Neuropsychologia 47 (2009) 1261–1268

Contents lists available at ScienceDirect

Neuropsychologia

journal homepage: www.elsevier.com/locate/neuropsychologia

Neural adaptation for novel objects during dynamic articulation

John A. Pyles*, Emily D. Grossman

Center for Cognitive Neuroscience & Department of Cognitive Sciences, University of California, Irvine, California 92697-5100, United States

ARTICLE INFO

Article history:
Received 23 October 2008
Received in revised form 18 December 2008
Accepted 7 January 2009
Available online 17 January 2009

Keywords:
Object perception
Ventral temporal cortex
fMRI
fMR-adaptation

ABSTRACT

Human observers readily identify objects with moving parts, and recognize their underlying structure even when the component parts undergo complex movement. This suggests the existence of neural representations that are invariant to motion and state of articulation, which together allow our visual system to maintain 'object constancy'. Ventral temporal cortex has previously been implicated in object perception and in coding object identity, but it is unclear where this is achieved when objects undergo motion-driven shape changes. In the present study, we use fMRI adaptation to probe the neural response properties when subjects view dynamic novel objects. Our results reveal neural selectivity for novel objects in the LOC region of the occipito-temporal lobe, even when those objects are viewed as moving and articulating. We also identify a bilateral area of posterior fusiform outside of the LOC with neural populations invariant to changes in the articulatory state of an object, a critical feature of object constancy. These results demonstrate the functional importance of ventral temporal cortex in the perception of moving objects, and the existence of neural populations coding for object constancy across movement and articulation.

1. Introduction

Human observers are readily able to recognize objects, even when seen from different viewpoints or when the object is moving. However, in natural viewing conditions objects often change form over time, such as trees blowing in the breeze, people walking about, and the movements of various machines (cranes, clocks, etc.). To build a stable representation of articulating objects requires a neural population that recognizes the underlying structure of an object across the various changes in shape it may undergo. Where is this accomplished in visual cortex?

Neurophysiological and neuroimaging studies typically implicate cortical areas in ventral temporal and occipital cortex as critical for object recognition, namely the lateral occipital complex (LOC, e.g. Grill-Spector, Kushnir, Edelman, Itzchak, & Malach, 1998; Grill-Spector, Kourtzi, & Kanwisher, 2001). The LOC and adjacent regions of human and monkey occipito-temporal cortex have been found to contain neural populations invariant to (i.e. robust against changes in) manipulations of object size, position, depiction (e.g. photograph versus drawing), and to a lesser extent viewpoint (Desimone, Albright, Gross, & Bruce, 1984; Ewbank, Schluppeck, & Andrews, 2005; Grill-Spector et al., 1999; Ito, Tamura, Fujita, & Tanaka, 1995; James, Humphrey, Gati, Menon, & Goodale, 2002; Kourtzi & Kanwisher, 2000; Sary, Vogels, & Orban, 1993; Tanaka, 2003). These object-tuned neuronal ensembles likely form the basis for object representations that are stable across many of the perspective changes associated with natural viewing conditions in the real world (e.g. Logothetis & Sheinberg, 1996; Riesenhuber & Poggio, 2002; Wallis & Bulthoff, 1999).

It is unclear, however, what parts of the brain support stability in the neural representations when object motion is accompanied by articulation. Articulation, specifically the movement of object parts relative to the object whole, is a problem for object representations that are form-based or viewpoint dependent. Accurate object identification during articulation requires a stable representation of object structure that ignores the accompanying changes in form over time. To the extent that human observers can easily recognize and discriminate articulating objects, there should be some neural populations that are invariant across these changes.

In this study, we use fMR-adaptation to identify neural populations invariant to changes in object shape that occur during the articulation of novel objects (Framsticks). In previous studies, these Framsticks have been found to activate a region of posterior ventral temporal cortex along the inferior occipital sulcus (Pyles, Garcia, Hoffman, & Grossman, 2007), implicating similar cortical regions as other studies of novel and common object recognition (Ewbank et al., 2005; James et al., 2002; Sarkheil, Vuong, Bulthoff, & Noppeney, 2008; Weigelt, Kourtzi, Kohler, Singer, & Muckli, 2007). In two experiments, we measure neural adaptation to changes in Framstick articulation, either viewed as animated sequences or as stationary frames (static snapshots, or poses, selected from the animations). We find regions of occipito-temporal cortex that adapt to

* Corresponding author. Tel.: +1 949 824 8378; fax: +1 949 824 2307.
E-mail address: jpyles@uci.edu (J.A. Pyles).

0028-3932/$ – see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.neuropsychologia.2009.01.006

