

## Integrating Visual and Tactile Information in the Perirhinal Cortex

J. S. Holdstock<sup>1</sup>, J. Hocking<sup>2,3</sup>, P. Notley<sup>1</sup>, J. T. Devlin<sup>4</sup> and C. J. Price<sup>2</sup>

<sup>1</sup>School of Psychology, University of Liverpool, Liverpool L69 3BS, UK, <sup>2</sup>Wellcome Trust Centre for NeuroImaging, University College London, London WC1N 3BG, UK, <sup>3</sup>Centre for Magnetic Resonance, University of Queensland, Brisbane, Queensland 4072, Australia and <sup>4</sup>Cognitive, Perceptual and Brain Sciences, University College London, London WC1H 0AP, UK

**By virtue of its widespread afferent projections, perirhinal cortex is thought to bind polymodal information into abstract object-level representations. Consistent with this proposal, deficits in cross-modal integration have been reported after perirhinal lesions in nonhuman primates. It is therefore surprising that imaging studies of humans have not observed perirhinal activation during visual-tactile object matching. Critically, however, these studies did not differentiate between congruent and incongruent trials. This is important because successful integration can only occur when polymodal information indicates a single object (congruent) rather than different objects (incongruent). We scanned neurologically intact individuals using functional magnetic resonance imaging (fMRI) while they matched shapes. We found higher perirhinal activation bilaterally for cross-modal (visual-tactile) than unimodal (visual-visual or tactile-tactile) matching, but only when visual and tactile attributes were congruent. Our results demonstrate that the human perirhinal cortex is involved in cross-modal, visual-tactile, integration and, thus, indicate a functional homology between human and monkey perirhinal cortices.**

**Keywords:** cross-modal, fMRI, integration, perception, visual-tactile

The perirhinal cortex lies on the ventral surface of the anterior medial temporal lobe and, in nonhuman primates, receives afferent projections from adjacent unimodal visual processing regions TE and TEO, somatosensory association areas of the insula cortex, auditory association cortex in superior temporal gyrus, and polymodal association cortices such as the orbito-frontal cortex and the dorsal bank of the superior temporal sulcus (Friedman et al. 1986; Suzuki and Amaral 1994a; Carmichael and Price 1995; Suzuki 1996a, 1996b; Insausti et al. 1998). It also has strong reciprocal connections with the hippocampus, via the entorhinal cortex, and with the amygdala (Van Hoesen and Pandya 1975; Suzuki and Amaral 1994a, 1994b; Suzuki 1996a, 1996b). It is therefore well placed to combine inputs from different sensory modalities and to interact closely with memory-related regions (Murray and Bussey 1999; Murray and Richmond 2001). Indeed, in nonhuman primates, removal of rhinal cortex impairs tactile-visual delayed non-matching-to-sample (Goulet and Murray 2001) and flavor-visual association learning (Parker and Gaffan 1998).

In humans, however, the evidence for the role of the perirhinal cortex in visual-tactile processing is inconsistent. One group of patients with perirhinal damage, subsequent to herpes simplex encephalitis, was significantly more impaired on visual-tactile matching than visual-visual or tactile-tactile matching (Shaw et al. 1990), but this is not always the case (Nahm et al. 1993). Moreover, damage in patients with herpes simplex encephalitis is not limited to the perirhinal cortex but

includes many other medial and lateral anterior temporal lobe regions. Therefore, there are no examples of patient studies that tested visual-tactile integration following damage limited to the perirhinal cortex. Likewise, there are no reports of anterior medial temporal activation in functional imaging studies of visual-tactile integration. Instead, activation for visual-tactile relative to unimodal matching was observed in the right or left insula/clastrum (Hadjikhani and Roland 1998; Banati et al. 2000), the anterior intraparietal sulcus (Grefkes et al. 2002), and the posterior intraparietal sulcus (Saito et al. 2003). In summary, in nonhuman primates, perirhinal cortex appears to be involved in cross-modal integration of both learned associations between objects and when binding polymodal information from single objects, but it is currently unclear whether the perirhinal cortex has a similar role in humans.

We hypothesized that the absence of perirhinal activation in previous functional imaging studies of visual-tactile matching may partly be due to the choice of experimental design and analyses, which summed over congruent and incongruent cross-modal trials. This approach is not sensitive to activation that depends on whether visuo-tactile integration is successful or not. That is, when tactile and visual information come from the same stimulus, then the 2 information streams are congruent and can be integrated into an object-level representation, whereas when the information comes from 2 different objects, an integrated object-level representation cannot be generated. It follows that, if perirhinal activation is greater for congruent than incongruent stimuli, this would be consistent with a role either in the integration process itself or subsequent processing of the integrated object-level representation. In contrast, greater perirhinal activation for incongruent than congruent stimuli could reflect an unsuccessful integration process, integration effort, or dual processing of 2 unintegrated stimuli relative to 1 integrated stimulus (see Hocking and Price 2008 for a discussion of this prediction).

We therefore evaluated this hypothesis by investigating human perirhinal activation during congruent versus incongruent cross-modal (visual-tactile) versus unimodal (visual-visual and tactile-tactile) shape matching. Our expectation was that perirhinal activation would be higher on congruent than incongruent visual-tactile trials. However, we also tested the reverse hypothesis (higher perirhinal activation for incongruent than congruent visual-tactile matching) because a previous functional imaging study of audio-visual integration (Taylor et al. 2006) reported higher activation just behind the left perirhinal cortex (according to the criteria of Insausti et al. 1998) when visual and auditory stimuli were different (incongruent) rather than the same (congruent).

## Materials and Methods

### Subjects

Eighteen healthy, right-handed volunteers (mean age 21.3 years; standard deviation [SD], 3.1; 13 male) with normal vision gave informed written consent to participate. The study was approved by the joint ethics committee of the Institute of Neurology and University College London Hospital.

### Task Design

Scanning was conducted in 4 10-min 35-s runs. Throughout the experiment, subjects were instructed to use a 2-choice foot press response to indicate whether or not 2 simultaneously presented stimuli had identical shapes or not. One foot indicated a same response and the other foot indicated a different response. Within each run, there were 4 conditions of interest corresponding to a  $2 \times 2$  design with Modality (unimodal, cross-modal) and Congruency (congruent, incongruent) as independent factors. On half the unimodal trials, 2 visual stimuli were presented on a rear projection screen, 1 left of center, the other right of center; on the other half, 2 tactile stimuli were presented, 1 to each hand. During cross-modal matching, 1 visual and 1 tactile stimulus was presented, with the hand and side of the screen systematically varied (right hand and right side of the screen, left hand and left side of the screen, right hand and left side of the screen, and left hand and right side of the screen). Preliminary analyses indicated that perirhinal activation was not influenced by hand or screen side. Within all conditions, the 2 stimuli were congruent (i.e., had identical shapes) on 50% of trials and incongruent (i.e., had different shapes) on the other 50% of trials. Each individual subject responded with 1 foot for the congruent response and 1 foot for the incongruent response. The motor responses were controlled because a motor response was made on both congruent and incongruent trials during both cross-modal and unimodal matching (i.e., a motor response was made on every trial irrespective of which condition it was).

