



## Supplement

# Action relationships concatenate representations of separate objects in the ventral visual system

Katherine L. Roberts\*, Glyn W. Humphreys

*Brain and Behavioural Sciences, School of Psychology, University of Birmingham, B15 2TT*

## ARTICLE INFO

## Article history:

Received 5 January 2010

Revised 10 May 2010

Accepted 14 May 2010

Available online 24 May 2010

## ABSTRACT

Objects in the real world are encountered in contexts where they interact together. Though it is known that neurons in the ventral visual stream mediate the recognition of individual objects, we have minimal knowledge of how multiple objects are processed at a neural level. We examined the neural response to pairs of objects using functional magnetic resonance imaging (fMRI). Objects positioned to interact together activated bilateral lateral occipital complex (LOC) and fusiform gyrus. This occurred irrespective of whether the objects were attended. In LOC, the effect of positioning objects for action was found regardless of whether the objects formed a familiar or unfamiliar action pair. In the fusiform gyrus activation was found when objects formed a familiar action pair. No changes were apparent in visuomotor (premotor and parietal) regions which might reflect a motor-based response to objects. These results show that ventral-stream regions respond to the interaction between objects, as well as to the sensory and functional properties of individual objects.

© 2010 Elsevier Inc. All rights reserved.

## Introduction

There is overwhelming evidence that neurons in the ventral visual stream are necessary for visual object recognition. Damage to ventral visual regions is associated with modality-specific deficits in visual object recognition (Humphreys and Riddoch, 2006). In normal observers, viewing objects, relative to textures or contours, leads to enhanced activation in the lateral occipital complex within ventral occipitotemporal cortex (Grill-Spector, 2003; Kourtzi and Kanwisher, 2001; Spiridon et al., 2006). A more medial ventral region, the fusiform gyrus, is activated when visual elements organise into shape (Op de Beeck et al., 2000), while also showing some selectivity in responding to particular objects (e.g., faces and words (Kanwisher et al., 1997; Nobre et al., 1994)). Interestingly the presence of multiple objects very often reduces neural activity in object selective regions (Jacques and Rossion, 2004; Kastner et al., 2001; Rolls and Tovee, 1995). This is consistent with stimuli competing to control the neural response to individual objects (Desimone and Duncan, 1995).

Although the vast majority of studies emphasise the recognition of individual objects, objects encountered in the real world very often appear within a context with other objects. At a behavioural level, object recognition improves when objects are coherently rather than randomly positioned within a scene (Biederman, 1972). There is also evidence for benefits in processing when objects are positioned to interact together. In the neuropsychological phenomenon of extinc-

tion patients are able to report a single stimulus appearing in the field contralateral to their lesion, but they can fail to detect the same object when another item appears simultaneously on the ipsilesional side. This spatial bias to select the ipsilesional stimulus is reduced when patients are presented with objects that are co-located for a common action (e.g., wine bottle pouring into a glass). Extinction returns, though, if the same objects are not co-located for action (Riddoch et al., 2003): in this case the identities of the objects are less critical than their being positioned together for a common action. Indeed, extinction reduces with objects that are unfamiliar as a pair but which still interact together (e.g., wine bottle pouring into a bucket) (Riddoch et al., 2006) and when unrelated objects have implied motion from one to the other (Riddoch et al., 2010). Recovery from extinction is found when objects are integrated via Gestalt grouping cues such as collinearity or common contrast polarity (Gilchrist et al., 1996). Riddoch et al.'s (2003) data indicate that interacting objects are similarly assimilated as a single perceptual unit, reducing spatial biases in selection after brain lesion. Interestingly, effects of the action relation between objects can also occur on trials when patients report only one of the two objects present. When objects are interacting, patients show preferential report of the object that is 'active' in the action (e.g., the wine bottle rather than the glass) irrespective of its spatial position. When objects do not interact, the bias to report the ipsilesional item returns. This suggests that the action relationship between the objects is coded pre-attentively, and influences which of the two objects is selected first.