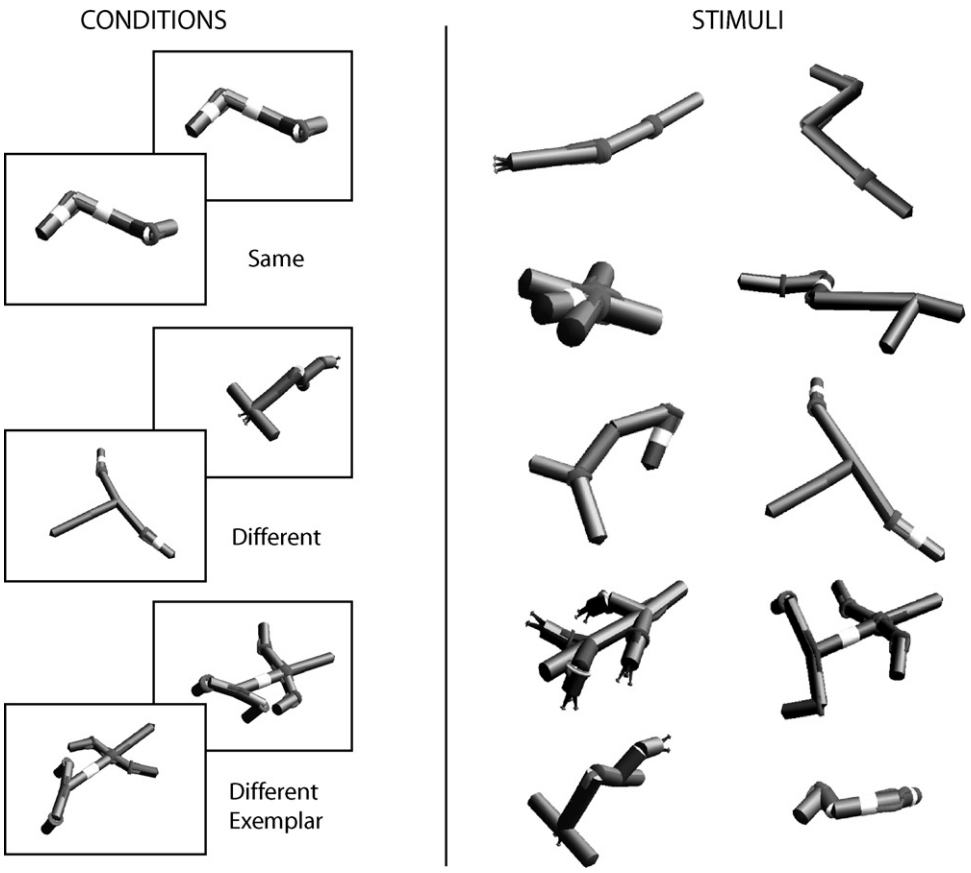

Fig. 1. Schematic depicting the Same, Different, and Different Exemplar conditions of the experiment, and single frame snapshots of the Framstick objects used in these experiments.

repeated exposures to these objects, when viewed both as stationary and while moving, but we only find evidence for adaptation across different examples of the same object when the Framstick is viewed in motion. Notably, this region with neural responses invariant to state of articulation is adjacent to, but not within, the brain area often identified as critical for object recognition (the LOC). Together these results support the idea that articulating movements facilitate stable representations of object structure, and suggest that regions in the ventral stream outside of the LOC may have mechanisms specialized for the perception of moving objects.

2. Materials and methods

2.1. Participants

Nine observers (5 males, 4 females) participated in these experiments and were financially compensated for their time. All participants had normal or corrected to normal vision, and gave informed consent as approved by the University of California, Irvine Institutional Review Board.

2.2. Stimuli

Novel, articulating objects were generated using Framsticks artificial evolution and life simulation software (Fig. 1) (Komosinski & Ulatowski, 1999; Pyles et al., 2007). As a result of the artificial evolution process, each object has a unique structure and pattern of articulation. The novel nature of the stimuli should mean that the objects are not representative of any existing category, nor should they have any name associated with them.

Image sequences of the articulating objects were exported from the Framsticks software and converted to grayscale. The stationary condition of the experiments depicted individual frames taken from the image sequences, while the animated condition depicted articulation throughout an entire 2 s image sequence (displayed as a movie file using Quicktime software). For both conditions, stimuli subtended approximately 6° of visual angle, and during articulation the objects traversed as much as 12° of visual angle.

Stimuli were displayed with Matlab and the Psychophysics Toolbox (Brainard, 1997) controlled by a Macintosh G4 computer. In the laboratory, these stimuli were viewed on a ViewSonic CRT. In the scanner the images were projected via a Christie DLV1400-DX DLP projector onto a screen at the head end of the bore, which participants viewed with a periscope mirror mounted on the headcoil.

2.3. MRI acquisition

Scanning was conducted on the 3T Philips Achieva MR scanner at the University of California, Irvine, equipped with an 8 channel SENSE imaging headcoil. High-resolution anatomical scans were acquired for each observer (T1-weighted 3D MPRAGE, 1 mm^3, TR = 8.4 ms, TE = 3.7 ms, flip = 8°, SENSE factor = 2.4) and used for coregistration of the functional data. Subjects participated in two types of functional scans (both T2-weighted, gradient echo imaging, SENSE factor = 1.5, right-left phase encoding, TE = 30 ms, flip = 90°). The rapid event-related scans were designed to most accurately estimate the amplitude of the hemodynamic response functions. These fMR-adaptation scans were acquired rapidly with slices positioned across the temporal occipital cortex (TR = 1 s, 17 axial slices, 2.40 mm × 2.44 mm × 4 mm voxels, 1 mm gap). Localizer scans were designed to identify regions of interest (ROIs) in the lateral occipital cortex, and were acquired using more standard, whole brain imaging parameters (28 axial slices with 2.05 mm × 2.05 mm × 4 mm voxels, no gap, TR = 2 s).

2.4. Procedure

Neural adaptation was measured in a rapid event-related design in which subjects viewed one of three trial types: (1) Same condition, in which participants were presented with the same object twice, (2) Different condition, in which participants viewed two different objects, or (3) Different Exemplar condition, in which subjects viewed two different depictions of the same object, shown in a different state of articulation. For each of these trials, the object was either stationary or animated, and scans were blocked by stimulus type. Static images were presented for 950 ms each, while animated sequences were presented for 2000 ms. Stimuli in each pair were presented in rapid succession, with 100 ms interstimulus intervals in the stationary scans, and no interstimulus interval for the animated pairs. For each trial, participants indicated if the two objects were the same or different by pressing one of two buttons on an MR compatible button pad. The "same" response was correct for both the Same and Different Exemplar conditions.

Table 1
Talairach coordinates of the centers of mass areas of significant adaptation in the cortex-aligned whole brain analysis. Significant areas were back projected from the cortex aligned surface reconstruction onto a representative subject's brain in Talaraich space.

ROI	Talairach coordinates					
	Left			Right		
	X	*Y*	*Z*	*X*	*Y*	*Z*
Repeated still adaptation						
LO	−32	−81	8	27	−79	5
pFs	−39	−57	−11	34	−56	−11
DLPFC	−42	10	29	41	16	30
IFG	−36	21	8	40	22	4
Repeated animated adaptation						
LO	−29	−86	−9	41	−71	−6
pFs	−35	−64	−11	34	−50	−13
DLPFC	−41	11	29	40	4	35
IFG	−35	20	9	38	23	6
Animated different exemplar adaptation						
FG	−21	−79	−12	18	−90	−4

Abbreviations: LO, lateral occiptal; pFs, posterior fusiform; DLPFC, dorsolateral prefrontal cortex; IFG, inferior frontal gyrus; FG, fusiform gyrus.