To minimize sensory processing differences between conditions and facilitate the presentation of the stimuli, bilateral visual and tactile processing was maintained on every trial by presenting wooden spheres and/or visually presented circles in addition to the stimuli being matched. Thus, in visual-visual matching trials, subjects also simultaneously manipulated 2 wooden spheres, 1 in each hand; in tactile-tactile matching trials, subjects simultaneously looked at 2 circles (1 on either side of the screen); and in visual-tactile matching conditions, a wooden sphere and a circle were presented in the hand and on the side of the screen that were not occupied by the stimuli being matched. In other words, subjects received 2 visual stimuli and 2 tactile stimuli on every trial. This approach provided continuous and predictable tactile and visual stimulation. The subjects were instructed to focus on the shape of the rectangular wooden blocks/visual shapes and ignore the wooden spheres and circles. Therefore, the congruent versus incongruent decisions always concerned the rectangles. They never involved the spheres/circles and they never involved the combination of rectangles and circles.

During a fifth condition, subjects viewed 2 circles and manipulated 2 wooden spheres while making alternate foot movements (i.e., no matching decision was required). As the shape of the visual and tactile stimuli was always congruent (circles and spheres), this condition was expected to show maximum activation in cross-modal integration areas. It can therefore be considered a ceiling (as opposed to a baseline) condition. It was included so that we could measure relative activation changes for each of our conditions of interest. The choice of baseline had no effect on our findings because our findings were entirely based on differences between our experimental conditions that were unrelated and independent of the baseline. Our experimental design did not include blocks of fixation with no stimulation; therefore, we cannot report activation for the baseline that is independent of the other conditions. Nor would this be useful or relevant to our findings. What is relevant is that the baseline involved continuous cross-modal congruency. Therefore, we can plot activation for each of our experimental conditions independently and aid interpretation by illustrating the relative effect sizes.

Within each scanning run, there were 2 blocks of each condition (i.e., 48 trials per type in total). All conditions were fully counterbalanced over subjects and within and between runs. In addition, we counterbalanced, across subjects, whether the left foot indicated a same response and the right foot a different response (or vice versa).

### Stimuli

Tactile stimuli were rectangular wooden blocks (10 cm  $\times$  2.5 cm  $\times$  1.2 cm) from which small rectangles and triangles were removed (see Fig. 1). Visual stimuli were silhouettes based on the same rectangular structure as the wooden blocks. Visual-tactile stimuli were silhouettes and wooden blocks. To help equate performance across conditions stimuli for the tactile-tactile and visual-tactile trials had fewer and slightly less similar features than the stimuli for the visual-visual trials (see Fig. 1).

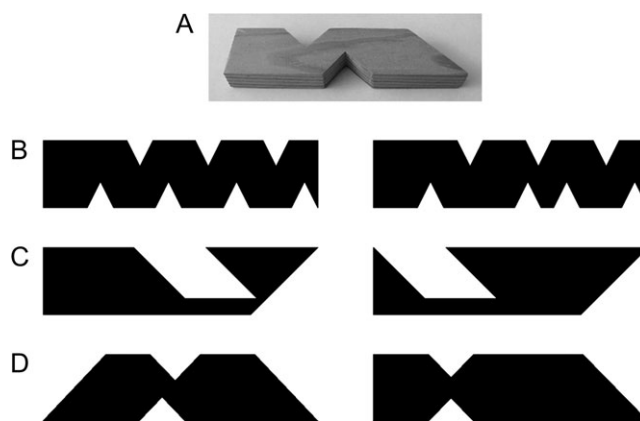
There were 72 visual shapes and 72 tactile shapes in total. Each stimulus was presented once per run, but the same stimuli were then repeated in subsequent runs (i.e., each stimulus was seen 4 times in total during scanning). To minimize visual and tactile repetition effects across scanning runs, subjects were presented with the full set of stimuli prior to scanning and instructed to practice the matching task. The order of the conditions and the order of the stimulus pairs within a condition in the prescan practice were different to the order in the scanner. Critically, all conditions (both congruent and incongruent trials for visual-tactile, visual-visual and tactile-tactile) were practiced in equal proportions; therefore, this prescan experience should not differentially influence 1 condition over another.

### fMRI Procedure

Tactile stimuli were mounted on card and presented on an angled magnetic resonance compatible table. The stimuli were placed under the subject's outstretched hands at the appropriate time by 2 experimenters coordinated by an auditory cue. During a block, 3 congruent and 3 incongruent trials of the same modality type (i.e., visual-visual, tactile-tactile, or 1 of the 4 cross-modal sets) were presented in a randomly intermixed order. Behavioral piloting found that at interstimulus intervals (ISIs) less than 4 s, accuracy on the tactile-tactile task was poor; therefore, we held the ISI at 4.23 s with 3.6 s of stimulus presentation and 0.63 s of fixation. Each block was followed by 16.2 s of fixation and then 3.78 s of visual instructions before the next block.

### Data Acquisition

Data were acquired on a Siemens 1.5-T scanner (Siemens, Erlangen, Germany). Functional images used a T2\*-weighted echo-planar imaging sequence for blood oxygen level-dependent (BOLD) contrast with  $3 \times 3$  mm in plane resolution, 2-mm slice thickness, and a 1-mm slice interval. Thirty-six slices were collected, resulting in an effective repetition time (TR) of 3.24 s. After the 4 functional runs, a T1-weighted anatomical



**Figure 1.** Example stimuli. (A) Photograph of a tactile stimulus; (B–D) examples of stimulus pairs from an incongruent visual-visual trial, an incongruent tactile-tactile trial, and an incongruent visual-tactile trial, respectively.

volume image was acquired from all subjects using a 3-dimensional (modified driven equilibrium Fourier transform) sequence and 176 sagittal partitions with an image matrix of  $256 \times 224$  and a final resolution of  $1 \text{ mm}^3$  (repetition time [TR]/echo time [TE]/inversion time [TI], 12.24/3.56/530 ms).

