

# Multisensory visual–tactile object related network in humans: insights gained using a novel crossmodal adaptation approach

Noa Tal · Amir Amedi

Received: 31 October 2008 / Accepted: 7 July 2009 / Published online: 4 August 2009  
© The Author(s) 2009. This article is published with open access at Springerlink.com

**Abstract** Neuroimaging techniques have provided ample evidence for multisensory integration in humans. However, it is not clear whether this integration occurs at the neuronal level or whether it reflects areal convergence without such integration. To examine this issue as regards visuo-tactile object integration we used the repetition suppression effect, also known as the fMRI-based adaptation paradigm (fMR-A). Under some assumptions, fMR-A can tag specific neuronal populations within an area and investigate their characteristics. This technique has been used extensively in unisensory studies. Here we applied it for the first time to study multisensory integration and identified a network of occipital (LOtv and calcarine sulcus), parietal (aIPS), and prefrontal (precentral sulcus and the insula) areas all showing a clear crossmodal repetition suppression effect. These results provide a crucial first insight into the neuronal basis of visuo-haptic integration of objects in humans and highlight the power of using fMR-A to study multisensory integration using non-invasive neuroimaging techniques.

**Keywords** Multisensory integration · Neuroimaging · Visual object recognition · Tactile object recognition · fMR-adaptation · Repetition suppression

## Introduction

Humans are equipped with multiple sensory channels through which they experience objects in the environment, an ability which is of prime importance to humans and animals. The neural correlate of object processing is the focus of the current work. Of the various properties that can be extracted from an object, shape is critical to human visual (Marr 1982) and tactile (Amedi et al. 2001; Klatzky et al. 1987) object recognition.

Recent research has dispelled the established orthodoxy that the brain is organized around parallel processing of discrete sensory inputs, and has provided strong evidence for “metamodal” brain areas with a multisensory task-based organization (Amedi et al. 2001, 2007; Pascual-Leone and Hamilton 2001). For example, it is now well known that cortical regions previously considered to be specialized for processing various aspects of visual input are also activated during analogous tactile or haptic tasks (reviewed in Sathian and Lacey 2007). Several cortical areas have been implicated in visuo-haptic shape processing in humans; principally the lateral occipital tactile-visual region (LOtv; Amedi et al. 2002), a sub-region within the human lateral occipital complex (LOC; Malach et al. 1995), and the anterior intraparietal sulcus (aIPS).

The LOtv is shape-selective both during haptic 3D perception (Amedi et al. 2001; Stilla and Sathian 2008; Zhang et al. 2004) and tactile 2D perception (Stoesz et al. 2003; Prather et al. 2004). Case studies suggest that the LOC is necessary for both haptic and visual shape perceptions.

---

**Electronic supplementary material** The online version of this article (doi:10.1007/s00221-009-1949-4) contains supplementary material, which is available to authorized users.

---

N. Tal · A. Amedi  
Department of Medical Neurobiology,  
Institute for Medical Research Israel-Canada (IMRIC),  
Hebrew University, Hadassah Medical School,  
91220 Jerusalem, Israel

A. Amedi (✉)  
Cognitive Science Program,  
The Hebrew University of Jerusalem,  
91220 Jerusalem, Israel  
e-mail: amir.amed@ekmd.huji.ac.il  
URL: http://brain.huji.ac.il/

Feinberg et al. (1986) reported a patient with a lesion to the left occipito-temporal cortex, which likely included the LOC, who exhibited both tactile and visual agnosia (inability to recognize objects), although the somatosensory cortex and basic somatosensation were spared. Another patient with bilateral lesions to the LOC was unable to learn new objects by either vision or touch (James et al. 2006). LOtv can be characterized as a processor of geometric shapes, since it is unresponsive during conventional auditory object recognition triggered by object-specific sounds (Amedi et al. 2002) or in tasks that call for learning associations between arbitrary sounds and names (Amedi et al. 2007).

Interestingly, the neighboring MT area (middle temporal/area V5; Tootell et al. 1995) is located in humans in the lateral occipital cortex a bit dorsally to LOtv and is known to be a complex motion area which responds better to moving than to stationary stimuli. It was recently reported that area MT has visuo-haptic convergence, but for moving stimuli rather than for shape (Blake et al. 2004; Hagen et al. 2002). This might extend the “division of labor” principle (Zeki 1978) for metamodal representations across sensory modalities.

All these and many other neuroimaging studies provide evidence for multisensory convergence of visuo-haptic representations in several cortical sites. However, due to the low spatial resolution of fMRI (within each fMRI voxel of 3 mm<sup>2</sup>, fMRI samples average activity across about one million neurons via hemodynamics; Braendgaard et al. 1990; Levy et al. 2004; Rockel et al. 1980), convergence within a particular area or fMRI voxel may not conclusively show that the signals converge onto the same individual neurons. In other words, these studies do not prove the existence of a modality-independent shape representation reflecting multisensory integration at the neuronal level, as opposed to the inter-digitation of unisensory neurons receiving information from visual or haptic input. This point is important for a better understanding of the shared or common properties and transformations of information across sensory modalities, because for multisensory stimulation to maximally affect perceptual and behavioral responses, information from the different sensory systems needs to converge on individual neurons. Using standard fMRI approaches it is impossible to assess whether the source of the signal is the activity of a combination of neuronal populations, each tuned to a different property, or whether it is the outcome of the activity of a homogeneous group of neurons which share a common property, the golden standard of real multisensory integration. When studying multisensory processing, this differentiation is crucial since these two alternatives represent completely different mechanisms.

Convergence of different sensory inputs on an area, but without synapsing on the same neurons, is referred to in

multisensory single unit literature as “areal convergence”. When the inputs converge in the same area and also synapse on the same neurons, this is termed “neuronal convergence”, and results in integration of the different sensory signals (Meredith 2002). Each of these different multisensory effects, determined by the nature of the convergence that produced them, may ultimately contribute to distinct types of multisensory perceptions or behaviors.

To disentangle these alternatives for visual and tactile object recognition, we used an fMRI-based adaptation paradigm (fMR-A) which relies on extremely well-documented repetition suppression effects. The term “repetition suppression” refers to reduced activation, for instance, in object-selective cortex, with repeated object presentation (Sayers and Grill-Spector 2005). This neural repetition effect has been reported at multiple spatial scales, from the level of individual cortical neurons in monkeys (Li et al. 1993; Miller and Desimone 1994; Sobotka and Ringo 1996) to the level of hemodynamic changes measuring the pooled activation of millions of neurons in humans using fMRI (Demb et al. 1995; Buckner et al. 1995; Stern et al. 1996; Grill-Spector et al. 1999; Jiang et al. 2000; Henson et al. 2000; Naccache and Dehaene 2001). So far the fMR-A paradigm has been used successfully to test for functional specialization in vision, for example, for convergence of different cues underlying a visual stimulus, shape processing (Kourtzi et al. 2003; Kourtzi and Kanwisher 2000, 2001), face encoding (Winston et al. 2004; Sigala and Rainer 2007), and the integration of color and motion (Self and Zeki 2005).

This method takes advantage of the repetition suppression effect to reveal generalizations across particular transformations; in our case, the cross modal transformation of object shapes across the visual and tactile modalities. The neuronal mechanisms underlying the repetition effect are not fully understood at this stage, but a straightforward interpretation is neuronal adaptation, namely, a reduction in the average spike rate of the adapted neuronal population (see Grill-Spector et al. 2006 for candidate neural models accounting for the repetition effect; Krekelberg et al. 2006 for a review comparing single-cell recordings with functional imaging).

Here we used a novel multi-sensory fMR-A paradigm, based on the original fMR-A which was first described by Grill-Spector et al. 1999. Our aim was to investigate whether fMRI adaptation would reveal multisensory voxels showing multisensory repetition suppression following the repetition of the same visual and haptic object (see also Amedi et al. 2001 and Driver and Noesselt 2008, which suggested manipulating the BOLD repetition suppression measures to test for crossmodal relations rather than unisensory relations). In each condition the subject viewed a picture of an object, which was followed by a tactile

exploration of the same (VT-same, visual–tactile same) or a different (VT-diff) object. But, unlike previous studies, the visual stimulus was only presented once and was repeated sequentially (rather than simultaneously) in the other, tactile modality. If the neurons are truly crossmodal, presenting a visual object followed by a tactile representation of the same object should result in the suppression of the activation of neurons within the voxel that are elicited by both the visual stimulus and the tactile stimuli, and hence lead to a reduced fMRI signal. On the other hand, if the signal recovers, this would imply that this activity originates from a combination of neuronal populations, each tuned to the object exposure in the visual modality or the tactile modality. Each modality would then activate a new group of neurons, and the result would be a stronger, non-adapted fMRI signal. Similarly, bimodal neurons should show a more general repetition suppression effect for second versus first presentation of a VT event (regardless of whether it is a VT-same or a VT-diff condition). Finally, we also compared these effects to additional control experiments which aimed to further confirm that the activation purportedly representing areas which go through adaptation is specific to objects, and is not a result of arousal and attention effects or incongruency effects.

## Methods

### Subjects

Ten subjects (five women and five men aged 24–30, mean age 27.5), all neurologically normal and native Hebrew speakers, took part in this study after giving their informed consent. All subjects were right-handed and had normal or corrected-to-normal visual acuity. The experiment was approved by the local Helsinki committee in accordance with the 1964 Helsinki Declaration.

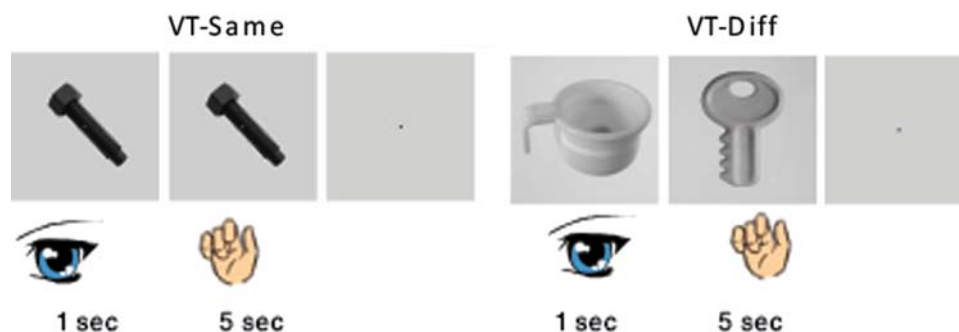
## Stimuli and experimental paradigms

### Experiment 1

In experiment 1 (the main experiment) two different experimental conditions were used in a slow event-related multi-sensory adaptation paradigm: the VT-diff condition (visual–tactile-different) and the VT-same (visual–tactile-same) condition. During both conditions subjects viewed photographs of two-dimensional objects for 1 s, and had 5 more seconds of haptic exploration since vision is much faster than touch which is serial. During the VT-same condition, the subjects touched the same object they were exposed to in the visual exposure. In the VT-diff condition, the subjects touched a different object (Fig. 1). Each multi-sensory event lasted 6 s, followed by 6 s of rest.