Each scan contained ten trials of each type (Same, Different, or Different Exemplar, all either stationary or animated). All trial pairs were separated by a minimum of 3 s. The trial order was determined in advance so as to maximize efficiency for extracting the underlying hemodynamic response functions (Serences, 2004). To achieve this, ten 4 s fixation trials were pseudorandomized within each scan, which served to create an exponential distribution of fixation periods and effectively jitter the intertrial interval. In addition, the initial 14 s and final 12 s of each scan were passive fixation intervals. In total each animated scan lasted 276 s, and each stationary scan lasted 216 s. Observers completed 8 scans for each of the animated and stationary experiments, for a total of 16 rapid event-related scans and 80 estimates per condition per subject, acquired over two scanning sessions on different days.

Subjects also participated in two localizer scans to identify the lateral occipital complex. The LOC was localized using a block design alternating 20 s blocks of stationary objects with 20 s blocks of scrambled objects, separated by 6 s of fixation (Grill-Spector et al., 2001; Kourtzi & Kanwisher, 2000). Stimuli consisted of a set of grayscale images of common household objects, contrasted with scrambled versions of the same images. Participants performed a one-back task to maintain attention. Seven blocks of each condition were viewed over the course of a 378 s scan.

2.5. MRI preprocessing

Preprocessing and analysis was performed using BrainVoyager QX (Brain Innovations, Inc.). The following was applied to functional data: (1) slice-timing correction, (2) 3D motion correction, and (3) linear trend removal and high-pass filtering (3 cycles per scan). Runs with significant fluctuations in overall mean were corrected for mean intensity shifts. All runs were manually co-registered to the high-resolution anatomical with the aid of high-resolution inplane scans acquired in the same locations as the functional slices.

Gray matter was manually segmented from white matter in all subjects' high-resolution anatomies and used to reconstruct the cortical surface, which was then smoothed and inflated. To provide the best possible alignment between subjects, and to maximize the correspondences across subjects in the group statistical analyses, the intersubject data was aligned into a standardized space using a cortical alignment approach (Formisano, Esposito, Di Salle, & Goebel, 2004; Goebel, Esposito, & Formisano, 2006). In this approach reconstructions of each subject's cortical surface are morphed onto a standardized sphere and then the sulcal/gyral patterns from all subjects are aligned to each other based on the best possible matching of all the cortical folding patterns. The identified regions of activation from the group statistical analyses (described below) are reported in Table 1 with the corresponding Talairach coordinates (Talairach & Tournoux, 1988). These coordinates were computed by projecting the activation maps generated from the dynamic group average onto a representative participant's brain, then transforming that cortex into standardized 3D space using a simple bounding box around the brain aligned along the anterior and posterior commissures.

2.6. fMRI analysis

The hemodynamic response for each trial type was estimated from each voxel's timeseries across the eight experimental scans using a deconvolution general linear model analysis (Dale & Buckner, 1997) with twenty predictors (sampled every second starting at stimulus onset) for each data point, for each condition (Same, Different, Different Exemplar). For all general linear model analyses, timecourses were normalized to percent signal change from the mean baseline intensity. This analysis makes no *a priori* assumptions as to the underlying shape or latency of the underlying responses, and nonetheless yielded classic hemodynamic response functions across visual cortex. The peak latency for the BOLD response when subjects viewed the stationary snapshots of the Framsticks was approximately 6–9 s post-stimulus onset, and approximately 7–10 s post-stimulus onset for the animated pairs. This latency difference likely reflects the longer stimulus duration in the animated sequences (4 s for animated versus 2 s for stationary).

fMR-adaptation was identified in those voxels with significantly lower BOLD response (computed from the four timepoints surrounding the peak of the hemodynamic response) during Same trials as compared to Different trials. Significance was assessed using the false discovery rate (FDR < .05), which controls the expected proportion of false positives among voxels in which the null hypothesis is rejected (Genovese, Lazar, & Nichols, 2002). Adaptation indices were computed as follows: (Same *or* Different Exemplar peak response − Different peak response)/Different peak response. The resulting index is therefore higher for greater adaptation.

In this study, changes in states of articulation in the Different Exemplar condition were accompanied by changes in object viewing perspective. To determine whether this factor influenced our findings, we did a subsequent deconvolution analysis that extracted the hemodynamic response functions from three types of viewpoint changes: small (45° shifts from the first exemplar to the second), moderate (90° perspective shifts) or large (135–180° perspective shifts). We found no statistical differences between these groups in our analyses, and thus have pooled them for the remainder of our study.

As a means for comparing our identified regions of interest with previous findings, we localized the LOC in each of our subjects individually as well as across the group. LOC was identified as the area of ventral cortex more selective for objects than scrambled objects using a hemodynamic response function convolved with a boxcar function, and contrasted using a general linear model (FDR < .01). This comparison yields a number of brain regions, including a large extent in occipito-temporal cortex that extends up the lateral surface, and a focus of activation in the inferior frontal gyrus. The frontal activation likely reflects the working memory demands of the 1-back task subjects were performing during the localizer (Carlson et al., 1998), and the occipito-temporal activation was identified as the LOC. Because there may be some functional differences between subregions within the LOC (e.g. Grill-Spector, Kushnir, Hendler, & Malach, 2000; Kourtzi, Erb, Grodd, & Bulthoff, 2003), we also divided the LOC into two sub-regions: the more lateral and dorsal LO (lateral occipital), and the posterior fusiform (pFs). Adaptation effects were evaluated within these ROIs by calculating the average deconvolved timecourse for all voxels within the ROI, and then contrasting the peak timepoints for each condition. As is evident in Fig. 3, we found no differences between these two subregions in our analysis, and so in some analyses the neural activity from these two regions were pooled.

2.7. Behavioral measurements

To assess subjects' ability to discriminate these objects in the three conditions viewed in the scanner, we measured reaction times for our discrimination judgments in the laboratory using the same stimuli and parameters as viewed in the neuroimaging experiment (including visual angle, presentation time, and trial orders). Twelve subjects were shown pairs of stationary images or animations in the same conditions as in the fMRI experiment (Same, Different, Different Exemplar). As in the fMRI experiment, subjects were asked if the pair of objects depicted the 'same object' or 'two different objects' ('same' being the correct response for the Same and Different Exemplar conditions, and 'different' being the correct response for the Different condition). Subjects were instructed to respond as quickly as possible by pushing one of two buttons on a button box (Empirisoft DirectIN), without sacrificing accuracy. Reaction times were measured starting from the onset of the second stimulus in the pair. Simulating the fMRI experiment, trial types (stationary or animated) were blocked into 30 trials each (the number of trials in a single scan). Subjects completed ten blocks of each type yielding 100 trials in each of the three conditions for animated and stationary stimulus types. Results were analyzed using a repeated measures ANOVA.