### Data Analysis

Functional data were analyzed with statistical parametric mapping (SPM2, Wellcome Department of Imaging Neuroscience, London, UK) implemented in Matlab 7.1 (Mathworks, Sherborne, MA). Preprocessing began by excluding the first 4 dummy scans to allow for T1 equilibration effects, realigning and unwarping the time series using the first volume as the reference scan (Andersson et al. 2001), spatially normalizing the data to the Montreal Neurological Institute (MNI)-152 template (Friston et al. 1995) and spatially smoothing using an 8-mm full width half maximum isotropic Gaussian kernel. One subject was removed from the analysis due to excess head movement.

First-level statistical analyses (single subject and fixed effects) modeled each trial type independently by convolving the stimulus onset times with a canonical hemodynamic response function (Glover 1999). We were able to distinguish congruent and incongruent trials using a constant ISI because their presentation order was randomized within block and the lengths of our ISI and TR differed such that the point of the hemodynamic response function at which the BOLD signal was sampled differed between trials (Veltman et al. 2002). However, we did not distinguish correct and incorrect responses in our analysis. The advantage of this approach is that we were able to match the number of trials for all possible variables (e.g., left vs. right hand/visual field, etc.). The disadvantage is that differences between conditions might be explained by differences in the number of correct/incorrect trials. However, in the present study, we can overcome this potential confound by showing that perirhinal activation was not correlated with correct or incorrect trials.

The data were high-pass filtered using a set of discrete cosine basis functions with a cut-off period of 128 s. Parameter estimates were calculated for all voxels using the general linear model. The parameter estimates were then fed into a second level random effects analysis implemented as analysis of variance (ANOVA) with 12 regressors (1 per condition summed over all 4 sessions) and the correction for nonsphericity implemented in SPM2. We computed linear contrasts to identify the main effects of Modality (cross-modal vs. unimodal) and Congruency (congruent vs. incongruent) and their interactions.

In addition, for each participant, we computed the effect of congruent visual-tactile trials relative to congruent unimodal trials across all sessions and for each session separately. Activation across all sessions was overlaid on the individuals' structural scans to confirm the presence of activation in the perirhinal cortex. Activation for each session separately was entered into a second group level analysis with 4 conditions (cross-modal congruent vs. unimodal congruent for each of the 4 sessions). We were then able to confirm that perirhinal activation was present in each session.

Activations were considered significant at  $P < 0.05$  using a voxelwise correction for multiple comparisons within the perirhinal region of interest. This was defined anatomically in both the left and right hemispheres of each participant according to the criteria proposed by Insausti et al. (1998). The perirhinal regions for the individual subjects were combined to produce an average for the group (see Devlin and Price 2007, for details of this procedure), which was used as the search space for the group effects of interest. In addition, an undirected whole brain search using a voxelwise correction for multiple comparisons identified significant activations outside of perirhinal cortex.

## Results

### Imaging Results

Preliminary investigation confirmed 1) the expected dissociation between visual-visual and tactile-tactile matching in occipital and somatosensory cortices, respectively; and 2) activation in bilateral occipito-temporal cortices for visual, tactile, and cross-modal

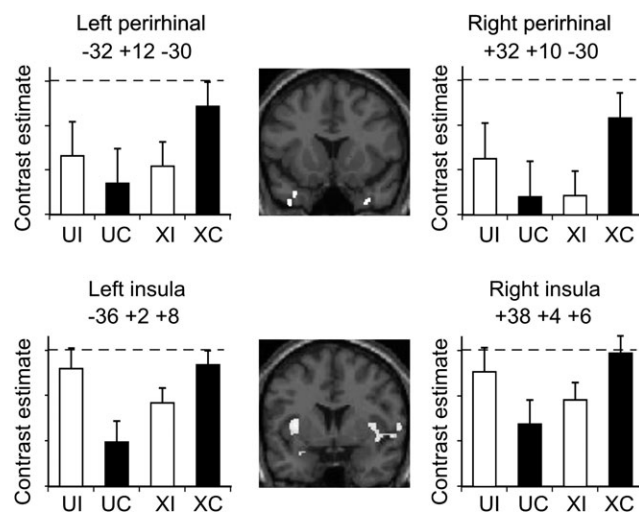
matching conditions (as predicted by Amedi et al. 2001), see Supplementary Material. Here we focus on the effect of cross-modal visual-tactile relative to unimodal matching in our perirhinal region of interest and the whole brain analysis.

### Perirhinal Region of Interest (ROI)

The functional imaging data were analyzed as a  $2 \times 2$  ANOVA. There were no significant main effects of modality (cross-modal vs. unimodal) or congruency (congruent vs. incongruent trials) within the perirhinal ROI. The critical comparison, however, was the interaction between congruency and modality and, as predicted, this comparison identified an area of significant activation located in perirhinal cortex bilaterally. Post hoc tests confirmed that perirhinal cortex activation was greater for congruent visual-tactile matching than either congruent and incongruent unimodal matching or incongruent visual-tactile matching (see Fig. 2 and Table 1 for details). There were no differences between congruent and incongruent trials for unimodal matching. In other words, consistent with our hypothesis, the interaction was driven by the condition with the greatest cross-modal integration demands, namely, congruent cross-modal matching.

To explore the anatomical localization of the perirhinal effects more precisely, we examined activation for congruent cross-modal trials relative to congruent unimodal matching in each individual subject within the preidentified subject-specific perirhinal region. These individual analyses clearly demonstrate activation peaks within the left and/or right perirhinal masks in all subjects (see Table 2). These individual effects are illustrated for 10 subjects in Figure 3.

Finally, we investigated whether perirhinal activation varied across the 4 different scanning sessions. The rationale here was that if perirhinal activation reflects the encoding of the stimulus into long-term memory rather than the integration of information (i.e., a perceptual process) then it might decrease



**Figure 2.** Perirhinal and insula activations from the group analysis. Central panels show coronal slices of group perirhinal ( $y = +14$ ) and insula ( $y = +4$ ) activations for cross-modal versus unimodal matching. The adjacent plots show estimates of the effect size for perirhinal and insular cortex activations in both left and right hemispheres for unimodal incongruent (UI), unimodal congruent (UC), cross-modal incongruent (XI), and cross-modal congruent (XC) conditions, with the ceiling condition indicated by the dashed line. The comparison of the cross-modal congruent condition with each of the other 3 conditions (XI, UC, and UI) was significant at  $P < 0.05$  in both the left and right perirhinal cortices.