The subjects were asked to focus on each stimulus, recognize it without explicit naming or any other motor action during all visual and haptic exploration presentations, and to continue palpating its features until the stop instructions appeared. In order to make sure that the instructions were clear, a set of training examples which included three epochs was administered in the scanner, right before the beginning of the experiment. Subjects were trained on this alternative set of stimuli, so that all stimuli in the experiment were novel (visually and tactually).

There were 18 different objects and 36 events in total. Every object was viewed and touched twice, once in the VT-diff condition, and once in the VT-same condition. VT-diff and VT-same events were pseudo-randomly interleaved so that subjects did not know which event to expect. Subjects did not receive any cues to guide them as to the nature of the upcoming condition (same or different), so they had to recognize both the visual and tactile objects in both types of events. This helped reduce arousal and attention confounds of the different events which is potentially an issue in any adaptation design.



**Fig. 1** The experimental protocol used in the main experiment. Two conditions were interleaved in a slow event-related design: VT-same and VT-diff. In the same condition (VT-same), the subject saw visual objects and touched the same somatosensory objects. In a different

condition (VT-diff), the subject saw and touched different stimuli. Subjects viewed the visual objects for 1 s, and touched the tactile object for 5 s, followed by a 6-s rest

All visual stimuli were presented in gray-scale, with a red fixation point and were placed in the center of the screen. The visual stimuli measured  $350 \times 350$  pixels. In each haptic part of each trial, an experimenter placed the stimulus in the subject's right hand. During haptic exploration, the subjects' eyes were open, and they were instructed to maintain fixation on the center red point. It is important to note that the palpated object was not visible to the subject during exploration. Once the color of the fixation point turned from red to black, the subjects knew that the experimenter was about to remove the tactile object from their hand. The sequence and timing of object presentation were guided by pre-programmed instructions displayed to the subject on a computer screen using the Presentation software package, which also recorded responses.

To test for VT repetition suppression effects at larger time scales and to minimize task and congruency differences while still looking for fMR-A of VT functional adaptation, we further analyzed our data using a secondary contrast. For this contrast, we subdivided the experimental conditions according to their first and second presentations. The fact that we had 18 different objects that were presented twice, once in a VT-diff condition, and once in a VT-same condition, enabled us to do so. Thus, for this analysis we actually had four conditions: VT-diff-1, VT-same-1, VT-diff-2, and VT-same-2. To test for repetition suppression across both congruent and incongruent task, we applied the following tests: VT-diff-1 and VT-same-1 > VT-diff-2 and VT-same-2.

#### Additional experiments

Experiments 2 and 3 were control experiments, in which we further tested for the possible contribution of other factors to the activation in the specified ROIs found in the main experiment.

#### *Experiment 2*

The goal of this experiment, in which we were able to invite back five subjects from the original experiment, was to (1). Test for the specificity of the effect to objects (in comparison to textures) (2). Use a one-back task to further control for arousal and attention effects and to minimize the difference between the different and same events. The experiment included tactile objects and tactile-textures. We designed four conditions that were used in a slow event-related paradigm: tactile-object-diff (TO-diff), tactile-object-same (TO-same), tactile-textures-diff (TT-diff), tactile-textures-same (TT-same). Each event lasted 6 s followed by 9 s of rest. Each condition was repeated eight times. In all the conditions, subjects haptically explored a tactile object or texture for 3 s, and had three more seconds of haptic exploration of

the same or a different object or texture, respectively. In all conditions, the subjects performed a one-back comparison between the currently presented stimulus and the previous one, indicating if the two were the same or different by pressing the left or right button on a response box. The objects and textures were only presented once and were not repeated, except for the "same" conditions. The different conditions were pseudo-randomly interleaved so that subjects did not know which event to expect. In each trial, an experimenter placed the stimulus in the subject's right hand. During haptic exploration, the subjects' eyes were open, and they were instructed to maintain fixation on the center red point. Once the color of the fixation point turned from red to black, the subjects knew that the experimenter was about to remove the tactile object from their hand.

#### *Experiment 3*

The experiment included only tactile objects. Its main aim was to test for tactile adaptation in the entire brain without dealing with congruency effects (see also below). The existence of tactile responses in visual or prefrontal areas for instance can also provide additional indirect support for the crossmodal adaptation effects reported in the main experiment without the involvement of a congruency effect. During the scan, the subjects were exposed to various objects that were handed to them for tactile exploration. Each event lasted 3 s followed by 9 s of rest. Every object was repeated once (TO-1), twice (TO-2), or four times (TO-4) sequentially but always with a 9-s rest interval after each object. The different objects (i.e. TO-1, TO-2 and TO-4 sequences), were presented in a pseudo random order which was not known to the subjects. In all conditions the subjects performed a one-back comparison between the currently presented stimulus and the previous one, indicating if the two were the same or different by pressing the left or right button on a response box. In each trial, an experimenter placed the stimulus in the subject's right hand. During haptic exploration, the subjects' eyes were open, and they were instructed to maintain fixation on the center red point. Once the color of the fixation point changed, the subjects knew that the experimenter was about to remove the tactile object from their hand. As noted earlier, the aim of this experiment was to avoid any incongruency effects by comparing the first with the second and fourth repetition of the same objects and looking for tactile adaptation. If there are indeed neurons responding to tactile input (in LOTv for instance), we would expect to see tactile adaptation for objects, but this time without the putative confound of the incongruency effect. This was designed to enable us to determine which area subsumes tactile object related neurons, to further supplement the results presented in the main crossmodal adaptation experiment.

## MR scanning

MR scans were performed on a 3 T Magnetom Trio scanner (Siemens, Germany), using a standard quadrature head coil. Functional images with blood oxygenation level-dependent (BOLD) contrast were acquired using a T2\*-weighted single-shot gradient-recalled echo planar imaging (EPI) sequence. Axial slices of 3-mm thickness were acquired to provide full-brain coverage with the following parameters: repetition time (TR) 3 s, echo time (TE) 30 ms, flip angle (FA) 85°, in-plane resolution 3 mm × 3 mm, in-plane matrix 80 × 80, FOV = 240 mm. 42 slices with 3 mm slice thickness, and a 1-mm gap were oriented approximately in the axial position, covering the whole brain. The subject lay supine in the scanner with the right arm outstretched beside the body. When advanced to the scan position (head centered in the magnet bore), the subject's right hand was located at the flared magnet aperture, fully accessible and free to haptically explore objects. Foam padding under the body and under the right hand was used to minimize movement and transfer of vibration from the gradient coils and ensure the subject's comfort. A mirror positioned above the subject's eyes provided unobstructed visualization of images projected on a screen at the rear magnet aperture. Head restraint straps and foam blocks were utilized to minimize head movement. Sound-attenuating headphones were used to muffle scanner noise.

## Image processing and analysis

Image processing and analysis was performed using BrainVoyager QX v10 (Brain Innovation, Maastricht, the Netherlands). Each subject's BOLD images were realigned to the first image of the series using a rigid-body transformation procedure. Functional 2D images were pre-processed utilizing trilinear interpolation for 3D motion correction, sinc interpolation for slice scan time correction, and linear trend removal and high-pass temporal filtering of 3 cycles per experiment scan time.

Anatomical 3D images were processed, co-registered with the functional data, and transformed into Talairach space (Talairach et al. 1988). The imaging data were spatially smoothed (spatial Gaussian smoothing, FWHM = 8 mm). Statistical group analysis in the general linear model (GLM) was calculated using hierarchical random effects model (RFX) analysis allowing for generalization of the results to the population level. Each contrast was performed with correction for multiple comparisons using a cluster size threshold adjustment based on the Forman et al. 1995 Monte Carlo stimulation approach (1,000 iterations), extended to 3D data set cortical voxels using the threshold size plug-in in BrainVoyager QX. This method takes into account the probability of a false detection for any given

cluster at a specific activation threshold and spatial smoothing parameters (spatial Gaussian smoothing was the same for all maps, FWHM = 8 mm). The minimum significance level,  $t(9) > 2.40$ , for individual voxels was combined with a cluster size  $>17$  functional voxels resulting in a map which was corrected for multiple comparisons with a minimum statistical threshold of  $p < 0.05$ . Supplementary Fig. 1 also presents a statistical parametric map using a higher initial threshold: at a minimum significance level,  $t(9) > 3.70$ , for individual voxel was combined with a cluster size  $>5$  voxels resulting in a map which was corrected for multiple comparisons with a minimum statistical threshold of  $p < 0.05$ .

## Group level GLM analysis

Statistical group analysis in the general linear model (GLM, Friston et al. 1995, 1999) was calculated using hierarchical random effects model analysis allowing for generalization of the results to the population level. A contrast of the VT-diff and VT-same conditions was used to identify regions that were going through adaptation. The contrast VT-diff versus VT-same was applied in conjunction with the VT-diff versus fixation baseline condition in order to avoid negative activations (VT-diff  $>$  VT-same, conjunction with VT-diff  $>$  baseline). Statistical parameter maps are presented on MNI's full Talairach-normalized inflated brain.

Since adaptation effects last from milliseconds (Sobotka and Ringo 1996) to minutes (Henson et al. 2000) and even days (van Turenout et al. 2000) in some cases (as also reviewed in the Grill-Spector work in 2006), we looked for such tactile adaptation effects of repeated stimuli within each run at an order of magnitude of tens of seconds. Since each object was presented twice in a pseudo-randomized order, once in a VT-same condition, and once in a VT-diff condition, this allowed us to carry out a second contrast between the first presentations of the object versus the second presentation of the object.

## MT localizer

The MT localizer was used to further enhance our ability to compare our results to the known architecture of the visual cortex. The MT ROI was defined according to a specific localizer mapping in nine subjects who also participated in the main experiment. Low-contrast moving versus stationary ring stimuli was used to localize MT according to standard practices (Tootell et al. 1995). Statistical parametric maps of activation were calculated using a random effect GLM analysis. The minimum significance level, corrected for multiple comparisons for any given cluster, was  $p < 0.05$ .

## Visual objects localizer

To localize visual object-selective areas over the entire brain, we conducted an independent statistical test that searched for voxels activated preferentially during visual object epochs compared to visual scrambled epochs, in conjunction with the objects versus fixation baseline in order to avoid negative activations (linear regression, object > scrambled and blanks).

Statistical parametric maps of activation ( $n = 5$  participants who also participated in the main VT adaptation experiment) were calculated using a fixed-effect GLM analysis. The minimum significance level, corrected for any given cluster, was  $p < 0.001$  (a stricter threshold was applied here due to the use of a fixed effect GLM). This statistical parametric map served as a mask which was used for our relative contribution analysis (see next paragraph).

## Relative contribution maps

Two-color maps of VT-diff versus VT-same were performed (Fig. 4). The marked voxels are those for which the relevant predictors explained a significant portion of the variance in the voxel's activation time course, using the normalized ratio between the GLM beta weights of the two conditions. It is important to note that this analysis does not assign any a-priori preference to either of the two conditions; the statistical test is equivalent to performing a VT-same + VT-diff > fixation baseline contrast, and only then the relative contribution is assigned.

This two-color analysis appears over regions that were preferably activated for visual objects versus scrambled figures (using the visual object localizer described above). This differential activation for visual objects versus scram-

bled ones, in conjunction with visual objects, defined our regions of interest for the two-color relative contribution maps (using a FFX analysis,  $p < 0.001$  corrected for multiple comparisons).