3. Results

3.1. Behavior

To determine the overall difficulty of discriminating the objects used in this study, and to ascertain any task differences across our conditions, we measured subjects' reaction times when identifying two objects as either the same or different. A two-way repeated measures ANOVA for stimulus type (stationary or animated) and trial type (Same, Different or Different Exemplar) yielded a significant main effect of both factors (Fig. 2). Subjects required more time to correctly judge the animated trials as compared to the stationary trials ($p < .001$), and reaction times were faster for the Same condition than for the Different and Different Exemplar

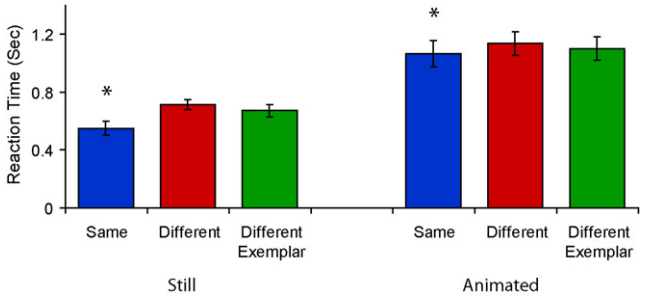

Fig. 2. Group reaction times for all conditions used in the neuroimaging experiments. Asterisks (*) indicate a significantly lower time for the Same condition (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article).

conditions (stationary: $p < .001$; animated: $p < .01$). Reaction times were not significantly different between the Different and Different Exemplar conditions when the Framsticks were viewed as either stationary or animated (paired sample t-test, $p > .1$).

These data show that for these novel objects, subjects were faster to identify the repeated trials as being the same than they were for identifying two different exemplars as being the same, or as the two different objects as being different. We interpret these reaction time results as evidence for recognition enhancement as is typically seen in studies of repetition priming (e.g. Roediger & McDermott, 1993; Tulving & Schacter, 1990). Critically, these behavioral measures do not reveal any processing time differences between the trials with two different stimuli and two different exemplars of the same object. To the extent that behavioral RTs reflect processing load, these findings suggest that the 'same' judgments for the Different Exemplar condition were no more or less taxing than the 'different' judgments for the Different condition. Therefore any differences in neural activity we find for these two conditions is unlikely to reflect differential allocation of attention.

3.2. Adaptation for stationary and animated objects

To determine whether regions in ventral temporal cortex have neural populations with response selectivity for our novel, articulating objects, we tested for brain areas exhibiting fMR-adaptation. This contrast tests for voxels with suppressed neural signals when an object is repeated (Same condition) as compared to viewing two different objects (Different condition). Weaker BOLD response for repeated exposures is believed to reflect repetition suppression, or adaptation, of a neural population specifically tuned to the viewed object. This fMR-adaptation effect has been previously demonstrated in studies of neural response tuning for stationary images of common and novel objects (Ewbank et al., 2005; Grill-Spector & Malach, 2001; James et al., 2002; Kourtzi & Kanwisher, 2000; Kourtzi et al., 2003; van Turennout, Ellmore, & Martin, 2000).

A whole brain GLM analysis of the BOLD response reveals a pattern of adaptation across the ventral and lateral surfaces of the occipito-temporal cortex (Fig. 3A, Table 1). When viewing stationary images of the Framsticks, large regions of adaptation were identified extending from the lateral occipital lobe to the posterior ventral surface of the temporal lobe, and into the medial fusiform. When the objects were viewed as animated (Fig. 3B), the analyses revealed smaller pockets of adaptation to the repeated animations within the same region of cortex. In both cases, the magnitude of the adaptation effect (as revealed by the adaptation index) are similar to those found in previous reports (Hayworth & Biederman, 2006; Kourtzi & Huberle, 2005; Kourtzi & Kanwisher, 2001; Weigelt et al., 2007), with stronger suppression on the stationary trials than in the animated trials.

These results replicate findings of object selectivity in ventral temporal cortex for stationary common objects, and generalize them to novel objects. These results also demonstrate neural selectivity in ventral temporal cortex to novel objects in motion and undergoing complex articulation. Adaptation to articulating objects could result from either neurons that code the many snapshots of form that comprise the articulating object, or neurons that code for the underlying object structure and are invariant to changes associated with articulation. The latter hypothesis was tested in the following analysis.

3.3. Different exemplar adaptation

To test for a generalized object representation that codes for object structure and is invariant to changes in object shape as introduced by different states of articulation, we analyzed for brain regions with suppressed neural signals when subjects viewed two exemplars of the same object. For the stationary condition, the different exemplars were two snapshots of the same object in two different states of articulation. For the animated condition, the different sequences depicted two different examples of the same object articulating.

The statistical maps of adaptation to object structure were generated via a whole brain deconvolution GLM contrasting neural signals during the Different Exemplar trials against those elicited by the two different objects (Different trials). This statistical comparison yielded no significant voxels when subjects viewed the static snapshots of the Framsticks. In other words, we found no evidence of neural populations that generalized across the two states of articulation of the same object when that object was viewed as two static snapshots, separated in time.

For the animated trials, however, bilateral areas of the posterior fusiform revealed significantly lower BOLD response (FDR < .05) during the Different Exemplar trials as compared to the Different object trials (Fig. 3C). In the left hemisphere, the focus of adaptation to the Different Exemplars overlapped partially with those regions showing adaptation for the animated Same object condition. These results are evidence for a neural population selectively tuned to the novel animated objects, even when entirely different examples of the objects' articulation are seen. Because this neural suppression across different exemplars is only evident in the animated trials, these results stress the importance of temporal integration (e.g. Wallis, 2002). It is likely that the conjunction of form and motion (in this case, motion during articulation in the animated condition) plays an important role in building the neural representation of the object within this neural population.