**Table 1**

MNI coordinates and Z scores for the interaction between modality and congruency in the group data, and for post hoc tests of cross-modal versus unimodal matching for congruent trials only and congruent versus incongruent trials for cross-modal matching only

Interaction (cross-modal > unimodal) for (congruent > incongruent)				Congruent only visual-tactile > visual-visual and tactile-tactile				Visual-tactile only congruent > incongruent			
Co-ordinates			Z score	Coordinates			Z score	Coordinates			Z score
x	y	z		x	y	z		x	y	z	
Perirhinal mask											
-34	16	-32	(3.2)	-34	14	-32	(3.8)	-34	14	-32	(2.8)
				-32	12	-30	(2.9)	-32	12	-30	(2.7)
-30	8	-20	(4.1)	-32	8	-20	(3.3)	-30	12	-20	(2.5)
-38	16	-40	(3.6)	-38	14	-40	(2.9)	-38	14	-40	(3.3)
32	8	-28	(3.3)	32	10	-30	(2.9)	32	10	-30	(3.5)
38	12	-42	(3.1)	34	14	-38	(2.8)	32	12	-46	(4.2)
Insula											
-36	8	4	(5.0)	-36	2	8	(4.8)	-36	6	4	(3.4)
38	4	6	(3.8)	38	4	6	(3.8)	36	6	8	(3.7)

Note: Negative x = left hemisphere.

**Table 2**

MNI coordinates and Z scores for the comparison of cross-modal versus unimodal matching for congruent trials only, within the perirhinal cortex for each subject (labeled 1-17)

	Left perirhinal, y: -8 to +8				Right perirhinal, y: -8 to +8			
	x	y	z	Z score	x	y	z	Z score
1	-26	6	-34	2.2	38	-6	-38	1.8
2	-34	2	-44	2.7	<b>32</b>	<b>-2</b>	<b>-40</b>	<b>3.1</b>
	-36	-2	-32	2.0	36	-8	-30	2.6
3	<b>-24</b>	<b>8</b>	<b>-44</b>	<b>3.1</b>	30	0	-44	2.5
4	<b>-30</b>	<b>-8</b>	<b>-38</b>	<b>4.9</b>	<b>32</b>	<b>-6</b>	<b>-46</b>	<b>4.2</b>
	<b>-36</b>	<b>-6</b>	<b>-48</b>	<b>3.4</b>	<b>34</b>	<b>-2</b>	<b>-44</b>	<b>3.9</b>
5	<b>-30</b>	<b>-6</b>	<b>-34</b>	<b>3.1</b>	20	-4	-46	1.9
	-30	4	-38	2.4				
6	-26	2	-36	1.8	22	2	-38	2.2
7	<b>-32</b>	<b>2</b>	<b>-42</b>	<b>3.2</b>	26	-2	-46	2.4
8	-30	2	-38	2.3				
9	-26	8	-30	2.7	<b>28</b>	<b>-8</b>	<b>-42</b>	<b>3.4</b>
	-34	8	-46	2.2				
10	<b>-36</b>	<b>-6</b>	<b>-38</b>	<b>3.0</b>	20	2	-40	2.2
	-34	8	-22	2.3	30	2	-34	2.1
11	-28	-4	-46	<b>3.1</b>	30	-6	-48	2.3
	-28	4	-32	2.1				
12	<b>-32</b>	<b>-8</b>	<b>-30</b>	<b>3.3</b>	38	0	-46	1.7
	-26	6	-32	2.7				
13					30	-8	-46	1.7
14	-20	2	-36	2.2				
	-24	8	-34	2.1				
15	-24	-4	-36	2.8	<b>24</b>	<b>6</b>	<b>-26</b>	<b>3.4</b>
	-32	-4	-34	1.9				
16	-34	2	-42	2.0				
	-28	-2	-32	1.9				
17	-30	2	-36	1.8	38	-2	-42	1.7

Note: Bold =  $P < 0.001$  uncorrected.

across time as the subjects are reexposed to the same stimuli (Lee et al. 2008). However, this is not what we observed. Perirhinal activation was observed ( $P < 0.05$ ) for the effect of cross-modal congruent versus unimodal congruent in each session with a weak trend for more activation in Sessions 1 and 2 relative to Sessions 3 and 4 in the left perirhinal cortex ( $x = -28$ ,  $y = +14$ ;  $z = -26$ ; Z score = 2.2).

### Whole Brain Analysis

For the comparison of interest (the interaction between Congruency and Modality), there was 1 significant effect following correction for multiple comparisons across the whole brain. This was located in a region of the left insula (see Table 1 and Fig. 2) with a corresponding, but less

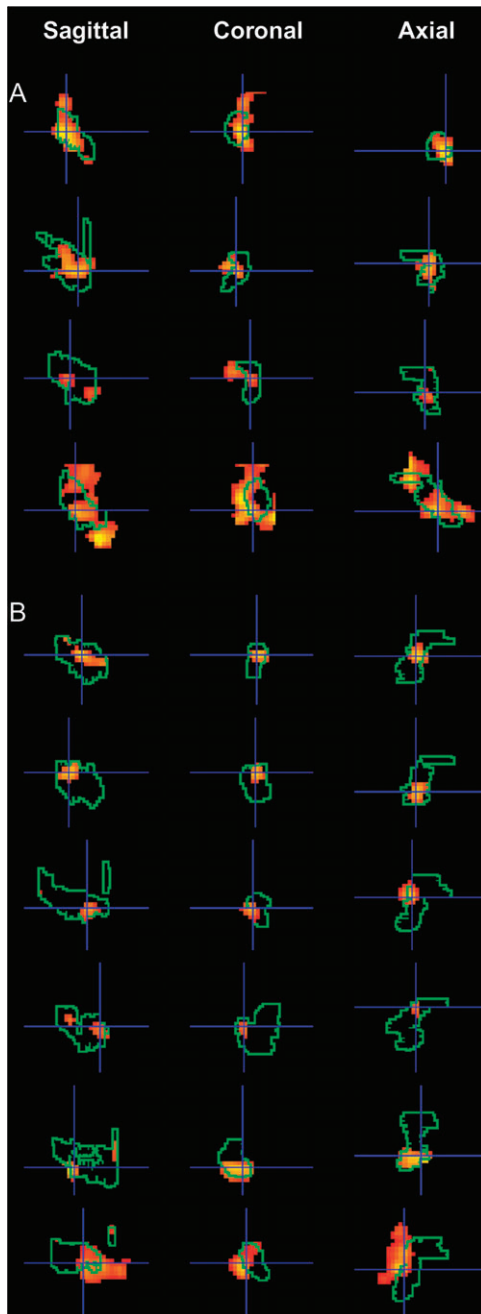
significant, effect in the right insula (see Table 1). Examination of the extent of these activations confirmed that they were located in the posterior dysgranular area in both the left and right insula. Although our activation bordered the claustrum where cross-modal processing has been identified in macaques (Hörster et al. 1989), it did not overlap the claustrum, and therefore, we think it is more likely to stem from the insula. The effect sizes per condition in the left and right insula regions are shown in Figure 2. Unlike the perirhinal cortex, which showed greater activation for congruent cross-modal trials than all other conditions, the activation within the insula was not significantly different during congruent cross-modal matching than incongruent unimodal matching. In short, the response profile for these 2 regions was not the same, despite both showing a significant interaction.