Within this visual localizer mask, the group results for two-color analysis were obtained in a fixed effects (FFX) analysis. These group results could not be generalized to the population level since the data for all the subjects were concatenated and analyzed as though they came from a single subject.

## Region of interest (ROI) approach

To quantitatively assess the adaptation contrast activation, we analyzed the time course of the activation in the areas which showed adaptation following the main contrast (VT-diff > VT-same; these were the peak LOTv, Calcarine, IPS and anterior Insula bilaterally, and the left PreCS clusters; Talairach coordinates for each peak ROI are presented in Table 1). In these same areas, we also analyzed the percent signal change of the contrast performed in this same experiment which compares the activation of the first object repetition versus the second repetition, and also the two control experiments.

To quantitatively assess possible fMR-A differences within visual areas of the gradient that we found using the relative contribution map of VT-same and VT-diff (Fig. 4), we analyzed the time course for three points along this dorsal-ventral axis. These three ROIs are (1). The MTG (middle temporal gyrus, using an anatomical landmark localizer) most dorsally; (2). Human MT (using the peak MT localizer); (3). LOC (using the peak visual objects localizer) located most ventrally. Note that since there is no inherent a-priori bias in the statistical maps in the two

**Table 1** Talairach coordinates, and statistical significance for peak ROIs selected from the activation of the contrast VT-diff > VT-same, and for the contrast VT-1st > VT-2nd (LOTv, IPS, insula and calcarine bilaterally, and the left PreCS)

Cortical region	X	Y	Z	Paired <i>t</i> test SO-diff > SO-same	One way ANOVA 1st versus 2nd repetition	Two way ANOVA same versus different effect (VT-same-1, VT-same-2 vs. VT-diff-1, VT-diff-2)	Two way ANOVA 1st/2nd repetition effect (VT-same-1, VT-diff-1 vs. VT-same-2, VT-diff-2)
LH LOTv	48	-54	-8	$t(9) = 7.59; p < \mathbf{0.0001}$	$F(3,36) = 7.07; p = \mathbf{0.003}$	$F(1,38) = 5.37; p = \mathbf{0.02}$	$F(1,38) = 3.72; p = 0.06$
IPS	30	-59	35	$t(9) = 6.01; p = \mathbf{0.0002}$	$F(3,36) = 3.11; p = \mathbf{0.04}$	$F(1,38) = 3.61; p = 0.06$	$F(1,38) = 2.75; p = 0.10$
PreCS	43	3	31	$t(9) = 5.90; p = \mathbf{0.0002}$	$F(3,36) = 10.66; p < \mathbf{0.0001}$	$F(1,38) = 7.61; p = \mathbf{0.01}$	$F(1,38) = 11.41; p = \mathbf{0.002}$
Ins	33	20	0	$t(9) = 4.11; p = \mathbf{0.0026}$	$F(3,36) = 11.08; p < \mathbf{0.0001}$	$F(1,38) = 4.65; p = \mathbf{0.03}$	$F(1,38) = 8.22; p = \mathbf{0.007}$
Calcarine	20	-70	9	$t(9) = 4.18; p = \mathbf{0.0024}$	$F(3,36) = 5.6; p = \mathbf{0.004}$	$F(1,38) = 5.76; p = \mathbf{0.02}$	$F(1,38) = 2.21; p = 0.14$
RH LOTv	-44	-58	-5	$t(9) = 3.45; p = \mathbf{0.0072}$	$F(3,36) = 3.89; p = \mathbf{0.02}$	$F(1,38) = 1.25; p = 0.27$	$F(1,38) = 2.84; p = 0.10$
IPS	-23	-51	43	$t(9) = 2.88; p = \mathbf{0.0182}$	$F(3,36) = 2.58; p = 0.07$	N/A (one way ANOVA n.s.)	N/A (one way ANOVA n.s.)
Ins	-40	17	6	$t(9) = 6.07; p = \mathbf{0.0002}$	$F(3,36) = 1.56; p = 0.22$	n.s. (one way ANOVA n.s.)	n.s. (one way ANOVA)
Calcarine	-15	-69	9	$t(9) = 3.81; p = \mathbf{0.0041}$	$F(3,36) = 4.66; p = \mathbf{0.0095}$	$F(1,38) = 3.5; p = 0.06$	$F(1,38) = 1.35; p = 0.25$

LH left hemisphere, RH right hemisphere, n.s. non significant

The significance of bold values is  $p > 0.05$

**Table 2** Talairach coordinates, number of voxels, and statistical significance for ROIs selected from the activation of the ROIs selected to present the gradient in the occipital area (MTG, area MT, and LO)

	Cortical region	X	Y	Z	Number of voxels	2 tail paired <i>t</i> test (VT-diff vs. VT-same)
Left hemisphere	LOC	48	−64	−6	100	<i>t</i> (9) = 3.96; <i>p</i> = <b>0.0033</b>
	MT	42	−65	7	94	<i>t</i> (9) = 2.09; <i>p</i> = 0.06
	MTG	51	−59	10	64	<i>t</i> (9) = −0.59; <i>p</i> = 0.57
Right hemisphere	LOC	−41	−67	−3	117	<i>t</i> (9) = 2.63; <i>p</i> = <b>0.0274</b>
	MT	−43	−66	6	104	<i>t</i> (9) = 1.48; <i>p</i> = 0.17
	MTG	−45	−58	12	54	<i>t</i> (9) = −1.48; <i>p</i> = 0.17

LOC lateral occipital complex, MT middle temporal, MTG middle temporal gyrus

The significance of bold values is  $p > 0.05$

relative contribution maps (see above), any differential response found in the time courses reflects a completely unbiased preference. Talairach coordinates for each peak ROI are presented in Table 2.

To quantitatively test for differences across stimulus condition sample from these ROIs, we entered the beta values of the different conditions in each experiment from the GLM performed in the BrainVoyager QX, and used SAS Version 8.2 statistical analysis software (SAS Institute Inc, Cary, NC) to conduct paired *t* tests (VT-diff vs. VT-same) on each peak ROI (Tables 1, 2). To test for a putative repetition suppression effect without the confound of (in)congruency, we also conducted independent one way ANOVAs that tested for differences across the two repetitions of VT-diff (VT-diff-1, VT-diff-2) and VT-same (VT-same-1, VT-same-2). In each ROI showing a significant one-way ANOVA effect we also tested using a two-way ANOVA for same versus different effect (VT-same-1, VT-same-2 vs. VT-diff-1, VT-diff-2), and for first versus second repetition effect (VT-same-1, VT-diff-1 vs. VT-same-2, VT-diff-2). The results for all of these statistical tests are summarized in Table 1.

## Results

In the main experiment, we investigated the patterns of cortical activation in ten subjects under two different experimental conditions: a visual–tactile same condition (VT-same) in which subjects viewed and touched the same object, and visual–tactile different condition (VT-diff) in which subjects viewed and touched a different object.

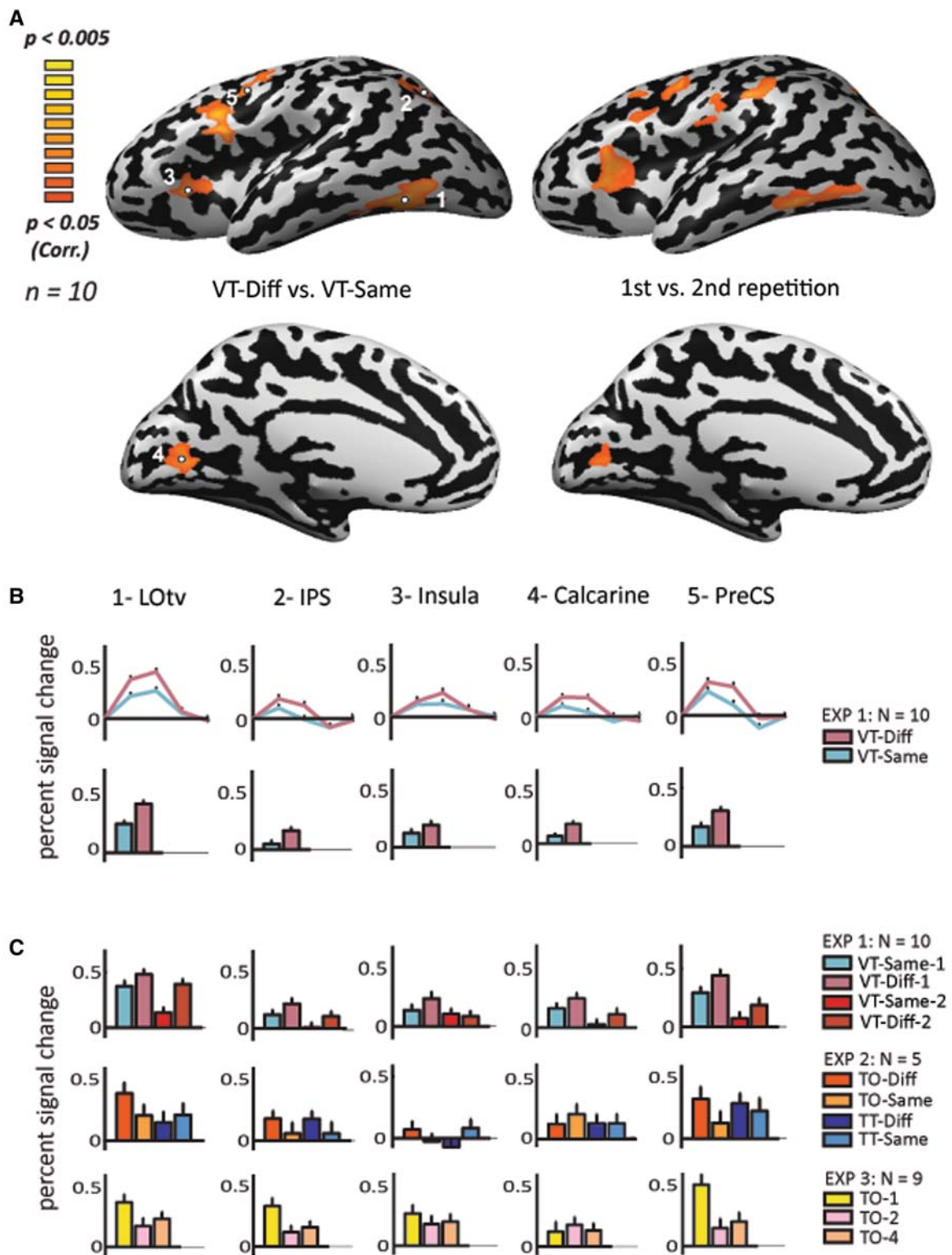
Data were analyzed at several levels. We first present the group analysis of the cortical activation patterns when contrasting VT-diff versus VT-same (Figs. 2a, 3a for left and right hemispheres, respectively). Since each VT-diff and VT-same event was repeated twice during the experiment, we also present a contrast of the first exposure versus the second exposure of any VT event (looking for tactile repeti-

tion suppression effects across the entire brain avoiding putative incongruency effects; see discussion). Finally, we present the activation from a two-color contribution map over all visual object-related areas that showed significant activation to objects versus scrambled Figures (Fig. 4). In all of these contrasts and maps we verified the magnitude of response in several regions of interest which showed an effect in the main contrast (Figs. 2, 3) and also looked at the magnitude of response in several areas defined by external localizers obtained for a subset of the scanned group.

