process. In a completely scrambled static image, there is nothing to initiate the grouping process. The grouping process is initiated when a cue that differentiates one set of elements from the rest, such as motion, is introduced. By measuring visual persistence, one obtains a measure of this grouping process.

Our findings of robust persisting activation in response to scrambled images are of particular interest in the context of the recent observations that the LOC responds to salient regions with perceived contours (Stanley and Rubin, 2003) and illusory contours (Mendola et al., 1999) with a similar level of activation. Here we show that perceived or illusory contours are not required to activate this brain region. Furthermore, we extend these findings by showing that even when the figure-ground segregation process is informed by motion it is still maintained by LO. This indicates that the LOC is involved in the relatively early stages of selecting candidate portions and salient features of the visual array for further processing.

Our second study showed that LO then stores – at least briefly – the motion-induced percept, but does not store the elements themselves. Recently, an adaptation paradigm was used in an elegant fMRI study to show that the neural response in the LOC decreased for the second stimulus in a pair of stimuli possessing the same perceived shape (Kourtzi & Kanwisher, 2001). Such a finding indicates that the LOC is sensitive to the sameness of the perceived shape and not to the particular features of the image. This adaptation effect is not unique to humans. Stimulus repetition results in decremental responses of neurons in the monkey's inferior temporal cortex (Fahy, Riches, & Brown, 1993). Even though these findings indicate that areas subserving perceptual processes store representations of stimuli, they do not address how these areas are involved in the on-line maintenance of a stimulus after it has been removed from view. When the original line-drawing in our experiment was removed and replaced by its complement, LO maintained its activation level. It is important to note in this context that the fragments of the complement object were only presented after the motion stopped and accordingly, did not move previously. Still, these new fragments subserve the maintenance of the shape. The motion-induced binding process is not limited to the very lines presented, but extends some distance around the original line fragments, thereby capturing neighbouring areas. Given that neurons in visual areas above V1 have relatively large receptive fields, one might argue that the prolonged fMRI activation in LO after the motion stopped may simply reflect the coding properties of the large receptive fields in neurons

in this brain area. While we believe that large receptive fields are a pre-requisite for binding processes extending beyond the physically presented features, we also believe that this account could not explain how the percept can be maintained in the first place. If *all* lines surrounding a given fragment were selected then the percept should disappear the moment the motion stopped because the displays contained more lines belonging to the background than to the figure, and all lines would essentially be weighted equally. Given that observers experience persistence and that LO shows persisting brain activation, it appears that LO binds only those elements with orientations approximately matching the orientation of the original elements. In this way, the binding process maintains a more abstract representation of the perceived shape. Our results show that LO can temporarily maintain this binding process even in the absence of the binding-inducing cue.

We did not find significant persistence of the fMRI signal in V1. This finding indicates that LO does not require feedback from V1 during the maintenance phase. This finding can be reconciled with recent observations regarding the processing of global shapes in V1 (Fitzpatrick, 2000 for review, Kourtzi, Tolias, Altmann, Augath, & Logothetis, 2003) when one considers that multiple randomly oriented lines placed in the receptive field surround of monkeys' V1 cause a decrease in a cell's response to an optimally oriented stimulus (Kapadia, Ito, Gilbert, & Westheimer, 1995). This situation is given after the motion stops in our displays. Even though the optimal stimulus is still physically present it cannot be distinguished anymore from random background noise. Also, in conjunction with absence of persistence in MT+ the absence of persistence in V1 provides further evidence against the idea that the observed persistence in LO is entirely subserved by an attentional effect. The binding-inducing cue in our paradigm is motion, accordingly one could argue that attention should be drawn to the motion. Importantly, the binding is maintained after the motion cue is removed. Accordingly, we argue that while attention is necessary to initiate the binding, attention is not required to maintain it. Moreover, based on the results of our scrambled shape-from-motion experiment we speculate that this storage mechanism is time-limited rather than capacity-limited. Furthermore, based on the time course of persistence we speculate that this storage-mechanism serves as a mediator between the iconic store and short-term memory.

Our interpretation that LO is subserving the persistence of the percept is in line with another recent observation that the LOC is critically involved in the maintenance of object-related percepts (Kleinschmidt, Büchel, Hutton, Friston, & Frackowiack, 2002). The authors correlated brain activation with the perceptual persistence of a percept in the face of parameter changes that favoured an alternative percept (perceptual hysteresis). The trial-by-trial event-related analysis revealed a relative activity increase in the LOC in trials with perceptual hysteresis. In other words, the prolonged maintenance of the percept was accompanied by an enhanced signal. This finding indicates that the LOC is involved in the main-

tenance of a percept even when the available sensory input favours a rivaling pattern.