The neural basis of this response to interacting objects is unknown. To test the response, we used fMRI to measure the blood oxygen level

\* Corresponding author. Fax: +44 0121 414 4897.

E-mail address: [k.l.roberts.1@bham.ac.uk](mailto:k.l.roberts.1@bham.ac.uk) (K.L. Roberts).

dependent (BOLD) response to objects that were either positioned correctly or incorrectly for action. To maximise the likelihood of objects being coded as a single perceptual unit, we used stimuli that co-occurred frequently as a pair (wine bottle and wine glass, screwdriver and screw), which were either positioned correctly (maximal integration) or not positioned correctly for action. We also repaired the members of these pairings to create stimuli where individual members could be positioned correctly or incorrectly for action (e.g., wine bottle and screw), but which should not form such a strong perceptual unit (see [Appendix 1](#) for a list of all the items and the methods section for ratings of the familiarity and positioning of the object pairs). These less plausible pairs test for effects of positioning individual objects for action when the familiarity of the object pairing is weak. Critically, comparisons of interest will contrast activation when the same pairs of objects are positioned correctly vs. when they are positioned incorrectly for action, allowing us to investigate effects of positioning independent of effects of abstract, semantic relatedness and/or the visual familiarity of the pairing. We also manipulated whether the objects were explicitly attended or not, to determine if objects must be attended for their action properties to be processed.

We ask whether regions known to mediate the recognition of single objects (LOC, the fusiform gyrus) also show a differential response to interacting compared with non-interacting objects, and whether they are sensitive to the strength of the action relationship between the objects (jug and glass, compared with jug and nail). We also examine whether effects of inter-object interaction are mediated by a visuomotor response to action. Prior imaging studies implicate posterior parietal and pre-motor cortices, particularly within the left hemisphere, in visuomotor responses to objects ([Chao and Martin, 2000](#); [Grèzes and Decety, 2002](#)). Are effects of co-locating objects for action based on visuomotor responses to objects which afford an action, or to a purely visual response within the ventral visual stream?

## Materials and Methods

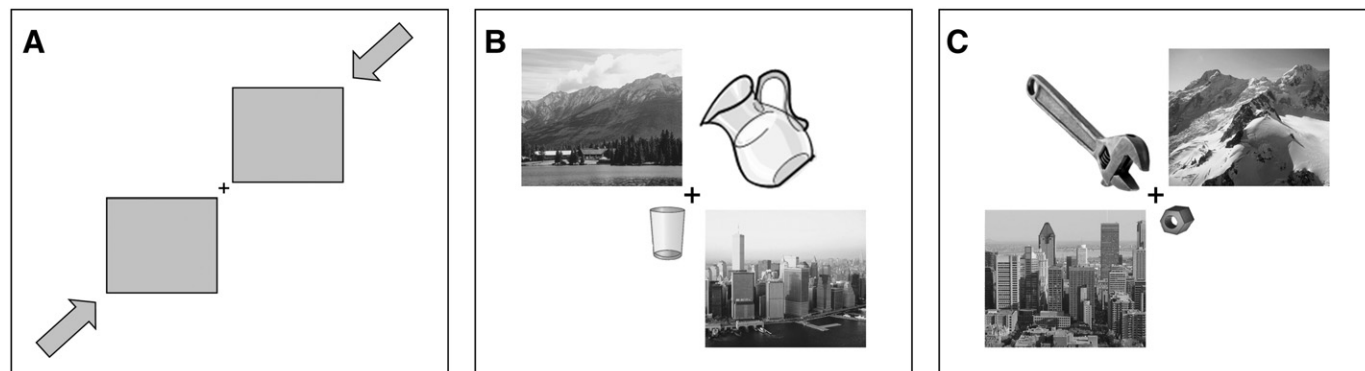
### Participants

Fourteen participants took part. Data from two participants had to be excluded due to excessive movement during the imaging session. The analyses are based on data from the remaining twelve participants (8 female, aged between 18 and 23 years, all reported being right-handed). All participants gave written informed consent and received either course credit or £15 compensation.