With respect to other effects, there was a main effect of Modality, with cross-modal relative to unimodal trials showing significant activation in a posterior part of the left insula ( $-38$   $-6$   $14$ ,  $Z = 4.7$ , with 135 voxels at  $P < 0.001$ ). The opposite contrast did not reveal any significant activations for unimodal relative to cross-modal activation. There was also a main effect of Congruency, with greater activation for incongruent than congruent trials located in the right posterior middle frontal gyrus ( $48$ ,  $14$ ,  $36$ ;  $Z = 4.7$ , with 101 voxels at  $P < 0.001$ ). No region showed the opposite effect (congruent > incongruent).

### Behavioral Results

A  $2 \times 2$  repeated measures ANOVA was performed on the accuracy data with Modality (cross-modal vs. unimodal) and Congruency (congruent vs. incongruent trials) as independent factors. There was a main effect of Modality ( $F(1,16) = 9.907$ ,  $P < 0.05$ ) with higher accuracy in the cross-modal than unimodal conditions (90.8% vs. 85.9%, see Table 3). This was driven by the fact that accuracy was lowest for tactile-tactile trials. In addition, there was a main effect of Congruency ( $F(1,16) = 7.761$ ,  $P < 0.05$ ) with significantly greater accuracy for congruent (92.4%) than incongruent trials (85.9%). However, there was no interaction between modality and congruency ( $F(1,16) = 2.589$ ,  $P > 0.05$ ); therefore, the perirhinal activation pattern could not be explained by accuracy differences.

The above analysis was then extended to include scanning run (1 to 4) as a third factor. A main effect of scanning run ( $F(1.75, 28.2) = 11.6$ ,  $P < 0.05$ ) indicated poorer performance



**Figure 3.** Perirhinal activations from the individual subject analysis. Subject-specific perirhinal activation for congruent visual-tactile trials relative to congruent unimodal trials superimposed onto subject-specific perirhinal masks in the sagittal, coronal, and axial planes. Each row shows activation in 1 hemisphere for 1 subject. Panel (A) shows activation of the right perirhinal cortex for subjects numbered 1–4 in Table 2. Panel (B) shows activation in the left perirhinal cortex for subjects numbered 5–10 in Table 2.

in the first than subsequent runs (82.6%, 88.5%, 91.2%, and 90.4%) suggesting that subjects became familiar with the stimuli over scanning runs. However, the absence of a significant interaction between run and modality ( $F(2.342, 37.47) = 0.755, P > 0.05$ ), run and congruency ( $F(1.894, 30.298) = 0.333, P > 0.05$ ), or run, modality and congruency ( $F(2.281, 36.489) = 1.158, P > 0.05$ ), (all  $P$  values Greenhouse-Geisser corrected for nonsphericity) provided no evidence that the effect of learning depended on the type of trial.

**Table 3**

Behavioral data

Condition	Accuracy		
	Total	Congruent	Incongruent
Unimodal matching			
Tactile-tactile	82.0 (10.7)	79.7 (11.1)	83.8 (16.7)
Visual-visual	89.7 (8.8)	96.1 (5.4)	84.1 (13.8)
Mean	85.9 (7.4)	87.9 (5.9)	84 (11.5)
Cross-modal matching			
Visual-tactile left/left	94.4 (5.1)	96.1 (7.6)	92.7 (6.8)
Visual-tactile left/right	89.0 (7.4)	95.1 (6.1)	82.6 (12.7)
Visual-tactile right/right	92.2 (7.5)	94.8 (5.4)	89.5 (12.2)
Visual-tactile right/left	87.9 (7.9)	92.8 (7.0)	83.0 (12.2)
Mean	90.8 (6.3)	94.7 (4.7)	86.9 (9.8)

Note: Mean ( $\pm$ SD) percent accuracy of 17 subjects, averaged over scanning runs.

Unfortunately, we were unable to collect reaction times during the experiment because we did not have the equipment to link a foot pedal to a timing device.

### Discussion

In this study, we show that perirhinal activation was higher during cross-modal than unimodal matching of visual and tactile information, but only when the inputs represented a single object. These findings support our hypothesis that the human perirhinal cortex plays a role in the successful integration of cross-modal perceptual information into an abstract object-level representation, and, thus, indicate a functional homology between the human and monkey perirhinal cortex. In addition, our observation that perirhinal activation was enhanced by congruent cross-modal trials explains why previous imaging studies of visuo-tactile matching in humans, which have combined activation over congruent and incongruent trials, have not reported activation in the perirhinal cortex.

The design and construction of our experiment allow us to eliminate a number of alternative explanations. The specificity of the response to congruent cross-modal rather than unimodal feature matching, despite controlling for task difficulty, suggests that the perirhinal activation we observed cannot be accounted for by a successful match per se. Nor can it be explained by comparisons of visible object representations with internal representations constructed from tactile information because such explanations would predict similar levels of activation on both incongruent and congruent cross-modal trials. It is also unlikely that our perirhinal activation can be explained by longer perceptual processing time in the congruent visual-tactile condition because response time was longest for tactile-tactile matching (see Experimental Procedures). Finally, it is also unlikely that our results can be explained in terms of semantic memory or verbal strategies because the stimuli used in our study had not been encountered by subjects prior to the experiment and had no meaning or verbal labels.