First, we present the group results of the cortical activation patterns when contrasting VT-diff versus VT-same in the left hemisphere (Fig. 2). By contrasting these two conditions, we show areas that go through visuo-haptic adaptation (crossmodal repetition suppression effect). Significant activation was found mainly in five regions: the LOtv, aIPS, the Insula, the Calcarine sulcus (all bilaterally), and the left Precentral sulcus in prefrontal cortex (Figs. 2, 3). Activation was bilateral in most areas, but with a clear trend for left hemisphere preference (Figs. 2, 3; see also Suppl. Fig. 1). From this network of areas, activation was most significant in left LOtv. The left aIPS showed the second strongest crossmodal adaptation (Figs. 2, 3; see also Suppl. Fig. 1).

In addition, the percent signal change for these object-selective ROIs was sampled and averaged across subjects. Time course analysis of activation in the five regions of interest (Fig. 2b) presents the average percent signal change across subjects for contrasting VT-diff versus VT-same. This analysis was statistically significant (see Table 1 for Talairach coordinates and statistical significance values).

Figure 2a also shows the group analysis of the cortical activation patterns of the left hemisphere when contrasting the first exposure to the second exposure of the objects. Since adaptation effects can last for minutes (Henson et al. 2000), we looked for such VT events of repeated stimuli within each run at an order of magnitude of tens of seconds. This analysis also examined whether any confounding factor, such as congruency, could explain our data. We present



**Fig. 2** Statistical parametric maps and magnitude of response in the left hemisphere. **a** Activation maps of the left hemisphere for VT-diff versus VT-same (*left*) and for first exposure versus the second exposure of all VT objects conditions (*right*). Statistical parametric maps of activation ( $n = 10$ ) using a random-effect GLM analysis. The data are presented on MNI full Talairach inflated brain. *Color* scale denotes significance (corrected for multiple comparisons). **b** Time course analysis

of activation and average percent signal change in the five regions of interest defined by the five significant VT-diff versus VT-same clusters presented in A. **c** Average percent signal change in the same ROIs for first exposure versus second exposure of all VT objects conditions ( $n = 10$ ), and in the two control experiments below (experiment 2,  $n = 5$ ; experiment 3,  $n = 9$ )



the histograms for the average percent signal change (Fig. 2c, upper row). All areas were found to be statistically significant with one-way ANOVAs (Table 1), and for that reason we ran two-way ANOVAs that returned a significant same/diff effect for the LOTv, PreCS, Calcarine and Insula, and a marginally significant same/diff effect for the IPS. A repetition effect (first vs. second repetition) was found in the PreCS and the Insula, and a marginal effect in the LOTv and IPS (Table 1).

A similar analysis was conducted for the right hemisphere, as shown in Fig. 3. The LOTv, aIPS, the Calcarine, and the Insula were found to be significantly active following the different versus same contrast (Fig. 3c for time courses and percent signal change, and Table 1, for one-way ANOVA statistics), and the LOTv and the Calcarine sulcus were found to be significantly active following the first versus second repetition in the right LOTv (Table 1). Here also, we present the percent signal change and histograms of the ROIs selected from the VT-diff versus VT-same contrast for the activation result from this contrast (Fig. 3b), and for the activation result from the first versus the second repetition contrast (Fig. 3c, upper row). A two-way ANOVA of the LOTv revealed no significant effect for same/diff or for the repetition effect.

When looking at the activation pattern in our predetermined ROIs, in particular the average activation in Experiment 2 which tested for tactile adaptation across objects and textures (Fig. 2c, 3c, two lowest rows), there was a clear trend in LOTv showing greater activation in the object trials than in the texture trials, and an adaptation in both objects and textures (diff > same). However, no additional statistical tests were conducted due to the small number of subjects. Note that these time course activations were sampled from the ROIs representing the peak activation in the main contrast (diff vs. same; see Table 1) and were used here to further illustrate that even when using a totally different design and task, we were able to find repetition suppression effects specific to objects (see also “Discussion”).

Experiment 2 included both tactile objects and textures and a controlled task within the scanner for the different/same events. Behaviorally, we found that subjects performed all four conditions similarly: a 95% binomial proportion confidence interval was calculated for the accuracy rate for subjects and trials (84–94%;  $H_0$  = no significant difference between accuracy rate across the different trials), and it was confirmed that the subjects' performance was similar in all conditions.

In experiment 3 (both right and left hemispheres), activation in the first repetition was higher than in the second and fourth presentations in most of the reported ROIs bilaterally. The activation of the first as compared to the second presentation was highly significant in bilateral LOTv, and

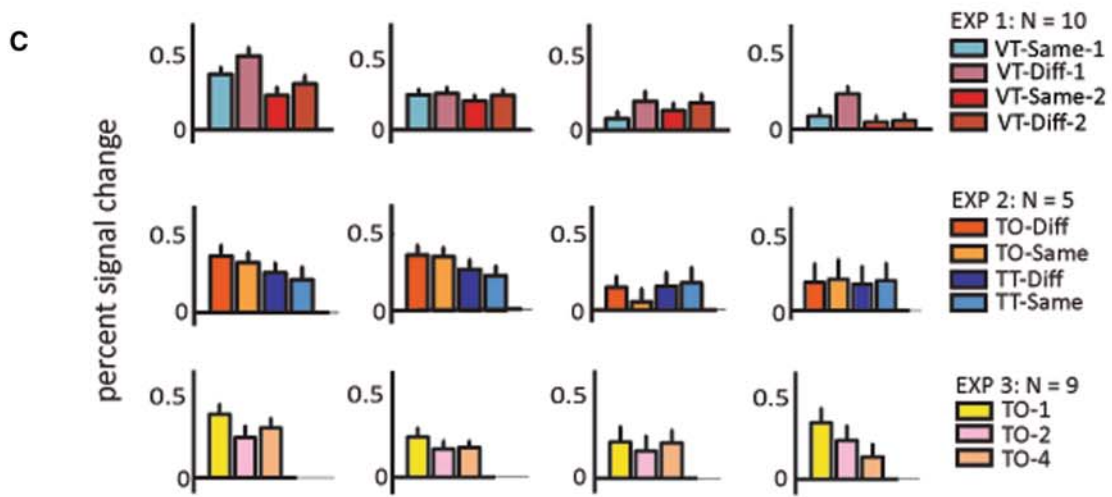
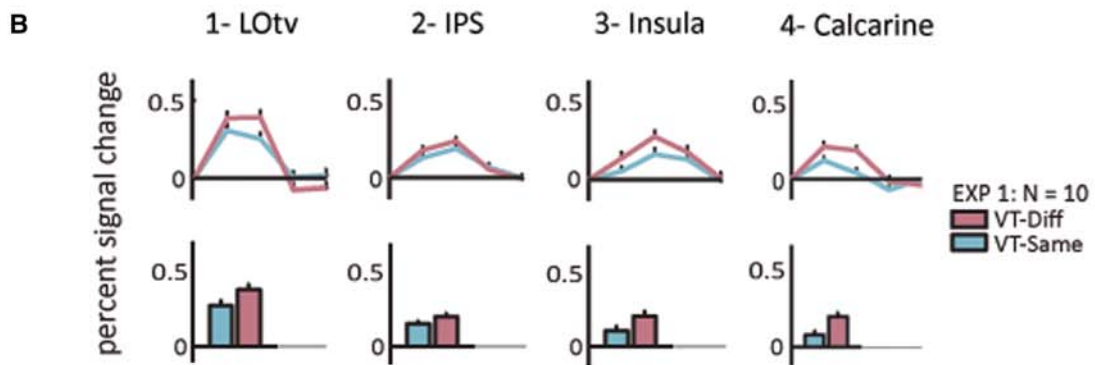
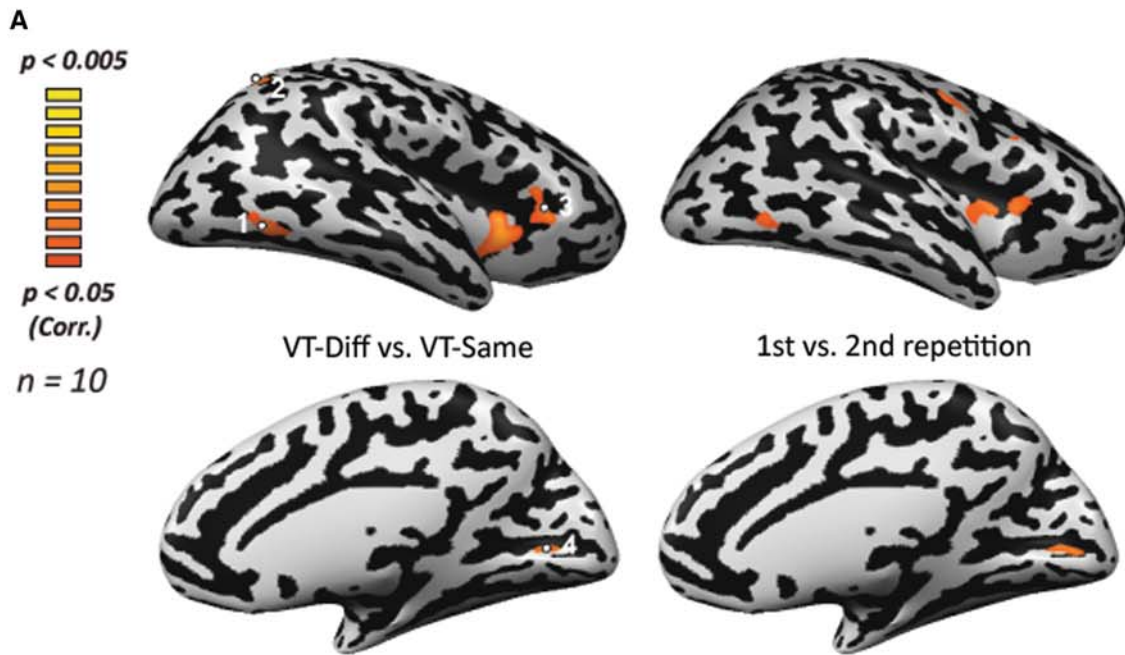
left aIPS and PreCS ( $t > 4.23$ ;  $p < 0.003$ ). It was also significant in the left Insula and the right aIPS ( $t > 2.53$ ;  $p < 0.035$ ). The activation of the first versus the fourth presentation was highly significant in the left LOTv, right Calcarine sulcus, left aIPS and left PreCS ( $t > 4.69$ ;  $p < 0.002$ ), and also significant in the left Insula and the right LOTv and right aIPS ( $t > 2.45$ ;  $p < 0.04$ ).