3.4. Comparison to LOC

Because most of the previously published studies examining the tuning properties of object representations have focused on the lateral occipital complex, we did a second analysis targeting this region. Specifically we asked whether the regions of adaptation we identified with our deconvolution analysis fall within the LOC proper.

To address this question, we independently identified the LOC in our subjects using the standard blocked localizer design that compares the neural response to objects and texture patterns (e.g. Kourtzi & Kanwisher, 2000). Across our subjects, we find a large region in occipito-temporal cortex that activates preferentially to stationary images of common objects. A comparison of this independently localized LOC to the Same adaptation maps reveals much overlap between the two (Fig. 4A), with subregions of the LOC adapting to repeated presentations of the both the stationary and moving objects. In contrast, Different Exemplar adaptation for the animated Framsticks revealed a region of cortex outside of the LOC.

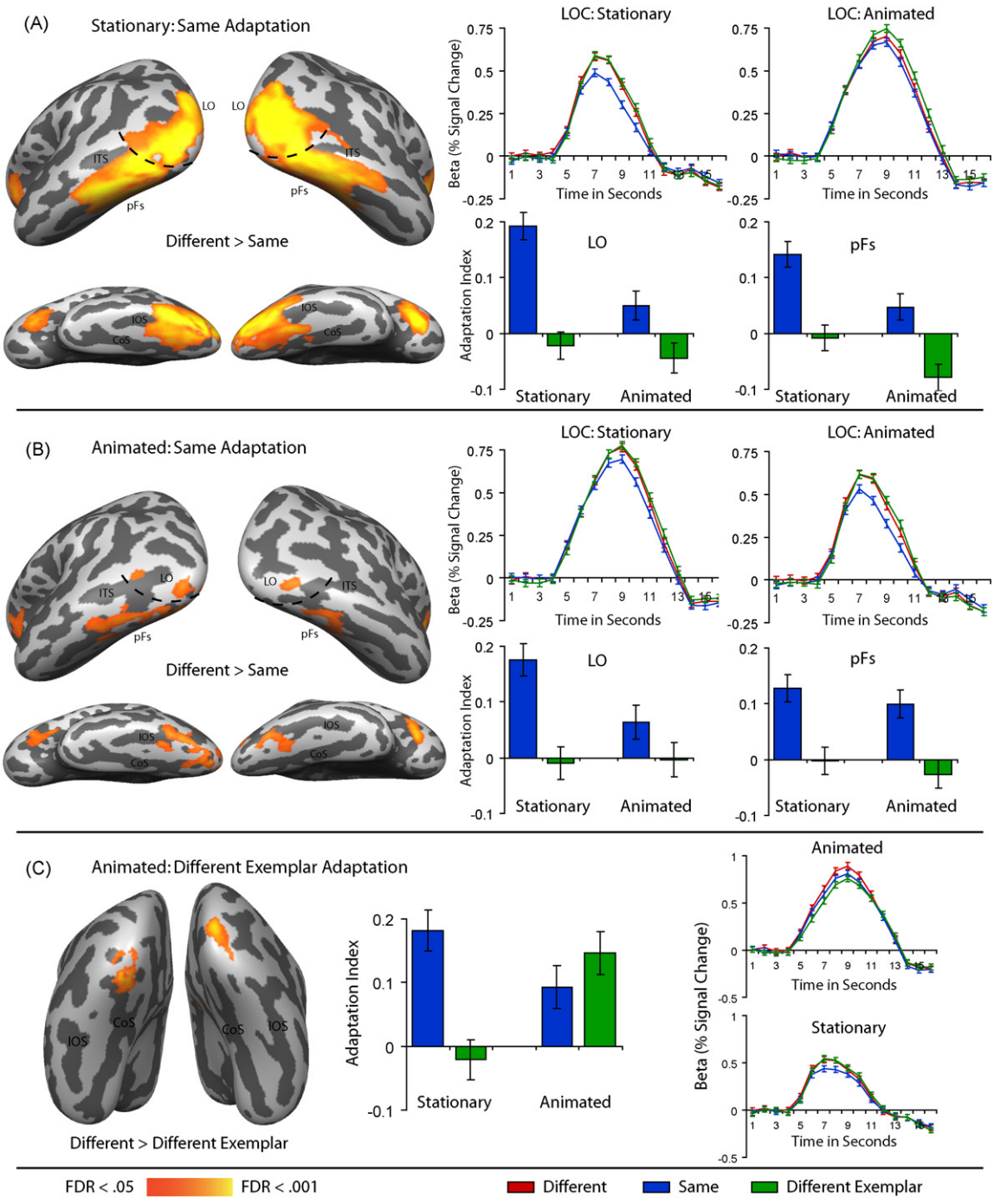

Fig. 3. Whole brain group analyses. (A) Group contrast maps based on cortically aligned data showing areas of adaptation to the stationary Same condition (Different > Same; FDR < .05). The group data is overlaid on an inflated brain from a representative subject. Inferior temporal sulculs (IOS), collateral sulcus (CoS), and inferior occipital sulcus (IOS) are labeled to aid in navigation. On the right are deconvolved hemodynamic responses for the stationary and animated data from significant clusters within the lateral occipital complex (LOC). Error bars indicate +/− one standard error. Also shown are bar graphs of the adaptation indices for the Same and Different Exemplar conditions from significant clusters from the lateral occipital (LO) and posterior fusiform (pFs) subregions of LOC. Error bars indicate one SEM estimated from the peak timepoint. (B) Group contrast maps showing areas of significant adaptation to the animated Same condition, with accompanying hemodynamic response timecourses. (C) Group contrast maps on the occipital lobe of a representative brain displaying bilateral areas of occipital cortex which showed significant adaptation to the Animated Different Exemplar condition. Hemodynamic response timecourses are averaged across hemispheres, along with the adaptation indices.