Returning to the functional homology between the human and monkey perirhinal cortex, our finding that perirhinal activation was greatest for successful (congruent) cross-modal matching is consistent with the view, derived from nonhuman primate studies, that the perirhinal cortex is involved in the integration of polymodal information into abstract object-level representations (Murray and Bussey 1999; Murray and Richmond 2001; Bussey and Saksida 2005; Buckley and Gaffan 2006). Thus, we are suggesting a comparable role in human and nonhuman

primates. Specifically, we are arguing that perirhinal activation plays a role at the level of object-level representations during congruent cross-modal matching. This cross-modal perspective compliments other work that has argued for a role in whole object processing in the context of manipulating the number of relevant features within the same modality (Bussey et al. 2002; Bussey and Saksida 2005). Object-level representations were not expected to be involved in unimodal matching in our study because it could be based on individual visual or tactile feature matching mediated by sensory-specific cortices (Bussey et al. 2002, 2003; Hampton and Murray 2002; Hampton 2005).

Others have also argued that the human perirhinal cortex is involved in cross-modal integration (Taylor et al. 2006); however, the findings on which these conclusions were based differ in 2 critical ways to ours. First, activation associated with tactile-visual perceptual matching in our study fell clearly within the perirhinal cortex, whereas the activation reported by Taylor et al., in response to auditory-visual conceptual matching, was more posterior and fell just behind the perirhinal cortex (according to the criteria of Insausti et al. 1998). Second, activation of the perirhinal cortex in our study was higher for congruent than incongruent tactile-visual matching, that is, higher when cross-modal inputs could be successfully integrated, whereas the more posterior activation reported by Taylor et al. was higher for incongruent than congruent auditory-visual conceptual matching. Therefore, our findings clearly demonstrate perirhinal activation that is associated with successful integration of cross-modal inputs into object-level representations. It remains to be determined why the findings of these studies differ, but the explanation may relate to differences in the nature of the task (perceptual vs. conceptual matching) and/or the nature of the stimuli (meaningless vs. meaningful; tactile and visual vs. auditory and visual).

Object-level representations may contribute to both declarative memory (Murray and Bussey 1999; Murray and Richmond 2001; Bussey et al. 2005; Buckley and Gaffan 2006) and object perception (Buckley and Gaffan 1997, 1998; Buckley et al. 2001; Bussey et al. 2002, 2003; Tyler et al. 2004; Buckley 2005; Lee, Barenese, and Graham, 2005; Lee, Buckley, et al. 2005, 2006; Lee, Bussey, et al. 2005; Lee, Bandelow, et al. 2006; Devlin and Price 2007; but see Buffalo et al. 1998; Holdstock et al. 2000; Stark and Squire 2000; Hampton 2005; Levy et al. 2005). In our study, we attempted to minimize the demands on learning, working and declarative memory by using simultaneous stimulus presentation with no delay between the decision process and the subject's response. Nevertheless, successful performance on all our matching conditions, be they cross-modal or unimodal, would have involved working memory and perhaps also incidental encoding into long-term memory. Indeed, a learning effect is apparent in the behavioral data that indicated greater accuracy during sessions 2-4 relative to session 1. However, the critical point is that this learning effect in the behavioral data was observed across all conditions. Learning was therefore not unique to the cross-modal congruent condition although we cannot exclude the possibility that perirhinal activation may reflect memory processes that are more engaged by cross-modal congruent trials than any other condition.

In summary, our data demonstrate that the perirhinal cortex is involved in cross-modal perceptual matching when the demands on learning, working and declarative memory were minimized. This is consistent with previous findings that this

region may contribute to object perception (Tyler et al. 2004; Lee, Buckley, et al. 2005; Lee, Bussey, et al. 2005; Lee, Bandelow, et al. 2006; Lee, Buckley, et al. 2006; Devlin and Price 2007) as well as memory (Buffalo et al. 1998; Davachi et al. 2003; Henson et al. 2003; Ranganath et al. 2003; O'Kane et al. 2005; Buffalo et al. 2006; Elfgren et al. 2006; Gold et al. 2006; Montaldi et al. 2006; Staresina and Davachi 2006). Future studies are now required to investigate whether the regions of human perirhinal cortex that we associate with cross-modal perceptual integration are anatomically distinct from those that activate during other perceptual, memory or learning tasks.

### **Whole Brain Analysis**

Although our study focused specifically on the perirhinal cortex, it is likely that a number of regions work together to integrate and represent multi-modal information about objects. Previous studies of visual-tactile matching have reported activation of the right or left insula/claustrum (Hadjikhani and Roland 1998; Banati et al. 2000), the anterior intraparietal sulcus (Grefkes et al. 2002) and the posterior intraparietal sulcus (Saito et al. 2003) during visual-tactile relative to unimodal matching. It has also been suggested that an area of the Lateral Occipital Complex (LOC or LOtv) is involved in visual-tactile integration because it is activated by recognition of real objects and perceptual processing of meaningless objects when they are presented both in the visual and tactile modality (Amedi et al. 2001; James et al. 2002). Likewise, the posterior inferior temporal gyrus close to region LOtv shows the same category-related patterns of response to manmade objects and faces when these are presented in the visual and tactile modalities (Pietrini et al. 2004).

In our study, the only 1 of these regions that was activated significantly more by visual-tactile than unimodal matching was the left insula, consistent with the findings of Banati et al. (2000). Here, we also report a new finding concerning the insula, which is that the effect of modality (cross-modal vs. unimodal) was modified by stimulus congruency (i.e., there is a significant interaction between modality and congruency). This has not been shown previously. It has been argued that the insula acts as a mediating region that enables communication and exchange of information between unimodal regions (Amedi et al. 2005). Our finding that the insula was activated more by congruent than incongruent visual-tactile stimulus pairs suggests either that greater exchange of information between unimodal regions is required on congruent trials or that the insula is involved in successful cross-modal integration itself rather than merely facilitating communication between unimodal regions. The fact that the insula was also activated by incongruent unimodal trials indicates that the insula, unlike the perirhinal cortex, was not exclusively responsive to cross-modal information in our study, and is more consistent with an involvement of this region in facilitating communication between other brain regions rather than a specific involvement in cross-modal integration.

Also consistent with previous studies, we found bilateral activation common to visual-tactile matching, visual-visual matching and tactile-tactile matching in a region close to the activations in LOtv reported by Amedi et al. (2001) and Pietrini et al. (2004) during visual and tactile object recognition (see Supplementary Material). Our data therefore support a role for this region in processing both visual and tactile information

about objects, show that it is activated during cross-modal as well as unimodal matching, and confirm the findings of James et al. (2002) that this region can be activated by meaningless as well as real objects. Further work is now needed to determine how the perirhinal cortex, LOTv, insula, and parietal cortex interact during visual-tactile processing, what the relative contributions of these regions are, and whether there are other, as yet unidentified, regions, particularly in the temporal lobe, that contribute to visual-tactile integration of object information. The contribution of the current study is to show that although posterior regions such as LOTv are involved in both cross-modal and unimodal visual and tactile processing, it is the perirhinal cortex and the insula that show enhanced activation for the successful integration of visual and tactile information.