Bear in mind that only part of the visual cortex showed significant crossmodal adaptation (i.e. part of LO but not the more dorsal and basal ventral areas; see Fig. 3a). Thus, next we focused on all visual object-related areas (including those not showing crossmodal adaptation) to test for possible large-scale organization of visual unisensory versus visuo-haptic multisensory areas. We used a combination of a fixed-effect visual object localizer, combined with a relative contribution analysis. This analysis was designed to find the relative contribution of two predictors in the model in terms of regression (using the normalized ratio between the GLM beta weights of the two conditions; see methods). However, this analysis does not assign any a-priori preference to either of the two conditions.

In order to define areas that showed selectivity to visual objects, we ran an independent statistical test that searched for voxels activated preferentially during visual object epochs compared to visual scrambled epochs on five of the ten subjects who participated in main Experiment 1. The object-selective areas included the ventral stream and partially the dorsal stream of the visual system. This is consistent with previous studies that suggested that both the ventral areas and the dorsal areas are engaged in different aspects of object processing (Goodale 2000; Shmuelof and Zohary 2005; see also “Discussion”). We present the statistical map resulting from the two sets of relative contribution analyses (Fig. 4a), within areas that showed significant preference for objects versus scrambled photos.

A spatial gradient, from non-adaptive areas to fully adaptive areas, was identified in the lateral occipital cortex. Other areas showing activation in the superior parietal cortex and in the frontal cortex did not show the same pattern of gradual suppression. This was also found in the examination of the average activation in the time course analysis. We found a ventral-to-dorsal gradient for areas that showed adaptation. Areas located more dorsally in the gradient, around and including area MT, showed less adaptation, whereas areas located more ventrally showed more adaptation, reaching a maximum crossmodal object adaptation in area LOTv.

To quantitatively assess the gradient and topographical specialization in the occipito-temporal cortex, we analyzed the time course of activation in the three regions of interest located within the occipital cortex. Figure 4b presents the average time course and histograms (across subjects) of activation in the visual object-selective voxels in the



◀ **Fig. 3** Statistical parametric maps and magnitude of response in the right hemisphere. **a** Activation maps of the right hemisphere for VT-diff versus VT-same (*left*) and for first exposure versus the second exposure of all VT objects conditions (*right*). Statistical parametric maps of activation ( $n = 10$ ) using a random-effect GLM analysis. The data are presented on MNI full Talairach inflated brain. *Color* scale denotes significance (corrected for multiple comparisons). **b** Time course analysis of activation and average percent signal change in the four regions of interest defined by the four significant VT-diff versus VT-same clusters presented in A. **c** Average percent signal change in the same ROIs for first exposure versus second exposure of all VT objects conditions ( $n = 10$ ), and in the two control experiments below (experiment 2,  $n = 5$ ; experiment 3,  $n = 9$ )

occipito-temporal region. The most dorsal ROI, the MTG/pSTS (Fig. 4b), showed a slight preference for the VT-same condition (see Table 2). We chose the second ROI to be the MT area that was located along the gradient found in the occipital cortex. MT did not show any significant preference in the left hemisphere (Table 2). The third ROI was the LOC, which is located in the most ventral area of the gradient, and showed a statistically significant preference for the VT-diff condition (Table 2). The MT localizer and object localizer pattern of activation are shown in Fig. 4a.

## Discussion

### Summary of results

In this study, we applied the fMR-A paradigm to identify the neuroanatomical basis for coding multisensory visuo-haptic object recognition. We found that five main regions showed clear crossmodal adaptation: the LOTv, the aIPS, the insula, the calcarine (all bilaterally), and the left precentral sulcus (PreCS). Different subsets of these five areas have been reported to have crossmodal features, (as summarized in Amedi et al. 2005; Lacey et al. 2007), but using our approach, which might be more sensitive, we were able to find significant effects in all five areas. More importantly, previous studies (including our own) could not clearly rule out either of the possibilities detailed in the Introduction; namely that: (1) This reflects truly bimodal neurons in these areas, or (2) This reflects a combination of visual and tactile neurons present in each voxel. Our results clearly support the first alternative in all five areas.

Furthermore, using this technique we were also able to demonstrate that in the visual cortex there is a gradient from no adaptation in the dorsal stream (around the MTG and area MT), and a gradual increase in multisensory preference for objects, which reaches its maximum multisensory adaptation in the LOTv area. Thus, our study suggests that there is an increase in the proportion of bimodal visuo-haptic object-related neurons when moving from the

middle temporal gyrus (MTG) towards the lateral occipital junction, peaking in the LOTv.

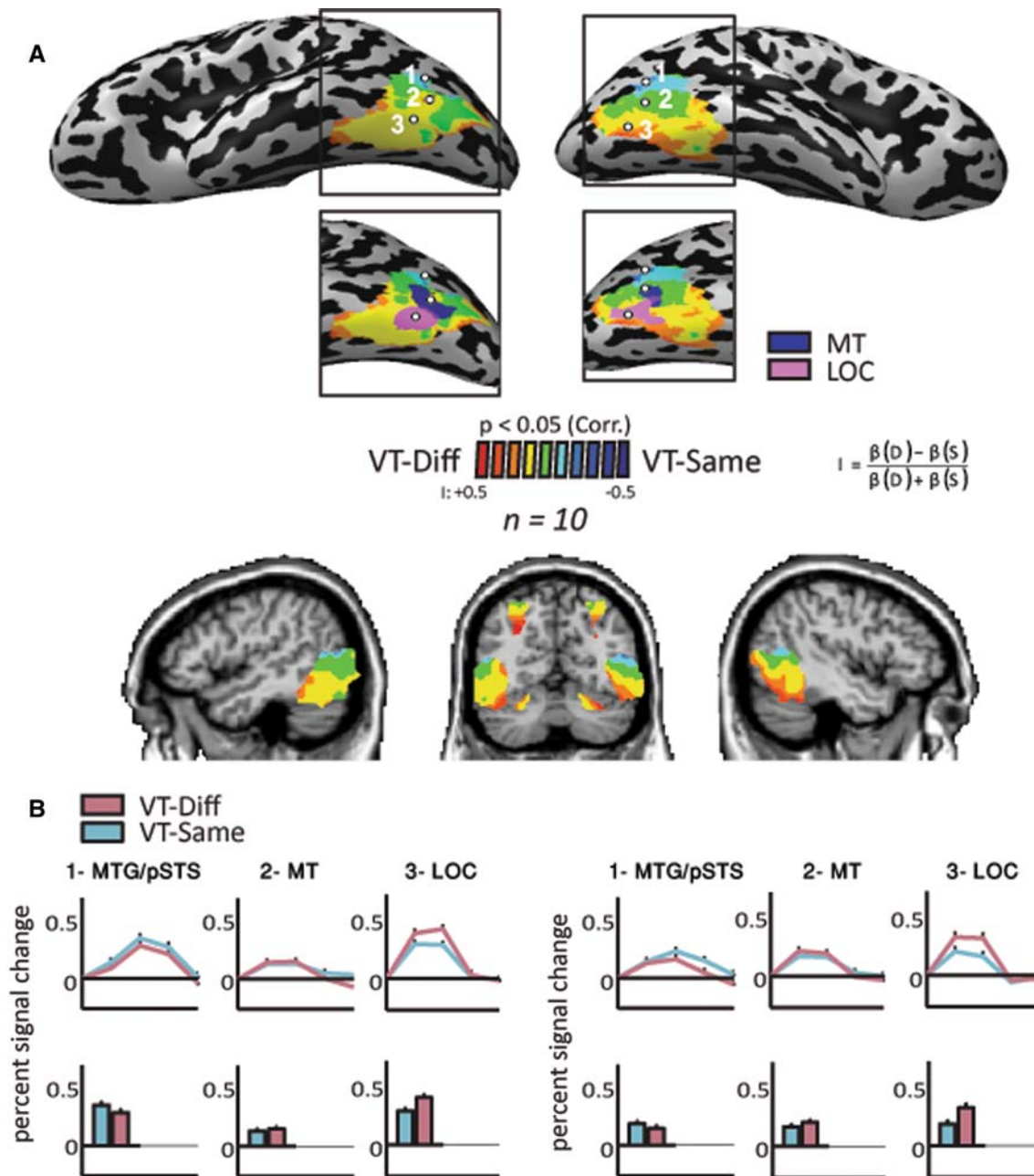
### The advantages of using fMR-A to study multisensory interactions

The invariant properties of human cortical neurons cannot be studied directly by fMRI due to its limited spatial resolution. One voxel obtained from an fMRI scan contains several hundred thousand neurons. Therefore, the fMRI signal may average out a heterogeneous group of highly selective neurons. The fMR-A method can tag specific neuronal populations within an area and investigate their functional properties non-invasively, and thus, can provide a powerful tool for assessing the functional properties of cortical neurons beyond the spatial resolution of several mm imposed by conventional fMRI. So far, this technique has successfully been used to explore many important issues in unisensory processing in humans, and specifically to study visual shape processing, but to the best of our knowledge, has never been used to study multisensory shape integration. We used a novel multi-sensory fMR-A paradigm to investigate whether fMRI adaptation would indicate multisensory voxels showing multisensory adaptation.

Our study highlights the power of using fMR-A to study crossmodal integration in humans, and may lead us to a better understanding of neural organization beyond what standard imaging techniques and fMR-A uni-modal studies can achieve, leading to a better grasp of the shared or common properties and transformations of information across sensory modalities. By manipulating experimental parameters and testing recovery from adaptation it should be possible to enhance our knowledge of the functional properties of cortical neurons which are beyond the spatial resolution limits imposed by fMRI. However, given the complexity of the neural processes, some cautionary comments are in order, such that experimental designs can be planned to circumvent these pitfalls as much as possible.

### fMRI-A putative caveats

Functional magnetic resonance imaging adaptation fMR-A is an increasingly popular method that aims to provide insights into the functional properties of subpopulations of neurons within an imaging voxel. However, when mapping between neural firing and hemodynamic response, caution should be exercised in interpreting the meaning of BOLD changes in adaptation paradigms such as this, both in general and for multisensory research in particular. It has been shown that local field potentials (LFPs) correlate with the BOLD signal better than multi- or single-unit activity in the macaque monkey (Logothetis et al. 2001). Thus, a region showing fMRI-A may not be transmitting fewer spikes but



**Fig. 4** Indications for a topographical gradient in VT-adaptation in the lateral-ventral occipital-temporal cortex. **a** Results of cross subject ( $n = 10$ ) GLM Two-color contribution analysis (see methods) are presented on MNI Talairach normalized inflated brain in a lateral-ventral view. Color scale denotes the relative preference of each voxel for one of the two conditions (VT-diff and VT-same). Voxels showing preference for VT-diff appear in red ( $I > 0$ ), and those with VT-same preference ( $I < 0$ ) are in blue. The inset present the same map aligned

with MT (blue) and LO (pink) Localizers. **b** We defined three ROIs along the dorsal-ventral axis of the occipital-temporal cortex: The most dorsal ROI is the MTG/pSTS (anatomical localizer), the most ventral ROI is LOC (using objects vs. scrambled images localizer) and intermediate MT ROI (using the MT localizer; see “Methods” for more details on localizers and ROI selection). The average time course (across subjects) and percent signal change histograms for these three ROIs are presented

may either be showing a reduced afferent input or reduced local processing. Adaptation may reflect a proportional reduction in firing rate to repetitions of a specific stimulus, a change in the tuning of neural responses to repeated stimuli, or shortening of the processing time for repeated

stimuli (for a recent review see Grill-Spector et al. 2006). The present study cannot differentiate these three forms of VT integration. This highlights one possible dissociation between fMRI-A and response suppression as recorded in single unit work in monkeys (Desimone 1996). On the

other hand, Sawamura et al. (2006) showed that two stimuli that activate the same neuron will elicit some adaptation. Conversely, if one finds fMRI-adaptation for a pair of stimuli, it is likely that these two stimuli activate the same neurons. Winston et al. (2004) suggested that fMRI experiments based on adaptation are not uniquely problematic in this regard, but that this is a more general interpretational issue in unifying electrophysiological and fMRI work (see Henson and Rugg 2003 for a more extensive discussion of hemodynamic decreases and response suppression).