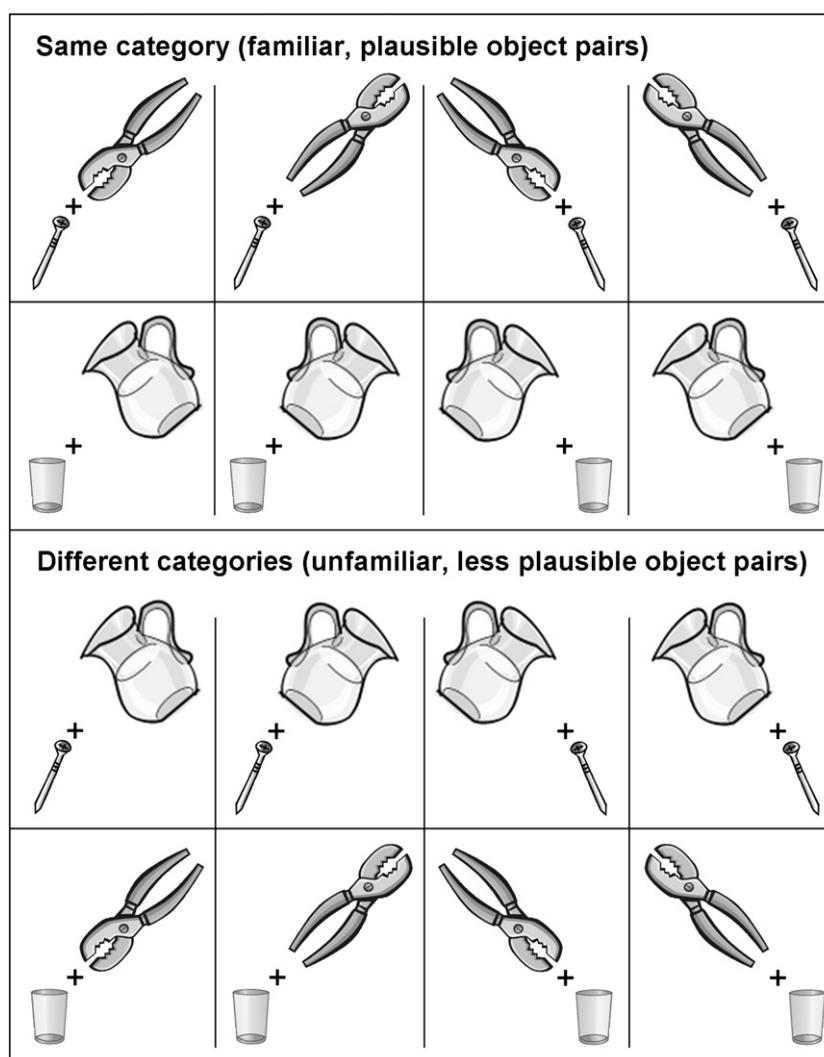
### Stimuli

On each trial participants saw a two by two array containing two objects along one diagonal (top right and bottom left or top left and bottom right) and two scenes along the other diagonal (see [Fig. 1](#) for example stimuli). On each block of trials participants were instructed to attend to the items along one of the diagonals, and to decide if the attended items, whether objects or scenes, were from the same category or from different categories. Objects could be either kitchen items or 'do-it-yourself' (DIY) objects (associated with household repairs). Scenes were either cityscapes or natural scenes. Object stimuli were greyscale clip-art style images of 16 'active' objects (8 kitchen and 8 DIY) paired with familiar 'passive' objects with which they frequently co-occurred (e.g. wine bottle and wine glass, screwdriver and screw). The objects were either positioned correctly or not positioned correctly for action. When the objects were incorrectly positioned for action, the active object was kept in the same location as before but was re-oriented to point away from the passive object ([Fig. 2](#)). Objects were re-paired across categories to create stimuli where the individual objects were in the same positions as in the familiar pairs, but without a strong action relationship (e.g. wine bottle and screw) (see [Appendix A](#) for a complete list of objects). A separate group of 16 participants (9 male, average age 23 years, all right-handed) rated each of the object pairs for, a) 'would these objects be used together?' and b) 'are these objects in the correct positions to be used together?', on a scale of 1 to 5 from 'definitely not' to 'definitely'. Familiar pairs of objects were considered more likely to be used together than unfamiliar pairs ( $F_{1,15} = 460.50$ ,  $p < 0.001$ ), particularly if they were also in correct positions for action (main effect of positioning:  $F_{1,15} = 8.27$ ,  $p < 0.05$ ; interaction between pairing and positioning:  $F_{1,15} = 5.39$ ,  $p < 0.05$ ). Familiar objects received average ratings of 4.75 (correct position) and 4.10 (incorrect position), while unfamiliar objects received average ratings of just 1.38 and 1.27, respectively. Similarly, objects that were positioned for action were considered more likely to be in the correct positions to be used together ( $F_{1,15} = 395.55$ ,  $p < 0.001$ ), and this was influenced by whether the objects formed a familiar pair (main effect of familiarity:  $F_{1,15} = 26.59$ ,  $p < 0.001$ ; interaction between pairing and positioning:  $F_{1,15} = 29.82$ ,  $p < 0.001$ ). Average ratings for correctly-positioned objects were 4.77 (familiar pairs) and 3.56 (unfamiliar pairs), while average ratings for incorrectly-positioned objects were 1.65 and 1.53, respectively. Neither rating was influenced by whether the objects were presented along the upward or downward diagonal.

Scene stimuli were greyscale photographs of 8 pairs of cityscapes and 8 pairs of natural scenes. Unrelated pairs were formed by swapping one of the scenes from each pair with a scene from the other category.