To compute the strength of adaptation within the LOC proper, we extracted the hemodynamic response functions from the independently localized LOC in each subjects, and computed the ROI-based deconvolution analysis of the BOLD response. Within the LOC proper and all its sub-regions, the peak response for the Same object condition was significantly less than for the Different object condition, regardless of whether the Framsticks were viewed as stationary ($p < .001$) or animated ($p < .01$). However, as suggested by the lack of intersect in the statistical maps, the LOC response failed to show evidence of invariance across the Different Exemplars of the Framsticks, either as stationary snapshots or as animated sequences (Fig. 4A and B). Thus while the neural response of the LOC as a whole reflects the repetition suppression for Same repeated objects, we find no evidence that LOC proper has the neural machinery capable of generating a stable object representation across states of articulation.

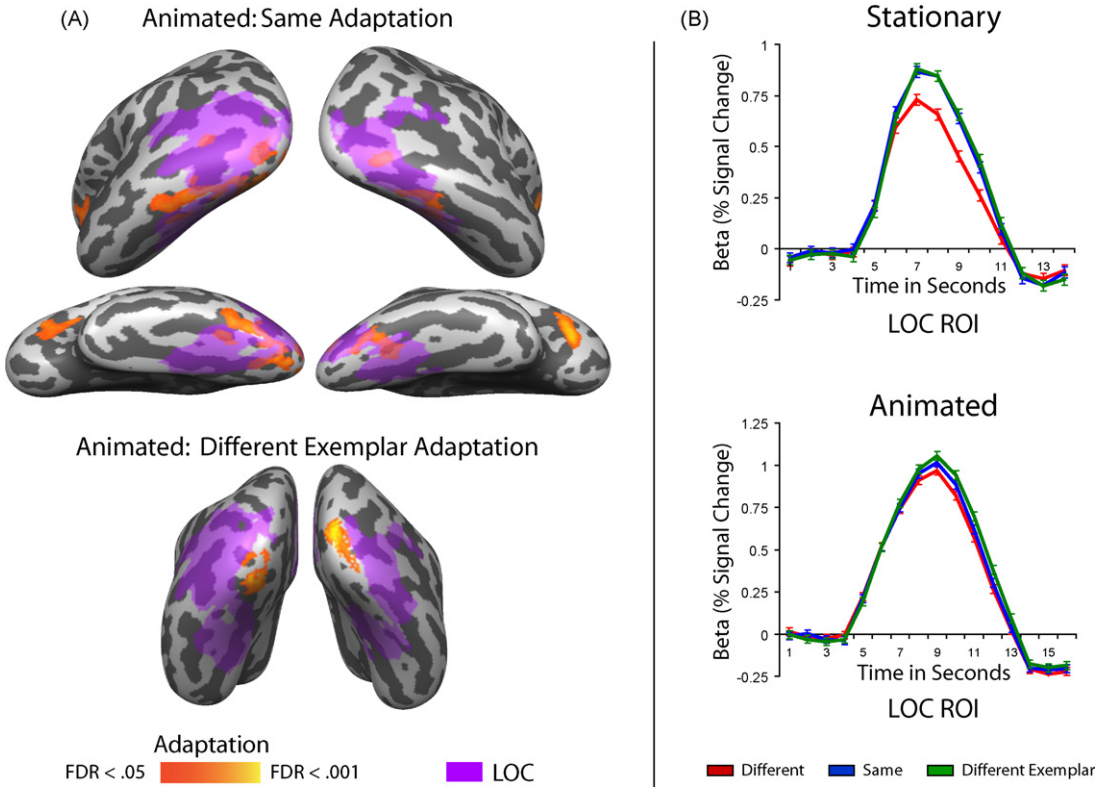

Fig. 4. (A) Statistical maps from the cortically aligned group analysis showing areas of significant (FDR < .05) adaptation to the animated Same and Different Exemplar conditions in orange, as compared to the group localized LOC (FDR < .001) in purple. (B) Group deconvolved timecourses from the stationary and animated data in LOC. Timecourses were extracted from the individually identified ROIs in each subject.

4. Discussion

The goal of this study was to identify neural mechanisms supporting recognition of objects that change shape during motion (e.g. articulate). Shape transformations accompanied by motion occur frequently in natural viewing conditions, evidence that our visual system is able to extract properties of objects that are invariant across these changes. Object constancy must somehow be instantiated in the neural representation, possibly within the visual pathway.

Our findings reveal areas of occipito-temporal cortex that have previously been implicated in the perception of stationary objects, to also have neural signals tuned to the perception of moving objects. Using fMR-adaptation, we find evidence for repetition suppression for repeated exposures to novel objects, both when viewed as static snapshots or as animated sequences. We interpret these findings as evidence for neural populations selectively encoding these novel objects.

We also identified a bilateral area outside of LOC with adapted BOLD response when subjects viewed two different exemplars of the same object, provided that the object was viewed in motion. Because the neural response to these Different Exemplar pairs shows more adaptation than for pairs of two different objects, the implication is that some neurons underlying the voxel responses in these regions are invariant to the changes induced by the different articulated states. In other words, some neurons are coding for the particular object being viewed, and not the specific viewing conditions in which it is being seen. These results are evidence for neural populations in the ventral visual processing stream coding object constancy.

Ventral temporal cortex is a highly complex region, associated with both specialized regions of cortical activity (such as the FFA, PPA, WFA and the LOC), and distributed patterns of neural activity supporting object recognition (e.g. Op de Beeck, Torfs, & Wagemans, 2008). The LOC in particular has been greatly investigated, and is identified as having neural signals invariant across changes in object size, position and viewpoint (e.g. Kourtzi & Kanwisher, 2001; Vuilleumier, Henson, Driver, & Dolan, 2002). Here we find evidence for adaptation within the LOC proper for repeated exemplars of the novel Framsticks, both when viewed as stationary or in motion.

The hallmark of object invariance is a neural representation that generalizes across different exemplars of the same object (for review, see Grill-Spector, 2003). For example, changes in object shape often accompany changes in viewpoint, the neural basis of which has been previously investigated (e.g. Grill-Spector et al., 1999; Kourtzi et al., 2003). In human cortex, the literature reports a range of findings in the extent to which neuronal populations in ventral temporal cortex and the LOC are viewpoint invariant for stationary objects. While some studies find evidence in favor of neural adaptation across viewpoints (Eger, Henson, Driver, & Dolan, 2006; James et al., 2002; Kourtzi et al., 2003), others find evidence against it (Grill-Spector et al., 1999; Vuilleumier et al., 2002). In general, it appears that viewpoint invariance is most typically found when the amount of rotation is relatively small or the image is mirror reversed, and in these cases changing the viewpoint has relatively little effect on the overall shape of the object (i.e. the same parts of the object are largely visible). In studies where larger changes in viewpoint were made, there is little or no evidence for adaptation.