Taken together, it is clear that the link between fMR-A and neuronal tuning is far from straightforward. Ascertaining the correct interpretation of fMR-A data will require further single-cell studies combined with fMR-A investigations in the monkey. This is even less clear for multisensory integration as there is little if any work that combines physiological recording of multisensory stimuli with an fMR-A paradigm. Thus even if the assumptions constituting the basis of this method are valid for unisensory processing, they will require verification in future crossmodal adaptation paradigms.

#### The role of possible confounding factors

Beyond the question of whether the fMR-A fully represents integration at the neuronal level, other confounding factors must be considered. For instance, as is the case in most fMR-A, our experimental paradigm includes a congruency (VT-same)/incongruency (VT-diff) component. It could be claimed that our VT-diff versus VT-same contrast findings simply stem from a classical multisensory congruent/incongruent manipulation. However, there is one critical difference between standard and adaptation paradigms using the congruency manipulation. To the best of our knowledge, most previous congruent/incongruent experiments have used a simultaneous presentation of visual-tactile, or the more widely used visual-auditory design (e.g. Hein et al. 2007, 2008; Naumer et al. 2008; Noppeney et al. 2007; van Atteveldt et al. 2004, 2007; Blau et al. 2008), at least in terms of the hemodynamic TR time resolution. This is, indeed, the right choice when testing for standard multisensory integration. However, in adaptation paradigms, as in our experiment, the crossmodal stimuli were sequential and not simultaneous. In fact, our prediction was to find opposite results in simultaneous integration versus sequential (adaptation) designs, in bimodal-responsive voxels. Thus, we expected that areas containing bimodal visuo-haptic neurons would show a higher BOLD response for simultaneous congruent stimuli in similar areas in which we found reduced activation for visual-haptic sequential stimuli. Preliminary results from the James group suggested this was indeed the case for VT object integration (increase in the

bimodal response in LOTv and IPS for congruent visual-tactile objects reported by Sunha et al. (2008), whereas in similar areas we found reduced activation for congruent visual-haptic sequential stimuli). One way to circumvent this problem is to look for repetition suppression of the same repeating bimodal events. This is exactly what we did when comparing the first and second exposures of each VT event (averaged across congruent and incongruent VT event). Our results (Figs. 2, 3) clearly showed that fMR adaptation is also evident in this case, irrespective of the incongruency factor. Alternatively, this effect could be explained by VV and TT unisensory repetition suppression effects, since the objects were also repeated within-modality from the first to the second presentation.

In addition, we carried out an additional experiment, Experiment 3, which was designed to avoid incongruency effects by comparing the first with the second and fourth repetition of the same objects by searching for voxels in the entire brain that show fMR-A. Our reasoning was that there are neurons responding to tactile input (in LOTv for instance), we would expect to see tactile adaptation for objects, but this time without the putative confound of the incongruency effect. We analyzed all the ROIs found in the main contrast for such fMRI-A effects and found that in similar areas which showed visual-tactile adaptation, there was higher activation for the first exposure of the object versus the second and fourth exposures (Figs. 2c, 3c). This strengthens our claim that the areas activated in the main contrast represent areas which go through adaptation and not a congruency effect.

Another possible confounding factor might be task differences. The subjects' task in the visual-tactile adaptation experiment was to identify both the visual and the tactile objects that were presented to them sequentially. To counter the argument that there were task differences between the visual and tactile events it is important to stress that VT-diff and VT-same events were pseudo-randomly interleaved so that the subjects did not know which event to expect. The fact that subjects had no cues to guide them meant they had to recognize both the visual and tactile objects in all types of events. This helped reduce task difference confounds to the different events which are inherent to any adaptation design. In addition, we also ran one more analysis of our data set to test whether the results were the sole outcome of task selections between the two main conditions (VT-diff and VT-same). We looked for VT adaptation effects of repeated stimuli within each run at an order of magnitude of tens of seconds. We found adaptation in all of the reported regions which showed visual-tactile adaptation, as shown in Figs. 2 and 3 and generally a very large overlap between the two types of analysis (VT-diff > VT-same and VT-1st > VT-2nd presentation). Finally, to further address this issue, we conducted two additional

experiments, Experiments 2 and 3, which both used a one-back comparison between the currently presented stimulus and the previous one, which were aimed to test the control and arousal effects and the specificity of our effects. These experiments were purely tactile. Subjects touched the same or different objects, or the same or different textures. Performance was similar in all conditions, which further strengthens the conclusion that the differences between the conditions were not due to attention and arousal confounds. When analyzing the time course of these experiments in the ROIs chosen as the peak activation areas for the main contrast, we found supportive evidence for our main results indicating tactile adaptation responses in the network of areas showing visuo-tactile adaptation responses. The findings from these control experiments further support (to some extent) our conclusions that real tactile processing of objects occurs in the areas showing visuo-tactile adaptation responses, including in area LOTv, regardless of a possible congruency effect (see Figs. 2, 3).

Multisensory responses in the LOTv and the nature of object representation in the ventral visual pathway

Visual–tactile convergence of object-related information occurs in the LOTv (Amedi et al. 2001), which is a sub-region within the human LOC. The defining feature of this region is that it is robustly activated during both visual and tactile object recognition. It shows a preference for objects compared to textures and scrambled objects in both modalities both in our results and in others (Amedi et al. 2002; Pietrini et al. 2004; Peltier et al. 2007; Sathian 2005; Weisser et al. 2004) and is only negligibly activated by the motor, naming, and visual imagery components of object recognition. As shown by Pietrini et al. (2004), these category-related responses are correlated across touch and vision, suggesting that a common representation of 3D objects is activated by both these modalities. In order to reduce the potential influence of naming and visual imagery, Reed et al. (2004) used abstract and nonsense 3D shapes combined with familiar and meaningful 3D objects to show activation of the LOTv in both cases. James et al. (2002) also found fMRI activation in occipital areas during haptic exploration of novel abstract objects. He further demonstrated that the magnitude of tactile-to-visual priming was similar to that of visual-to-visual priming, supporting the idea that vision and touch share common representations (see also Easton et al. 1997; Reales and Ballesteros 1999). Interestingly, while some studies refer to repetition suppression as a possible neural correlate of priming (e.g. Grill-Spector et al. 2006; Henson and Rugg 2003), several other studies have shown that in some cases, priming can be associated with increased activity rather than reduction (Dolan et al. 1997; Grill-Spector et al. 2006;

Henson 2003; Henson and Rugg 2003). Several studies even reported a combination of repetition suppression in some brain areas and repetition enhancement in others when performing cross modal priming (e.g. Badgaiyan et al. 2001, 1999; Schacter et al. 1999). Here we tested for such a repetition enhancement, and found no significant effect. Taken together, we decided to focus, in the current study, on crossmodal repetition suppression as a marker of crossmodal adaptation rather than priming. Future studies should test for a crossmodal repetition suppression effect of more than one crossmodal repetition (like Grill-Spector et al. 1999; Sayers and Grill-Spector 2005 in the unisensory domain), which might help to better differentiate between priming and repetition suppression effects.

To conclude, by using fMR-A to study crossmodal representations of objects we showed that area LOTv, which is known for its crossmodal nature, is truly a region with bimodal neurons in humans and in fact evidenced the most significant VT adaptation effect (Figs. 2, 3; see also Suppl. Fig. 1).

Multisensory responses in area MT

The MT area is located in the lateral occipital cortex and is known to be a complex motion area which responds better to moving than to stationary stimuli (Dupont et al. 1994; Tootell et al. 1995; Watson et al. 1993; Zeki et al. 1991). Two recent human functional neuroimaging studies suggest that MT might play a role in processing motion in the tactile modality. Hagen et al. (2002) found greater activity in area MT when a small brush stroked the length of a subject's arms than that during a fixation control. Blake et al. (2004) found greater activity in area MT as subjects grasped a rotating plastic ball than when they grasped a stationary ball. Regardless of the source of tactile responses in MT, they are relatively weak. The reported MT activations are much weaker for moving tactile than moving visual stimuli (0.8% MR signal change versus 0.2% in the Blake study). Our study, however, suggests that while area MT shows robust tactile responses, it does not show crossmodal adaptation for objects. It would be interesting to test whether this area shows cross modal adaptation to moving stimuli rather than objects.

Gradient in the occipito-temporal cortex from no to full adaptation

In the visual cortex, we found a gradient from no adaptation at all in the dorsal stream, and a gradual increase in multi-sensory preference reaching its maximum in area LOTv (Fig. 4). This suggests that the LO part of the LOC (i.e. the LOTv) is the peak adaptation; hence, we expect to find the maximum proportion of bimodal visuo-haptic

object-related neurons in this area, an effect that drops off when moving from LOTv both dorsally (to area MT) and ventrally. The origins of this highly specific bimodal processing in LOTv are still not clear although effective connectivity studies by Sathian's group (Deshpande et al. 2008; Peltier et al. 2007) provide a crucial link to understanding the role of these bimodal neurons in the visual cortex.

#### Tactile responses in the calcarine sulcus

Using fMR-A, we showed an adaptation effect in the Calcarine sulcus (Figs. 2, 3). Previous studies have demonstrated robust activation of the calcarine sulcus in the blind (but not in sighted subjects) during tactile tasks such as Braille character recognition, discrimination of vibrotactile stimuli, fine spatial discrimination, and haptic object recognition (Amedi et al. 2002; Buchel et al. 1998; Hadjikhani and Roland 1998; Prather et al. 2004; Sadato et al. 1996, 1998). Furthermore, Wittenberg et al. (2004) show enhanced connectivity between primary somatosensory cortex and primary visual cortex in early blind subjects. In contrast, a few other recent studies showed sporadic but significant tactile activation in the Calcarine sulcus in sighted individuals (Burton et al. 2004, 2006; Merabet et al. 2007).

Furthermore, complete and transient visual deprivation in sighted subjects (i.e., 5 days of blindfolding) seems to be sufficient to lead to recruitment of the primary visual cortex for tactile and auditory processing (Pascual-Leone and Hamilton 2001; Merabet et al. 2008), including tactile object recognition (Amedi et al. 2006). The rapidity of the neuroplastic changes supports the view of preexisting connections projecting to visual cortical areas which may be selectively activated for the crossmodal recruitment of visual cortical areas for non-visual tasks. After prolonged visual deprivation, these same connections may become enhanced through synaptic reinforcement or greater recruitment of contributory neural networks based on task demands. The activation in the Calcarine sulcus reported here (see Figs. 2, 3) offers further support to this view.