**Fig. 1.** Example stimuli. At the start of each block participants were instructed to attend to items on one of the diagonals (panel A). On each trial they decided if the attended images were from the same category (as in panel B: both kitchen objects) or different categories (as in panel C: one cityscape, one natural scene).



**Fig. 2.** Example stimuli from the experimental conditions. Objects were seen in familiar, plausible action pairs (same category: top two rows) and unfamiliar, less plausible pairs (different categories: bottom two rows). The objects were either correctly positioned for action (columns 1 and 3) or incorrectly positioned for action (columns 2 and 4), and were oriented along each diagonal (left- vs. right-hand columns).

Scenes were viewed in their original orientation on one diagonal and horizontally flipped along the other diagonal. Object and scene pairs were combined so that the objects and scenes indicated the same response on 50% of trials and different responses on 50% of trials.

#### *Design and procedure*

Prior to the imaging session participants completed a training session to ensure that they could identify the stimuli and understood the task. The imaging session included a high-resolution anatomical scan and four 8-minute blocks of the experimental task. At the start of each block participants were told which diagonal to attend to (i.e. the images top right and bottom left, or the images top left and bottom right; see Fig. 1). The attended diagonal alternated across blocks so that participants completed two blocks on each diagonal, with the initial diagonal counterbalanced across participants. One participant mistakenly attended to the same diagonal for three blocks, and one participant only completed three blocks. On each experimental trial participants viewed a fixation cross for 750 ms, followed by a 500-ms target. A question mark then appeared on screen for 1750 ms, and participants made an index- or middle-finger response to indicate whether the attended images were from the same category or from different categories. The response hand was counterbalanced across participants. The attended images could either be objects or scenes.

Whether attended or unattended, the objects formed a familiar or unfamiliar pair, and were positioned correctly or incorrectly for action. This produced 8 ( $2 \times 2 \times 2$ ) experimental conditions, where object pairs could be familiar and correctly positioned (FC), familiar and incorrectly positioned (FI), unfamiliar and correctly positioned (UC), unfamiliar and incorrectly positioned (UI), and in each case either attended or unattended. In addition to the experimental trials we included a baseline condition in which participants saw a fixation dot instead of the target and question mark, and were instructed to do nothing. Each block comprised 160 trials (32 per condition), the order of which was determined using optseq2 (Dale, 1999), a tool for optimising event-related fMRI designs. Images were presented using E-prime presentation software (PST Inc., Pittsburgh, PA). Stimuli were projected onto a screen at the back of the scanner and were viewed from a distance of approximately 65 cm via a mirror placed on the head coil. Each scene measured approximately 5.5 cm by 4.7 cm, with the inner corner 4 mm above or below and to the left or right of the centre of the fixation cross. Objects varied in size but were located within the same regions as the scenes.