To our knowledge, ours is the first investigation of neural coding invariant to changes in shape that result from articulation. In these experiments, the articulation of the Framsticks causes dramatic changes in shape over time, with the result that two different snapshots of the same object depict quite different 2D projections of object shape. These transformations are more similar to large changes in viewpoint, and in the context of the previous literature, it perhaps should not be surprising that we found no adaptation

to the different stationary exemplars of the Framsticks. This lack of adaptation for the stationary exemplars must be interpreted in the context of our behavioral findings, in which subjects clearly recognized the snapshots as depicting the same object, and were able to make those recognitions with equal speed as recognizing different objects. Therefore although observers can infer object constancy for the stationary stimuli across the transformations of state of articulation, this is not robustly encoded across the neuronal ensemble in ventral temporal cortex.

We did, however, find evidence for repetition suppression across different exemplars of the Framstick objects when they were animated. This represents a unique finding to the literature, but some basis for this finding can be drawn from previous studies of viewpoint adaptation for objects undergoing rigid apparent motion (Weigelt et al., 2007). In that study, only viewpoints from within the path of the apparent motion elicited adaptation in the ventral stream. Our study broadens those findings to viewpoints outside of the object's motion path when the object undergoes complex articulation. The interpretation of our findings suggest that the neural representation supporting recognition of these objects is invariant to different examples of the object's motion pattern, which is typically accompanied by changes in viewpoint and location in the visual field.

More broadly, our result shows that dynamic stimuli can be used to explore the response properties of ventral cortex using an adaptation paradigm, which opens the door to a number of interesting investigations on the temporal properties of object invariance. In particular, this raises the question as to how moving objects are encoded in ventral temporal cortex. Are moving objects treated as a series of stationary 'frames', or is the motion sequence encapsulated in some way and processed as a whole? Previous studies have suggested that temporal association of different views of an object through motion contributes to building an invariant representation of that object (Kourtzi & Shiffrar, 1999; Wallis, 2002). That is, viewpoints and other information gathered through viewing a moving object are bound together in such as way as to later inform recognition of that object across any of the experienced viewing conditions. These previous studies examine this idea with relation to rotating an object and the rotation's effect on viewing the object under different viewpoints.

Our data seem to suggest regions in ventral temporal cortex support both stationary form information, and may use motion to bind object representations across different articulatory states. We found, for example, that the adaptation index for the stationary objects was generally much stronger than for the animated objects, an indicator of many neurons coding for specific instances of stationary object snapshots. However, we only found adaptation when the articulating objects were viewed as animated sequences, not when viewed as stationary snapshots. Thus it appears that motion is playing a critical role in binding together the different shape representations that occur during object articulation. It seems reasonable, therefore, that the contributions to object recognition observed by rotational motion could be extended to articulatory motion as well.

One could argue that a possible reason we failed to find evidence for shape invariance across the stationary snapshots of the Framsticks may be simply that the neural population supporting recognition of these objects is too small to be identifiable with fMR-adaptation. This is a plausible concern given that far fewer neurons may be encoding across viewpoints, as suggested by models of object perception (Riesenhuber & Poggio, 1999). We were, however, able to find a neural population that adapts to different exemplars of the Framsticks when animated. We therefore do not interpret the lack of adaptation for the stationary images as an issue of power.

A final point is that the patterns of adaptation revealed in these findings were not strictly confined to the functionally localized lateral occipital complex or its subregions. In particular, the animated Same adaptation revealed clusters of adaptation within the LOC, but no obvious link to known functional subdivision. Thus our findings reflect the functional heterogeneity of occipito-temporal cortex that has previously been investigated by other experimental manipulations (e.g. Bar et al., 2001; Grill-Spector et al., 1999; Kourtzi et al., 2003), and illustrate that the combination of an ROI analysis with a whole brain analysis provides a more complete picture of the roles of different neural populations.

In summary, we have shown that ventral cortex is actively involved in the perception of moving objects and as with stationary objects, some areas of cortex show responses demonstrating object constancy and some do not. We argue these results strongly suggest a network of neural machinery in ventral cortex that is involved in the processing of dynamic objects and is homologous to the network for processing stationary objects that has been studied in the LOC. Further research should clarify if and how these networks interact, and what resources and processing tasks they might share.

Acknowledgements

We thank Alyssa Brewer, John Serences and Javier Garcia for helpful discussions, and Nicole Jardine for assistance in data collection.

References

Bar, M., Tootell, R. B., Schacter, D. L., Greve, D. N., Fischl, B., Mendola, J. D., et al. (2001). Cortical mechanisms specific to explicit visual object recognition. *Neuron*, *29*(2), 529–535.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*(4), 433–436.

Carlson, S., Martinkauppi, S., Rama, P., Salli, E., Korvenoja, A., & Aronen, H. J. (1998). Distribution of cortical activation during visuospatial n-back tasks as revealed by functional magnetic resonance imaging. *Cerebral Cortex*, *8*(8), 743–752.

Dale, A. M., & Buckner, R. L. (1997). Selective averaging of rapidly presented individual trials using fMRI. *Human Brain Mapping*, *5*, 329–340.

Desimone, R., Albright, T. D., Gross, C. G., & Bruce, C. (1984). Stimulus-selective properties of inferior temporal neurons in the macaque. *Journal of Neuroscience*, *4*(8), 2051–2062.

Eger, E., Henson, R. N., Driver, J., & Dolan, R. J. (2006). Mechanisms of top-down facilitation in perception of visual objects studied by fMRI. *Cerebral Cortex*.

Ewbank, M. P., Schluppeck, D., & Andrews, T. J. (2005). fMR-adaptation reveals a distributed representation of inanimate objects and places in human visual cortex. *Neuroimage*, *28*(1), 268–279.

Formisano, E., Esposito, F., Di Salle, F., & Goebel, R. (2004). Cortex-based independent component analysis of fMRI time series. *Journal of Magnetic Resonance Imaging*, *22*(10), 1493–1504.