## Conclusions

We conclude that, as predicted on the basis of anatomical connectivity and lesion studies in nonhuman primates, the human perirhinal cortex is involved in cross-modal perceptual matching of congruent stimuli, and we suggest that this is because it plays a role in the successful integration of visual and tactile information into an abstract multi-modal object representation.

## Supplementary Material

Supplementary material can be found at: <http://www.cercor.oxfordjournals.org/>

## Funding

The Wellcome Trust (072282/Z/03/Z to C.P.).

## Notes

We would like to thank Essie Li for her help with the behavioral piloting of the tasks and Keith May for his help with figure preparation.

*Conflict of Interest:* None declared.

Address correspondence to email: [juliet@julietholdstock.com](mailto:juliet@julietholdstock.com).

## References

- Amedi A, Malach R, Hendler T, Peled S, Zohary E. 2001. Visuo-haptic object-related activation in the ventral visual pathway. *Nat Neurosci.* 4:324-330.
- Amedi A, von Kriegstein K, van Atteveldt NM, Beauchamp MS, Naumer MJ. 2005. Functional imaging of human crossmodal identification and object recognition. *Exp Brain Res.* 166:559-571.
- Andersson JL, Hutton C, Ashburner J, Turner R, Friston K. 2001. Modeling geometric deformations in EPI time series. *Neuroimage.* 13:903-919.
- 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.
- Buckley MJ. 2005. The role of the perirhinal cortex and hippocampus in learning, memory, and perception. *Quart J Exp Psychol.* 58B:246-268.
- Buckley MJ, Booth MCA, Rolls ET, Gaffan D. 2001. Selective perceptual impairments after perirhinal cortex ablation. *J Neurosci.* 21:9824-9836.
- Buckley MJ, Gaffan D. 1997. Impairment of visual object-discrimination learning after perirhinal cortex ablation. *Behav Neurosci.* 111:467-475.
- Buckley MJ, Gaffan D. 1998. Perirhinal cortex ablation impairs visual object identification. *J Neurosci.* 18:2268-2275.
- Buckley MJ, Gaffan D. 2006. Perirhinal cortical contributions to object perception. *Trends Cogn Sci.* 10:100-107.
- Buffalo EA, Bellgowan PSF, Martin A. 2006. Distinct roles for medial temporal lobe structures in memory for objects and their locations. *Learn Mem.* 13:638-643.
- Buffalo EA, Reber PJ, Squire LR. 1998. The human perirhinal cortex and recognition memory. *Hippocampus.* 8:330-339.
- Bussey TJ, Saksida LM. 2005. Object memory and perception in the medial temporal lobe: an alternative approach. *Curr Opin Neurobiol.* 15:730-737.
- Bussey TJ, Saksida LM, Murray EA. 2002. Perirhinal cortex resolves feature ambiguity in complex visual discriminations. *Eur J Neurosci.* 15:365-374.
- Bussey TJ, Saksida LM, Murray EA. 2003. Impairments in visual discrimination after perirhinal cortex lesions: testing 'declarative' vs. 'perceptual-mnemonic' views of perirhinal cortex function. *Eur J Neurosci.* 17:649-660.
- Bussey TJ, Saksida LM, Murray EA. 2005. The perceptual-mnemonic/feature conjunction model of perirhinal cortex function. *Quart J Exp Psychol.* 58B:269-282.
- Carmichael ST, Price JL. 1995. Limbic connections of the orbital and medial prefrontal cortex in macaque monkeys. *J Comp Neurol.* 363:615-641.
- Davachi L, Mitchell JP, Wagner AD. 2003. Multiple routes to memory: distinct medial temporal lobe processes build item and source memories. *Proc Natl Acad Sci USA.* 100:2157-2162.
- Devlin JT, Price CJ. 2007. Perirhinal contributions to human visual perception. *Curr Biol.* 17:1484-1488.
- Elfgrén C, van Westen D, Passant U, Larsson E-M, Mannfolk P, Fransson P. 2006. fMRI activity in the medial temporal lobe during famous face processing. *Neuroimage.* 30:609-616.
- Friedman DP, Murray EA, O'Neill JB, Mishkin M. 1986. Cortical connections of the somatosensory fields of the lateral sulcus of macaques: evidence for a corticolimbic pathway for touch. *J Comp Neurol.* 252:323-347.
- Friston KJ, Ashburner J, Frith CD, Poline J-B, Heather JD, Frackowiak RSJ. 1995. Spatial registration and normalization of images. *Hum Brain Mapp.* 2:165-189.
- Glover GH. 1999. Deconvolution of impulse response in event-related BOLD fMRI. *Neuroimage.* 9:416-429.
- Gold JJ, Smith CN, Bayley PJ, Shrager Y, Brewer JB, Stark CEL, Hopkins RO, Squire LR. 2006. Item memory, source memory, and the medial temporal lobe: concordant findings from fMRI and memory-impaired patients. *Proc Natl Acad Sci USA.* 103:9351-9356.
- Goulet S, Murray EA. 2001. Neural substrates of crossmodal association memory in monkeys: the amygdala versus the anterior rhinal cortex. *Behav Neurosci.* 115:271-284.
- Grefkes C, Weiss PH, Zilles K, Fink GR. 2002. Crossmodal processing of object features in human anterior intraparietal cortex: an fMRI study implies equivalencies between humans and monkeys. *Neuron.* 35:173-184.
- Hadjikhani N, Roland PE. 1998. Cross-modal transfer of information between tactile and the visual representations in the human brain: a positron emission tomographic study. *J Neurosci.* 18:1072-1084.
- Hampton RR. 2005. Monkey perirhinal cortex is critical for visual memory, but not for visual perception: reexamination of the behavioural evidence from monkeys. *Quart J Exp Psychol.* 58B:283-299.
- Hampton RR, Murray EA. 2002. Learning of discriminations is impaired, but generalization to altered views is intact, in monkeys (*macaca mulatta*) with perirhinal cortex removal. *Behav Neurosci.* 116:363-377.
- Henson RNA, Cansino S, Herron JE, Robb WGK, Rugg MD. 2003. A familiarity signal in human anterior medial temporal cortex? *Hippocampus.* 13:301-304.
- Hocking J, Price CJ. 2008. The role of the posterior superior temporal sulcus in audiovisual processing. *Cereb Cortex.* 18:2439-2449.
- Holdstock JS, Guitnikov SA, Gaffan D, Mayes AR. 2000. Perceptual and mnemonic matching-to-sample in humans: contributions of the hippocampus, perirhinal cortex and other medial temporal lobe cortices. *Cortex.* 36:301-322.
- Hörster W, Rivers A, Schuster B, Ettliger G, Skreczek W, Hesse W. 1989. The neural structures involved in cross-modal recognition and