#### *MRI data acquisition*

Imaging data were acquired using a Phillips 3 T Achieva scanner at Birmingham University Imaging Centre. A T1-weighted  $1 \times 1 \times 1$  mm

anatomical image was acquired for each participant. T2\*-weighted functional echoplanar imaging data were obtained using an eight-channel SENSE head coil with a sense factor of 2. For the experimental task data were acquired for 49 slices (2.5 mm<sup>3</sup> resolution, TR = 3 seconds, TE = 35 ms, flip angle = 35°).

### Data analysis

Imaging data were analysed using SPM5 (Wellcome Department of Imaging Neuroscience, London; [www.fil.ion.ucl.ac.uk/spm](http://www.fil.ion.ucl.ac.uk/spm)). Following slice timing correction, data were spatially realigned and unwarped to correct for motion artefact and distortions in the magnetic field, then transformed into MNI space and spatially smoothed using a Gaussian kernel of 8 mm full-width-at-half-maximum.

### Random effects and region of interest (ROI) analyses

Data were modelled at the individual level with regressors for each condition (attended-FC, attended-FI, attended-UC, attended-UI, unattended-FC, unattended-FI, unattended-UC, unattended-UI) convolved with the canonical haemodynamic response function (HRF) plus the temporal and dispersion derivatives. Additional regressors were included to account for incorrect responses and whether attention was directed towards the upward or downward diagonal. A 1/128 Hz high-pass filter was applied to remove low-frequency noise. Data for each experimental condition (based on the canonical HRF only), for each participant, were then entered into a second-level analysis. We first identified regions which were more active when participants attended to objects than when they attended to scenes ( $p < 0.001$  uncorrected, extent greater than 5 suprathreshold voxels). We then examined whether any of these object-processing regions were sensitive to whether the objects were familiar as a pair, whether the objects were positioned for action, and the interaction between these factors and the focus of attention (whether objects were attended or unattended). For each object-processing region (attend objects > attend scenes) we created a 5-mm radius spherical ROI centred on the peak coordinates, and extracted the parameter estimate across the ROI for each participant, for each condition. Since the primary aim of the study was to identify activation associated with the perceptual integration of plausibly-interacting objects, we additionally looked for regions that were significantly more active when attending to familiar pairs of correctly-positioned objects (attended-FC) compared with the same objects positioned incorrectly-for action (attended-FI) ( $p < 0.001$  uncorrected, extent greater than 5 suprathreshold voxels), regardless of whether activation was enhanced overall when objects were attended. As before, we examined activation at the peak of these regions (5-mm radius spherical ROI) to determine if they were responsive only to familiar pairs of interacting objects (interaction between pairing and positioning), and only when the objects were explicitly attended (interaction between the focus of attention and the position of the objects).

## Results

### Behavioural data

Participants were highly accurate. Individual accuracy rates for the twelve participants ranged from 85% to 99% (mean = 93%). Trials with incorrect responses were excluded from the reaction time (RT) analyses, which are based on participants' median RTs in each condition.

The object and scene decision tasks were well matched. To establish this, we entered the data into a 2 (attend objects, attend scenes)  $\times$  2 (same category, different categories) repeated-measures analysis of