However, in contrast to activations found in the LOTv, activation in the Calcarine sulcus showed less consistency in control Experiment 2 with no tactile object selectivity (Figs. 2, 3) and no visual object selectivity (not shown). This is consistent with a recent fMRI study (Merabet et al. 2007) reporting clear tactile responses in the primary visual cortex (confirmed by the clear tactile adaptation effect found in the current study in Experiment 3), but with no clear selectivity for any of the tactile tasks also used in that experiment (distance or roughness judgment). Tactile responses in V1 might thus reflect top-down modulation (see for instance Lacey et al. 2007) or anticipatory effects (Yevgeniy and Aniruddha 2009).

#### Cross modality in the aIPS

Sathian et al. (1997) reported robust tactile activation in the parieto-occipital cortex, a region previously regarded as part of the dorsal visual stream. In more recent studies, this group also found a number of bilateral parietal regions that were shape-selective for haptic as well as visual stimuli peaking in the aIPS and multiple parts of the IPS (Peltier et al. 2007). These findings are in keeping with reports of multisensory shape-selectivity and suggest that the IPS plays a crucial role in binding visual and tactile information (Amedi et al. 2005, 2007; Grefkes et al. 2002; Saito et al. 2003), and multisensory responses in the IPS in monkeys (Iriki et al. 1996). This is in line with our findings that show that the IPS is activated during visual object epochs compared to visual scrambled epochs (Fig. 4a). This emphasizes its shape-selectivity, and is in line with our findings that show the second most significant crossmodal repetition effect in the left aIPS (Figs. 2, 3; see also Supplementary Fig. 1).

#### Multisensory responses in the precentral sulcus and the insula

In contrast to area LOTv and the aIPS, the question as to whether the precentral sulcus and the insula are multisensory visuo-haptic object-relation integration sites is still highly debated. While many studies (including some of our own) did not find any such response (Amedi et al. 2001, 2002; James et al. (2002); Peltier et al. 2007; Pietrini et al. 2004; Saito et al. 2003; Weisser et al. 2004), there is sporadic evidence suggesting both regions do contribute to various aspects of multisensory integration of objects and possibly also other types of stimuli.

For instance, Hein and colleagues found multisensory integration and congruency effects for auditory and visual objects in the neighboring (probably more ventral) prefrontal cortex (Hein et al. 2007). We recently found convergence of activation in the precentral sulcus on a tactile shape recognition task, and visual objects converted into auditory soundscapes created using a visual-to-auditory sensory substitution algorithm (Amedi et al. 2007). Taken together with the results presented here showing significant fMR-A in the precentral sulcus in prefrontal cortex (though somewhat smaller in magnitude compared to the adaptation found in the occipital cortex (see Figs. 2, 3), we suggest that this area might indeed contain a combination of bimodal or even trimodal neurons supporting visual, tactile, and auditory object integration.

Some studies have suggested that the right Insula might play a role in binding visual and tactile information by serving as a mediating area enabling unisensory areas to communicate and exchange (Hadjikhani and Roland 1998;

Prather et al. 2004), and the notion that the Insula is densely connected to various sensory information (Amedi et al. 2005, 2007). This is based on multisensory matching experiments areas and thus might be an ideal candidate for this function. While there are some reports of right lateralization of activation in the Insula, other studies indicate left lateralization of activation during visual–tactile integration (Banati et al. 2000).

The suggestion that the Insula is more involved in the transfer and/or binding of information between sensory modalities, but is perhaps less involved in object shape analysis and recognition per-se (Amedi et al. 2005, 2007), is in line with our findings that the Insula shows clear fMR-A (Figs. 2, 3), but did not demonstrate significant activation during unimodal visual object epochs compared to visual scrambled epochs (the MASK used for Figs. 2a, 3a).

## Conclusion

Our study suggests that a network of areas with bi-modal neurons form an important part of the visuo-haptic integration of objects. Using fMR-A, we were able to confirm that the LOtv and the aIPS show a repetition suppression effect, suggesting that they are bimodal visuo-haptic object-related integration sites. We also suggest that other areas, namely the left precentral sulcus, the insula, and the calcarine sulcus, might also contain a large proportion of visuo-haptic neurons. Another interesting finding is the gradient of adaptation found in the lateral occipital cortex.

Furthermore, fMR-A has successfully been used to study many important questions in unisensory (mostly visual) research in humans. We suggest here that fMR-A is a powerful tool for multisensory research in humans as it provides an elegant way to study multisensory integration to gain insights beyond voxel resolution in normal healthy subjects (though there are some questions and uncertainties about using this method rather than directly recording from neurons in humans as is possible in some clinical cases, see Krekelberg et al. 2006 and “Discussion”). This method directly responds to questions about crossmodal interactions and integration which are the core enigmas motivating multisensory research in humans.

**Acknowledgments** We thank Ran Geva and Uri Hertz for their help in running the experiments. We thank Zohar Tal, Ella Striem and Haim Azulai for their help in preparing and analyzing the control experiments. We wish to thank Ehud Zohary and Rafael Malach for their valuable advice on an early version of the experimental design. This work was supported by the Human Frontiers Science Program Career Development Award (to AA), an EU-FP7 MC International Reintegration Grant, Israel Science Foundation grant (ISF 1684/08) and GIF-young grant (To AA) and the generous support of the Moscona foundation.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

## References

- Amedi A, Malach R, Hendler T et al (2001) Visuo-haptic object-related activation in the ventral visual pathway. *Nat Neurosci* 4:324–330
- Amedi A, Jacobson G, Hendler T et al (2002) Convergence of visual and tactile shape processing in the human lateral occipital complex. *Cereb Cortex* 12:1202–1212
- Amedi A, Von Kriegstein K, Van Atteveldt N et al (2005) Functional imaging of human crossmodal identification and object recognition. *Exp Brain Res* 166:559–571
- Amedi A, Camprodon J, Merabet L, Bempohl F, Haligan E, Bass-Pit-skel N, Ronen I, Kim D, Pascual-Leone A (2006) Highly transient activation of primary visual cortex (V1) for tactile processing in sighted following 5 days of blindfolding. Abstracts of the Annual Meeting of the Society for Neuroscience (Atlanta, USA)
- Amedi A, Stern WM, Camprodon JA et al (2007) Shape conveyed by visual-to-auditory sensory substitution activates the lateral occipital complex. *Nat Neurosci* 10:687–689
- Badgaiyan RD, Schacter DL, Alpert NM (1999) Auditory priming within and across modalities: evidence from positron emission tomography. *J Cogn Neurosci* 11:337–348
- Badgaiyan RD, Schacter DL, Alpert NM (2001) Priming within and across modalities: exploring the nature of rCBF increases and decreases. *Neuroimage* 13:272–282
- Banati RB, Goerres GW, Tjoa C, Aggleton JP, Grasby P (2000) The functional anatomy of visual–tactile integration in man: a study using positron emission tomography. *Neuropsychologia* 38:115–124
- Blake R, Sobel KV, James TW (2004) Neural synergy between kinetic vision and touch. *Psychol Sci* 15:397–402
- Blau V, van Atteveldt N, Formisano E, Goebel R, Blomert L (2008) Task-irrelevant visual letters interact with the processing of speech sounds in heteromodal and unimodal cortex. *Eur J Neurosci* 28:500–509
- Braendgraard H, Evans SM, Howard CV, Gundersen HJ (1990) The total number of neurons in the human neocortex unbiasedly estimated using optical disectors. *J Microsc* 157:285–304
- Buchel C, Price C, Frackowiak RS, Friston K (1998) Different activation patterns in the visual cortex of late and congenitally blind subjects. *Brain* 121(Pt 3):409–419
- Buckner RL, Petersen SE, Ojemann JG, Miezin FM, Squire LR, Raichle ME (1995) Functional anatomical studies of explicit and implicit memory retrieval tasks. *J Neurosci* 15:12–29
- Burton H, Sinclair RJ, McLaren DG (2004) Cortical activity to vibrotactile stimulation: an fMRI study in blind and sighted individuals. *Hum Brain Mapp* 23:210–228
- Burton H, McLaren DG, Sinclair RJ (2006) Reading embossed capital letters: an fMRI study in blind and sighted individuals. *Hum Brain Mapp* 27:325–339
- Demb JB, Desmond JE, Wagner AD, Vaidya CJ, Glover GH, Gabrieli JD (1995) Semantic encoding and retrieval in the left inferior prefrontal cortex: a functional MRI study of task difficulty and process specificity. *J Neurosci* 15:5870–5878
- Deshpande G, Hu X, Stilla R et al (2008) Effective connectivity during haptic perception: a study using Granger causality analysis of functional magnetic resonance imaging data. *Neuroimage* 40:1807–1814