- tactile discrimination performance: an investigation using 2-DG. *Behav Brain Res.* 33:209-227.
- Insausti R, Juottonen K, Soininen H, Insausti AM, Partanen K, Vainio P, Laakso MP, Pitkanen A. 1998. MR volumetric analysis of the human entorhinal, perirhinal, and temporopolar cortices. *Am J Neuroradiol.* 19:659-671.
- James TW, Humphrey GK, Gati JS, Servos P, Menon RS, Goodale MA. 2002. Haptic study of three-dimensional objects activates extrastriate visual areas. *Neuropsychologia.* 40:1706-1714.
- Lee ACH, Bandelow S, Schwarzbauer C, Henson RNA, Graham KS. 2006. Perirhinal cortex activity during visual object discrimination: an event-related fMRI study. *NeuroImage.* 33:362-373.
- Lee ACH, Barense MD, Graham KS. 2005. The contribution of the human medial temporal lobe to perception: bridging the gap between animal and human studies. *Quart J Exp Psychol.* 58B:300-325.
- Lee ACH, Buckley MJ, Gaffan D, Emery T, Hodges JR, Graham KS. 2006. Differentiating the roles of the Hippocampus and perirhinal cortex in processes beyond long-term declarative memory: a double dissociation in dementia. *J Neurosci.* 26:5198-5203.
- Lee ACH, Buckley MJ, Pegman SJ, Spiers H, Scahill VL, Gaffan D, Bussey TJ, Rhys Davies R, Kapur N, Hodges JR, et al. 2005. Specialization in the medial temporal lobe for processing of objects and scenes. *Hippocampus.* 15:782-797.
- Lee ACH, Bussey TJ, Murray EA, Saksida LM, Epstein RA, Kapur N, Hodges JR, Graham KS. 2005. Perceptual deficits in amnesia: challenging the medial temporal lobe 'mnemonic' view. *Neuropsychologia.* 43:1-11.
- Lee ACH, Scahill VL, Graham KS. 2008. Activating the medial temporal lobe during oddity judgment for faces and scenes. *Cereb Cortex.* 18:683-696.
- Levy DA, Shrager Y, Squire LR. 2005. Intact visual discrimination of complex and feature-ambiguous stimuli in the absence of perirhinal cortex. *Learn Mem.* 12:61-66.
- Montaldi D, Spencer TJ, Roberts N, Mayes AR. 2006. The neural system that mediates familiarity memory. *Hippocampus.* 16:504-520.
- Murray EA, Bussey TJ. 1999. Perceptual-mnemonic functions of the perirhinal cortex. *Trends Cogn Sci.* 3:142-151.
- Murray EA, Richmond BJ. 2001. Role of perirhinal cortex in object perception, memory, and associations. *Curr Opin Neurobiol.* 11:188-193.
- Nahm FKD, Tranel D, Damasio H, Damasio AR. 1993. Cross-modal associations and the human amygdala. *Neuropsychologia.* 8:727-744.
- O'Kane G, Insler RZ, Wagner AD. 2005. Conceptual and perceptual novelty effects in human medial temporal cortex. *Hippocampus.* 15:326-332.
- Parker A, Gaffan D. 1998. Lesions of the primate rhinal cortex cause deficits in flavour-visual associative memory. *Behav Brain Res.* 93:99-105.
- Pietrini P, Furey ML, Ricciardi E, Gobbi MI, Wu W-HC, Cohen L, Guazzelli M, Haxby JV. 2004. Beyond sensory images: object-based representation in the human ventral pathway. *Proc Natl Acad Sci USA.* 101:5658-5663.
- Ranganath C, Yonelinas AP, Cohen M, Dy CJ, Tomn SM, D'Esposito M. 2003. Dissociable correlates of recollection and familiarity within the medial temporal lobes. *Neuropsychologia.* 42:2-13.
- Saito DN, Okado T, Morita Y, Yonekura Y, Sadato N. 2003. Tactile-visual cross-modal shape matching: a functional MRI study. *Cogn Brain Res.* 17:14-25.
- Shaw C, Kentridge RW, Aggleton JP. 1990. Cross-modal matching by amnesic patients. *Neuropsychologia.* 28:665-671.
- Stark CE, Squire LR. 2000. Intact visual perceptual discrimination in humans in the absence of perirhinal cortex. *Learn Mem.* 7:273-278.
- Staresina BP, Davachi L. 2006. Differential encoding mechanisms for subsequent associative recognition and free recall. *J Neurosci.* 26:9162-9172.
- Suzuki WA. 1996a. The anatomy, physiology and functions of the perirhinal cortex. *Curr Opin Neurobiol.* 6:179-186.
- Suzuki WA. 1996b. Neuroanatomy of the monkey entorhinal, perirhinal and parahippocampal cortices: organization of cortical inputs and interconnections with amygdala and striatum. *Semin Neurosci.* 8:3-12.
- Suzuki WA, Amaral DG. 1994a. Perirhinal and parahippocampal cortices of the macaque monkey: cortical afferents. *J Comp Neurol.* 350:497-533.
- Suzuki WA, Amaral DG. 1994b. Topographic organization of the reciprocal connections between monkey entorhinal cortex and the perirhinal and parahippocampal cortices. *J Neurosci.* 14:1856-1877.
- Taylor KI, Moss HE, Stamatakis EA, Tyler LK. 2006. Binding crossmodal object features in perirhinal cortex. *Proc Natl Acad Sci USA.* 103:8239-8244.
- Tyler LK, Stamatakis EA, Bright P, Acres K, Abdallah S, Rodd JM, Moss HE. 2004. Processing objects at different levels of specificity. *J Cogn Neurosci.* 16:351-362.
- Van Hoesen GW, Pandya DN. 1975. Some connections of the entorhinal (area 28) and perirhinal (area 35) cortices of the rhesus monkey. 1. Temporal lobe afferents. *Brain Res.* 95:1-24.
- Veltman DJ, Mechelli A, Friston KJ, Price CJ. 2002. The importance of distributed sampling in blocked functional magnetic imaging designs. *NeuroImage.* 17:1203-1206.