Genovese, C. R., Lazar, N. A., & Nichols, T. (2002). Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *Neuroimage*, *15*(4), 870–878.

Goebel, R., Esposito, F., & Formisano, E. (2006). Analysis of functional image analysis contest (FIAC) data with brainvoyager QX: From single-subject to cortically aligned group general linear model analysis and self-organizing group independent component analysis. *Human Brain Mapping*, *27*(5), 392–401.

Grill-Spector, K. (2003). The neural basis of object perception. *Current Opinion in Neurobiology*, *13*(2), 159–166.

Grill-Spector, K., Kourtzi, Z., & Kanwisher, N. (2001). The lateral occipital complex and its role in object recognition. *Vision Research*, *41*(10–11), 1409–1422.

Grill-Spector, K., Kushnir, T., Edelman, S., Avidan, G., Itzchak, Y., & Malach, R. (1999). Differential processing of objects under various viewing conditions in the human lateral occipital complex. *Neuron*, *24*(1), 187–203.

Grill-Spector, K., Kushnir, T., Edelman, S., Itzchak, Y., & Malach, R. (1998). Cue-invariant activation in object-related areas of the human occipital lobe. *Neuron*, *21*(1), 191–202.

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*(8), 837–843.

Grill-Spector, K., & Malach, R. (2001). fMR-adaptation: A tool for studying the functional properties of human cortical neurons. *Acta Psychologica (Amsterdam)*, *107*(1–3), 293–321.

Haxby, J. V., Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., & Pietrini, P. (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science*, *293*(5539), 2425–2430.

Hayworth, K. J., & Biederman, I. (2006). Neural evidence for intermediate representations in object recognition. *Vision Research*, *46*(23), 4024–4031.

Ito, M., Tamura, H., Fujita, I., & Tanaka, K. (1995). Size and position invariance of neuronal responses in monkey inferotemporal cortex. *Journal of Neurophysiology*, *73*(1), 218–226.

James, T. W., Humphrey, G. K., Gati, J. S., Menon, R. S., & Goodale, M. A. (2002). Differential effects of viewpoint on object-driven activation in dorsal and ventral streams. *Neuron*, *35*(4), 793–801.

Komosinski, M., & Ulatowski, S. (1999). Framsticks: Towards a simulation of a nature-like world, creatures and evolution. Paper presented at the Advances in Artificial Life. In *Proceedings of the Fifth European Conference on Artificial Life*.

Kourtzi, Z., Erb, M., Grodd, W., & Bulthoff, H. H. (2003). Representation of the perceived 3-D object shape in the human lateral occipital complex. *Cerebral Cortex*, *13*(9), 911–920.

Kourtzi, Z., & Huberle, E. (2005). Spatiotemporal characteristics of form analysis in the human visual cortex revealed by rapid event-related fMRI adaptation. *Neuroimage*, *28*(2), 440–452.

Kourtzi, Z., & Kanwisher, N. (2000). Cortical regions involved in perceiving object shape. *Journal of Neuroscience*, *20*(9), 3310–3318.

Kourtzi, Z., & Kanwisher, N. (2001). Representation of perceived object shape by the human lateral occipital complex. *Science*, *293*(5534), 1506–1509.

Kourtzi, Z., & Shiffrar, M. (1999). The visual representation of three-dimensional, rotating objects. *Acta Psychologica (Amsterdam)*, *102*(2–3), 265–292.

Logothetis, N. K., & Sheinberg, D. L. (1996). Visual object recognition. *Annual Review of Neuroscience*, *19*, 577–621.

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, USA*, *92*(18), 8135–8139.

Op de Beeck, H. P., Torfs, K., & Wagemans, J. (2008). Perceived shape similarity among unfamiliar objects and the organization of the human object vision pathway. *Journal of Neuroscience*, *28*(40), 10111–10123.

Pyles, J. A., Garcia, J. O., Hoffman, D. D., & Grossman, E. D. (2007). Visual perception and neural correlates of novel 'biological motion'. *Vision Research*, *47*(21), 2786–2797.

Riesenhuber, M., & Poggio, T. (1999). Hierarchical models of object recognition in cortex. *Nature Neuroscience*, *2*(11), 1019–1025.

Riesenhuber, M., & Poggio, T. (2002). Neural mechanisms of object recognition. *Current Opinion in Neurobiology*, *12*(2), 162–168.

Roediger, H. L., & McDermott, K. B. (1993). Implicit memory in normal human subjects. *Handbook of Neuropsychology*, *8*, 63.

Sarkheil, P., Vuong, Q. C., Bulthoff, H. H., & Noppeney, U. (2008). The integration of higher order form and motion by the human brain. *Neuroimage*.

Sary, G., Vogels, R., & Orban, G. (1993). Cue-invariant shape selectivity of macaque inferior temporal neurons. *Science*, *260*(5110), 995–997.

Serences, J. T. (2004). A comparison of methods for characterizing the event-related BOLD timeseries in rapid fMRI. *Neuroimage*, *21*(4), 1690–1700.

Talairach, J., & Tournoux, P. (1988). *Co-planar stereotaxic atlas of the human brain*. New York: Thieme Medical.

Tanaka, K. (2003). Columns for complex visual object features in the inferotemporal cortex: Clustering of cells with similar but slightly different stimulus selectivities. *Cerebral Cortex*, *13*(1), 90–99.

Tulving, E., & Schacter, D. L. (1990). Priming and human memory systems. *Science*, *247*(4940), 301–306.

van Turennout, M., Ellmore, T., & Martin, A. (2000). Long-lasting cortical plasticity in the object naming system. *Nature Neuroscience*, *3*(12), 1329–1334.

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*(5), 491–499.

Wallis, G. (2002). The role of object motion in forging long-term representations of objects. *Visual Cognition*, *9*(1-2), 233–247.

Wallis, G., & Bulthoff, H. (1999). Learning to recognize objects. *Trends in Cognitive Science*, *3*(1), 22–31.

Weigelt, S., Kourtzi, Z., Kohler, A., Singer, W., & Muckli, L. (2007). The cortical representation of objects rotating in depth. *Journal of Neuroscience*, *27*(14), 3864–3874.