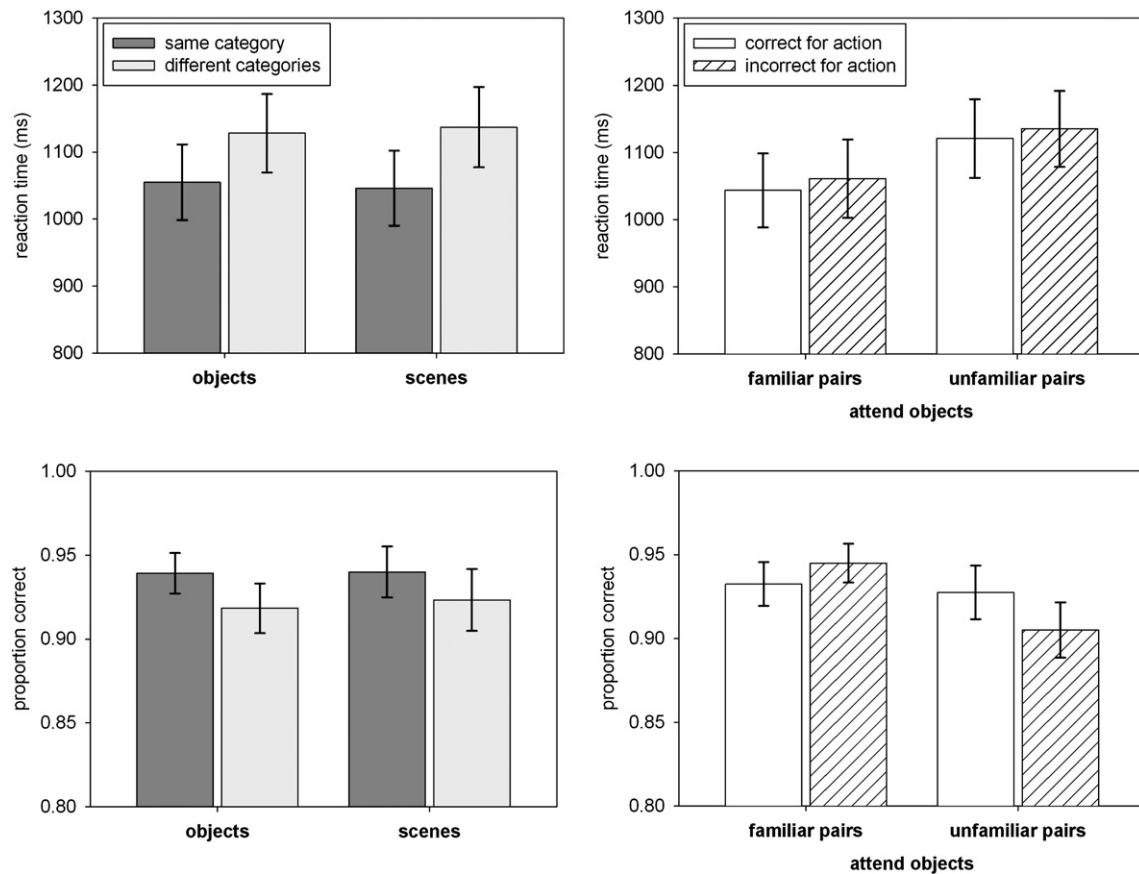
variance (ANOVA). The results indicated that participants were faster to decide that two images were from the same category than from different categories ( $F_{1,11} = 30.13$ ,  $p < 0.001$ ), but there was no difference in their response times to objects and scenes ( $F_{1,11} = 0.00$ ,  $p = 0.998$ ), and no interaction between the response and stimulus type ( $F_{1,11} = 0.50$ ,  $p = 0.49$ ). RT and accuracy data can be seen in Fig. 3. RTs appear slow as they are calculated from the onset of the target, not the onset of the question mark that prompted a response (500 ms after target onset). We next looked at RTs on trials when objects were attended to determine if there were effects of familiarity (same category, different categories) and positioning (correct or incorrect for action) (Fig. 3). Participants were faster to indicate that the objects were from the same category than different categories ( $F_{1,11} = 21.24$ ,  $p < 0.01$ ) and faster to respond when objects were positioned correctly for action ( $F_{1,11} = 6.22$ ,  $p < 0.05$ ), but there was no interaction ( $F_{1,11} = 0.01$ ,  $p = 0.91$ ). Finally, paired *t* tests were used to compare RTs on congruent trials (where the objects and scenes indicated the same response) with RTs on incongruent trials (where the objects and scenes indicated different responses). There was no evidence for conflict between the two responses, either when the objects were correctly positioned for action ( $t_{11} = 0.11$ ,  $p = 0.91$ ) or incorrectly positioned ( $t_{11} = 1.36$ ,  $p = 0.20$ ).