- Desimone R (1996) Neural mechanisms for visual memory and their role in attention. *Proc Natl Acad Sci* 93:13494–13499
- Dolan RJ, Fink GR, Rolls E, Booth M, Holmes A, Frackowiak RSJ, Friston KJ (1997) How the brain learns to see objects and faces in an impoverished context. *Nature* 389:596–599
- Driver J, Noesselt T (2008) Multisensory interplay reveals crossmodal influences on ‘sensory specific’ brain regions, neural responses, and judgments. *Neuron* 57:11–23
- Dupont P, Orban GA, De Bruyn B et al (1994) Many areas in the human brain respond to visual motion. *J Neurophysiol* 72:1420–1424
- Easton RD, Greene AJ, Srinivas K (1997) Transfer between vision and haptics: memory for 2-D patterns and 3-D objects. *Psychonom Bull Rev* 4:403–410
- Feinberg TE, Rothi LJ, Heilman KM (1986) Multimodal agnosia after unilateral left hemisphere lesion. *Neurology* 36:864–867
- Forman SD, Cohen JD, Fitzgerald M, Eddy WF, Mintun MA, Noll DC (1995) Improved assessment of significant activation in functional magnetic imaging (fMRI): use of cluster-size threshold. *Magn Reson Med* 33:636–647
- Friston KJ, Holmes AP, Worsley KJ et al (1995) Statistical parametric maps in functional imaging: a general linear approach. *Hum Brain Mapp* 2:189–210
- Friston KJ, Holmes AP, Worsley KJ (1999) How many subjects constitute a study? *Neuroimage* 10:1–5
- Goodale MA (2000) Occipitotemporal and occipitoparietal visual pathways in the primate brain. *Adv Neurol* 84:75–86
- Grefkes C, Weiss PH, Zilles K et al (2002) Crossmodal processing of object features in human anterior intraparietal cortex: an fMRI study implies equivalencies between humans and monkeys. *Neuron* 35:173–184
- Grill-Spector K, Kushnir T, Edelman S et al (1999) Differential processing of objects under various viewing conditions in the human lateral occipital complex. *Neuron* 24:187–203
- Grill-Spector K, Henson R, Martin A (2006) Repetition and the brain: neural models of stimulus-specific effects. *Trends Cogn Sci* 10:14–23
- Hadjikhani N, Roland PE (1998) Cross-modal transfer of information between the tactile and the visual representations in the human brain: a positron emission tomographic study. *J Neurosci* 18:1072–1084
- Hagen MC, Franzen O, McGlone F et al (2002) Tactile motion activates the human middle temporal/V5 (MT/V5) complex. *Eur J Neurosci* 16:957–964
- Hein G, Doehrmann O, Müller NG et al (2007) Object familiarity and semantic congruency modulate responses in cortical audiovisual integration areas. *J Neurosci* 27:7881–7887
- Henson RNA (2003) Neuroimaging studies of priming. *Prog Neurobiol* 70:53–81
- Henson RNA, Rugg MD (2003) Neural response suppression, haemodynamic repetition effects, and behavioural priming. *Neuropsychologia* 41:263–270
- Henson R, Shallice T, Dolan R (2000) Neuroimaging evidence for dissociable forms of repetition priming. *Science* 287:1269–1272
- Iriki A, Tanaka M, Iwamura Y (1996) Coding of modified body schema during tool use by macaque postcentral neurones. *Neuroreport* 7:2325–2330
- James TW, Humphrey GK, Gati JS et al (2002) Haptic study of three-dimensional objects activates extrastriate visual areas. *Neuropsychologia* 40:1706–1714
- James TW, James KH, Humphrey GK, Goodale MA (2006) Do visual and tactile object representations share the same neural substrate? In: Heller MA, Ballesteros S (eds) *Touch and blindness: psychology and neuroscience*. Lawrence Erlbaum Associates, Mahwah, NJ, pp 139–155
- Jiang Y, Haxby JV, Martin A, Ungerleider LG, Parasuraman R (2000) Complementary neural mechanisms for tracking items in human working memory. *Science* 287:643–646
- Klatzky RL, Lederman S, Reed C (1987) There’s more to touch than meets the eye: the salience of object attributes for haptics with and without vision. *J Exp Psychol* 116:356–369
- Kourtzi Z, Kanwisher N (2000) Cortical regions involved in perceiving object shape. *J Neurosci* 20:3310–3318
- Kourtzi Z, Kanwisher N (2001) Representation of perceived object shape by the human lateral occipital complex. *Science* 293:1506–1509
- Kourtzi Z, Tolias AS, Altmann CF et al (2003) Integration of local features into global shapes monkey and human fMRI studies. *Neuron* 37:333–346
- Krekelberg B, Boynton GM, van Wezel RJ (2006) Adaptation: from single cells to BOLD signals. *Trends Neurosci* 29:250–256
- Lacey S, Campbell C, Sathian K (2007) Vision and touch: multiple or multisensory representations of objects? *Perception* 36:1513–1521
- Levy I, Hasson U, Malach R (2004) One picture is worth at least a million neurons. *Curr Biol* 11:996–1001
- Li L, Miller EK, Desimone R (1993) The representation of stimulus familiarity in anterior inferior temporal cortex. *J Neurophysiol* 69:1918–1929
- Logothetis NK, Pauls J, Augath M et al (2001) Neurophysiological investigation of the basis of the fMRI signal. *Nature* 412:150–157
- Malach R, Reppas JB, Benson RR et al (1995) Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. *Proc Natl Acad Sci USA* 92:8135–8139
- Marr D (1982) *Vision: a computational investigation into the human representation and processing of visual information*. WH Freeman, San Francisco, CA
- Merabet LB, Swisher JD, McMains SA, Halko MA, Amedi A, Pascual-Leone A, Somers DC (2007) Combined activation and deactivations of visual cortex during tactile sensory processing. *J Neurophysiol* 97:1633–1641
- Merabet LB, Hamilton R, Schlaug G, Swisher JD, Kiriakopoulos ET, Pitskel NB, Kauffman T, Pascual-Leone A (2008) Rapid and reversible recruitment of early visual cortex for touch. *PlosOne* 3. doi:10.1371/journal.pone.0003046
- Meredith MA (2002) On the neuronal basis for multisensory convergence: a brief overview. *Cogn Brain Res* 14:31–40
- Miller EK, Desimone R (1994) Parallel neuronal mechanisms for short-term memory. *Science* 263:520–522
- Naccache L, Dehaene S (2001) The priming method: imaging unconscious repetition priming reveals an abstract representation of number in the parietal lobes. *Cereb Cortex* 11:966–974
- Naumer MJ, Doehrmann O, Müller NG, Muckli L, Kaiser J, Hein G (2008) Cortical plasticity of audio-visual object representations. *Cereb Cortex*. doi:10.1093/cercor/bhn200
- Noppeney U, Josephs O, Hocking J, Price CJ, Friston KJ (2007) The effect of prior visual information on recognition of speech and sounds. *Cereb Cortex* 18:598–609
- Pascual-Leone A, Hamilton RH (2001) The metamodal organization of the brain. *Prog Brain Res* 134:427–445
- Peltier S, Stilla R, Mariola E et al (2007) Activity and effective connectivity of parietal and occipital cortical regions during haptic shape perception. *Neuropsychologia* 45:476–483
- Pietrini P, Furey ML, Ricciardi E et al (2004) Beyond sensory images: object-based representation in the human ventral pathway. *Proc Natl Acad Sci USA* 101:5658–5663
- Prather SC, Votaw JR, Sathian K (2004) Task-specific recruitment of dorsal and ventral visual areas during tactile perception. *Neuropsychologia* 42:1079–1087

- Reales JM, Ballesteros S (1999) Implicit and explicit memory for visual and haptic objects: cross-modal priming depends on structural descriptions. *J Exp Psychol: Learn Mem Cogn* 25:644–663
- Reed CL, Shoham S, Halgren E (2004) Neural substrates of tactile object recognition: an fMRI study. *Hum Brain Mapp* 21:236–246
- Rockel AJ, Hiorns RW, Powell TPS (1980) The basic uniformity in structure of the neocortex. *Brain* 103:221–244
- Sadato N, Pascual-Leone A, Grafman J, Ibanez V, Deiber MP, Dold G, Hallett M (1996) Activation of the primary visual cortex by Braille reading in blind subjects. *Nature* 11:526–528
- Sadato N, Pascual-Leone A, Grafman J, Deiber MP, Ibanez V, Hallett M (1998) Neural networks for Braille reading by the blind. *Brain* 121:1213–1229
- Saito DN, Okada T, Morita Y et al (2003) Tactile–visual cross modal shape matching: a functional MRI study. *Brain Res Cogn Brain Res* 17:14–25
- Sathian K (2005) Visual cortical activity during tactile perception in the sighted and the visually deprived. *Dev Psychobiol* 46:279–286
- Sathian K, Lacey S (2007) Journeying beyond classical somatosensory cortex. *Can J Exp Psychol* 61:254–264
- Sathian K, Zangaladze A, Green J et al (1997) Tactile spatial acuity and roughness discrimination: impairments due to aging and Parkinson's disease. *Neurology* 49:168–177
- Sawamura H, Orban GA, Vogels R (2006) Selectivity of neuronal adaptation does not match response selectivity: a single-cell study of the fMRI adaptation paradigm. *Neuron* 49:307–318
- Sayers R, Grill-Spector K (2005) Object-selective cortex exhibits performance-independent repetition suppression. *J Neurophysiol* 95:995–1007
- Schacter DL, Badgaiyan RD, Alpert NM (1999) Visual word stem completion priming within and across modalities: A PET study. *Neuroreport* 10:2061–2065
- Self MW, Zeki S (2005) The integration of colour and motion by the human visual brain. *Cereb Cortex* 15:1270–1279
- Shmuelof L, Zohary E (2005) Dissociation between ventral and dorsal fMRI activation during object and action recognition. *Neuron* 47:457–470
- Sigala R, Rainer G (2007) Visual neuroscience: face-encoding mechanisms revealed by adaptation. *Curr Biol* 17:R20–R22
- Sobotka S, Ringo JL (1996) Mnemonic responses of single units recorded from monkey inferotemporal cortex, accessed via trans-commissural versus direct pathways: a dissociation between unit activity and behavior. *J Neurosci* 16:4222–4230
- Stern CE, Corkin S, González RG, Guimaraes AR, Baker JR, Jennings PJ, Carr CA, Sugiura RM, Vedantham V, Rosen BR (1996) The hippocampal formation participates in novel picture encoding: evidence from functional magnetic resonance imaging. *Proc Natl Acad Sci USA* 93:8660–8665
- Stilla R, Sathian K (2008) Selective visuo-haptic processing of shape and texture. *Hum Brain Mapp* 29:1123–1138
- Stoesz M, Zhang M, Weisser VD, Prather SC, Mao H, Sathian K (2003) Neural networks active during tactile form perception: common and differential activity during macrospatial and microspatial tasks. *Int J Psychophysiol* 50:41–49
- Sunha K, Eylath D, Stevenson RA, James TW (2008) Evidence for ventral and dorsal neural pathways for visuo-haptic object recognition. *IMRF abstract*
- Talairach J, Tournoux P, Rayport M (1988) Co-planar stereotaxic atlas of the human brain: 3-dimensional proportional system: an approach to cerebral imaging. Thieme, New York
- Tootell RB, Reppas JB, Kwong KK et al (1995) Functional analysis of human MT and related visual cortical areas using magnetic resonance imaging. *J Neurosci* 15:3215–3230
- van Atteveldt NM, Formisano E, Goebel R, Blomert L (2004) Integration of letters and speech sounds in the human brain. *Neuron* 43:271–282
- van Atteveldt NM, Formisano E, Blomert L, Goebel R (2007) The effect of temporal asynchrony on the multisensory integration of letters and speech sounds. *Cereb Cortex* 17:962–974
- van Turennout M, Ellmore T, Martin A (2000) Long-lasting cortical plasticity in the object naming system. *Nat Neurosci* 3:1329–1334
- Watson JD, Myers R, Frackowiak RS et al (1993) Area V5 of the human brain: evidence from a combined study using positron emission tomography and magnetic resonance imaging. *Cereb Cortex* 3:79–94
- Weisser VD, Stilla R, Prather SC et al (2004) Multisensory cortical processing of object shape and its relation to mental imagery. *Cogn Affect Behav Neurosci* 4:251–259
- Winston JS, Henson RN, Fine-Goulden MR et al (2004) fMRI adaptation reveals dissociable neural representations of identity and expression in face perception. *J Neurophysiol* 92:1830–1839
- Wittenberg GF, Werhahn KJ, Wassermann EM, Herscovitch P, Cohen LG (2004) Functional connectivity between somatosensory and visual cortex in early blind humans. *Eur J NeuroSci* 20:1923–1927
- Yevgeniy BS, Aniruddha D (2009) Anticipatory haemodynamic signals in sensory cortex not predicted by local neuronal activity. *Nature* 457. doi:10.1038/nature07664
- Zeki SM (1978) Functional specialisation in the visual cortex of the rhesus monkey. *Nature* 274:423–428
- Zeki S, Watson JD, Lueck CJ et al (1991) A direct demonstration of functional specialization in human visual cortex. *J Neurosci* 11:641–649
- Zhang M, Weisser VD, Stilla R et al (2004) Multisensory cortical processing of object shape and its relation to mental imagery. *Cogn Affect Behav Neurosci* 4:251–259