Object recognition, however, requires more than that. The percept must be compared with stored visual representations and eventually needs to be associated with a meaning. These processes are most likely subserved by the pre-frontal cortex and structures in the medial temporal lobes. Neurons in the inferior temporal cortex are directly or indirectly connected with early visual areas, the medial temporal lobe and pre-frontal areas. If the multi-staged object recognition process relies on feedback with iterative processing, then LO is ideally located and equipped to serve this function. However, it is interesting to note in this context that both experiments showed that the pattern of fMRI activations in LO did not follow the pattern of perceptually driven behavioural responses. One might argue that the limited temporal resolution of fMRI – even with our volume acquisition time of 500 ms – does not allow for the detection of such differences. This argument might hold true for experiment 2, in which the difference between the duration of the percepts in the two relevant conditions was less than 500 ms. In experiment 1, however, the behavioural difference between scrambled and real objects was more than a second. The difference between the perceptual persistence of scrambled object stop and scrambled object vanish was in the same range and the differences in persistence of brain activation clearly reflected the behavioural result. Thus, we believe that the temporal resolution of our fMRI scans was sufficient. Future studies using MEG or ERPs could elucidate the issue of top-down modulation in LO further. Given that the scrambled displays and the REPLACE displays led to shorter perceptual persistence which was not reflected in differences in persistence in LO one could argue that while LO maintains the binding process, persistence in frontal brain areas will be more closely correlated to the subjective percept. The important point in the context of this paper is, however, that we observed persistence both behaviourally and with fMRI for objects without contours and for new object fragments.

Functionally, LO is defined by subtracting the activation in response to images of scrambled objects from the activation in response to images of intact objects – in particular, images of animals (Chao, Haxby, & Martin, 1999; Grill-Spector, 2003). Anatomically, LO can be located as an area posterior and inferior to MT+. This anatomical proximity allows swift shape-from-motion perception. Also, infants as young as 4 months use common motion to group objects together into units (Kellman & Spelke, 1983) and are more likely to use shape information than colour cues to find object boundaries (Needham, 1999). Infants do not have complete myelination of neuronal projections between cortical areas, and it is reasonable to assume that projections that are needed the most and are not costly to maintain are myelinated first. It would be interesting to test whether infants respond to scrambled shape-from-motion and intact shape-from-motion in the same way.

In summary, the time course of activation we observed in LO corresponds with the view that LO perceptually analyzes the given visual information and is also involved in the retention of the stimulus. Importantly, LO does neither require contours nor object-sensitive binding cues. Regardless of the nature of the binding process, once a figure is segregated from the background LO can maintain its representation even in the absence of any contours or the fragments which initially informed the binding. This indicates that LO is involved in the intermediate stages of object recognition and has some mnemonic function, potentially as an intermediate buffer between sensory memory and short-term memory. The persisting activation might mediate between object recognition, maintenance of information, learning and decision-making. This finding provides evidence that sustaining a visual percept involves perceptual brain areas, which could also support the idea that the neural correlates of awareness of a given perceptual attribute rely on neural structures that analyze the perceptual information but that these neural correlates are not sufficient for awareness.

### Acknowledgments

This research was supported by a grant from the Canadian Institutes of Health Research awarded to T.V. and a grant from the Natural Science and Engineering Research Council of Canada awarded to S.F. The authors would like to thank Jody Culham, Kathy O'Craven and Adam K. Anderson for comments on an earlier version of the manuscript. We would like to acknowledge Stephen Emrich for his help with the pilot study for experiment 1.

### References

- Chao, L. L., Haxby, J. V., & Martin, A. (1999). Attribute-based neural substrates in temporal cortex for perceiving and knowing about objects. *Nature Neuroscience*, *2*, 913–919.
- Corbetta, M., Miezin, F. M., Dobmeyer, S., Shulman, G. L., & Petersen, S. E. (1990). Attentional modulation of neural processing of shape, color, and velocity in humans. *Science*, *22*, 1556–1559.
- Fahy, F. L., Riches, I. P., & Brown, M. W. (1993). Neuronal activity related to visual recognition memory: Long-term memory and the encoding of recency and familiarity information in the primate anterior and medial inferior temporal rhinal cortex. *Experimental Brain Research*, *96*, 457–472.
- Faillenot, I., Toni, I., Decety, J., Grégoire, M. C., & Jeannerod, M. (1997). Visual pathways for object-oriented action and object recognition: Functional anatomy with PET. *Cerebral Cortex*, *7*, 77–85.
- Ferber, S., Humphrey, G. K., & Vilis, T. (2003). The lateral occipital cortex subserves the perceptual persistence of motion-defined groupings. *Cerebral Cortex*, *13*, 716–721.
- Fitzpatrick, D. (2000). Seeing beyond the receptive field in primary visual cortex. *Current Opinion in Neurobiology*, *10*, 438–443.
- Grill-Spector, K., Kushnir, T., Hendler, T., Edelman, S., Itzhak, Y., & Malach, R. (1998). A sequence of object-processing stages revealed by fMRI in the human occipital lobe. *Human Brain Mapping*, *6*, 316–328.

- Grill-Spector, K., Kushnir, T., Edelman, S., Avidan, G., Itzhak, Y., & Malach, R. (1999). Differential processing of objects under various viewing conditions in the human lateral occipital complex. *Neuron*, *24*, 187–203.
- Grill-Spector, K., Kushnir, T., Hendler, T., & Malach, R. (2000). The dynamics of object-selective activation correlate with recognition performance in humans. *Nature Neuroscience*, *3*, 837–843.
- Grill-Spector, K., Kourtzi, Z., & Kanwisher, N. (2001). The lateral occipital complex and its role in object recognition. *Vision Research*, *41*, 1409–1422.
- Grill-Spector, K. (2003). The neural basis of object perception. *Current Opinion in Neurobiology*, *13*, 159–166.
- Halgren, E., Dale, A. M., Sereno, M. I., Tootell, R. B. H., Marinkovic, K., & Rosen, B. R. (1999). Location of human face-selective cortex with respect to retinotopic areas. *Human Brain Mapping*, *7*, 29–37.
- Huk, A. C., Ress, D., & Heeger, D. J. (2001). Neuronal basis of the motion after effect reconsidered. *Neuron*, *32*, 61–172.
- James, T. W., Humphrey, G. K., Gati, J. S., Menon, R. S., & Goodale, M. A. (2000). The effects of visual object priming on brain activation before and after recognition. *Current Biology*, *10*, 1017–1024.
- Kapadia, M. K., Ito, M., Gilbert, C. D., & Westheimer, G. (1995). Improvement in visual sensitivity by changes in local context: Parallel studies in human observers and in V1 of alert monkeys. *Neuron*, *15*, 843–856.
- Kanwisher, N., Chun, M. N., Mc Dermott, J., & Ledden, P. J. (1996). Functional imaging of human visual recognition. *Cognitive Brain Research*, *5*, 55–67.
- Kellman, P. J., & Spelke, E. S. (1983). Perception of partly occluded objects in infancy. *Cognitive Psychology*, *15*, 483–524.
- Kleinschmidt, A., Büchel, C., Hutton, C., Friston, K. J., & Frackowiak, R. S. J. (2002). The neural structures expressing perceptual hysteresis in visual letter recognition. *Neuron*, *34*, 659–666.
- Kraut, M., Hart, J., Jr., Soher, B. J., & Gordon, B. (1997). Object shape processing in the visual system evaluated using functional MRI. *Neurology*, *48*, 1416–1420.
- Kourtzi, Z., & Kanwisher, N. (2001). Representation of perceived object shape by the human lateral occipital complex. *Science*, *293*, 1506–1509.
- Kourtzi, Z., Tolias, A. S., Altmann, C. F., Augath, M., & Logothetis, N. K. (2003). Integration of local features into global shapes: Monkey and human fMRI studies. *Neuron*, *37*, 333–346.
- Malach, R., Reppas, J. B., Benson, R. R., Kwong, K. K., Jiang, H., Kennedy, W. A., et al. (1995). Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. *Proceedings of the National Academy of Science of the United States of America*, *92*, 8135–8139.
- Mendola, J. D., Dale, A. M., Fischl, B., Liu, A. K., & Tootell, R. B. H. (1999). The representation of illusory and real contours in human cortical visual areas revealed by functional magnetic resonance imaging. *Journal of Neuroscience*, *19*, 8560–8572.
- Needham, A. (1999). The role of shape in 4-month-old infants' object segregation. *Infant Behavior and Development*, *22*, 161–178.
- Regan, D. (2000). *Human perception of objects*. Sunderland, MA: Sinauer Associates.
- Sereno, M. I., Dale, A. M., Reppas, J. B., Kwong, K. K., Belliveau, J. W., Brady, T. J., et al. (1995). Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science*, *268*, 889–893.
- Sergent, J., Ohta, S., & MacDonald, B. (1992). Functional neuroanatomy of face and object processing. A positron emission tomography study. *Brain*, *115*, 15–36.
- Stanley, D. A., & Rubin, N. (2003). fMRI activation in response to illusory contours and salient regions in the human lateral occipital complex. *Neuron*, *37*, 323–331.
- Tootell, R. B. H., Reppas, J. B., Kwong, K. K., Malach, R., Born, R. T., Brady, T. J., et al. (1995). Functional analysis of human MT+ and related visual cortical areas using magnetic resonance imaging. *Journal of Neuroscience*, *15*, 3215–3220.
- Vogels, R. (1999). Effect of image scrambling on inferior temporal cortical responses. *NeuroReport*, *10*, 1811–1816.
- Vuilleumier, P., Henson, R. N., Driver, J., & Dolan, R. J. (2002). Multiple levels of visual object constancy revealed by event-related fMRI of repetition priming. *Nature Neuroscience*, *5*, 491–499.
- Zeki, S., Watson, J. D., & Frackowiak, R. S. (1993). Going beyond the information given: The relation of illusory visual motion to brain activity. *Proceedings of the Royal Society of London*, *252*, 215–222.