### Imaging data

Attending to the objects (attend objects > attend scenes,  $p < 0.001$  uncorrected, extent greater than 5 suprathreshold voxels) led to increased activation in object-processing regions, including the lateral occipital complex (LOC) and left inferior frontal gyrus (pars opercularis) (see Table 1). We conducted  $2 \times 2 \times 2$  repeated-measures ANOVAs on the estimated activation within a 5-mm radius spherical ROI placed at the peak coordinates of each of these regions, to determine if any of these object-processing regions were sensitive to the position of the objects, the familiarity/plausibility of the pairing, or the interaction between these factors and/or the direction of attention. Left and right LOC were significantly more active when participants viewed correctly-positioned objects relative to incorrectly positioned objects (left:  $F_{1,11} = 6.98$ ,  $p < 0.05$ ; right:  $F_{1,11} = 10.51$ ,  $p < 0.01$ ), but there was no interaction between the position of the objects and whether the objects formed a familiar pair and/or were attended (all *F*s less than 1) (Fig. 4; Supplementary Table 1). None of the other ROIs listed in Table 1 showed any significant effects involving object position or pairing.

Since the primary aim of the study was to identify activation associated with the perceptual integration of plausibly-interacting objects, we also looked for regions that were significantly more active when attending to familiar, correctly-positioned objects compared to the same objects positioned incorrectly for action ( $p < 0.001$  uncorrected, extent > 5 suprathreshold voxels). This was assessed regardless of whether activation was enhanced overall when objects were attended. This contrast revealed just one cluster of activation in the left fusiform gyrus (16 voxels with a peak at -35, -45, -18). As before, we were interested in whether the position-related increase in activation in this region was specific to attended, related objects, or also present with unattended and/or unrelated objects. Within this region, activation was increased when objects were positioned for action ( $F_{1,11} = 7.21$ ,  $p < 0.05$ ), and this increase in activation was enhanced when objects formed a familiar pair (interaction between pairing and positioning:  $F_{1,11} = 9.77$ ,  $p < 0.05$ ), with the same pattern of activation found irrespective of whether attention was directed towards objects or scenes (three way interaction:  $F < 1$ ) (Fig. 5; Supplementary Table 1). However, such an analysis is subject to the 'double dipping' problem (Kriegeskorte et al., 2009), as the additional comparisons are biased by the nature of the contrast used to select the ROI. To circumvent this problem, we defined the ROI for each participant based on data from the other 11 participants. In this way,





**Fig. 3.** Behavioural data. Graphs on the left show reaction times and accuracy when participants decided if pairs of objects or scenes were from the same or different categories. Graphs on the right show reaction times and accuracy when participants attended to objects that formed a familiar or unfamiliar pair, and were positioned correctly or incorrectly for action. Error bars show standard errors.

the data extracted from the ROI for each participant are independent from the data used to define the ROI. For eleven of the participants, the primary occipito-temporal cluster of activation revealed by this 'leave one out' approach was in the same location (with the same peak voxel) as that reported above ( $-35, -45, -18$ ). For these 11 participants, the ANOVA revealed a significant interaction between pairing and positioning, as before ( $F_{1,10} = 11.03$ ,  $p < 0.01$ ), with no three way interaction involving whether objects or scenes were attended ( $F < 1$ ). For the remaining participant, the primary occipito-temporal cluster was also located in fusiform gyrus, but with a peak at  $-28, -62, -12$  (and a subpeak at  $-35, -45, -18$ ). At this peak location, the data extracted from the ROI for the participant followed a different pattern from that of the other participants, resulting in only a trend towards an interaction between pairing and positioning for the whole group ( $F_{1,11} = 3.39$ ,  $p = 0.093$ ), again, with no three-way interaction involving attention to objects or scenes ( $F < 1$ ).

**Table 1**

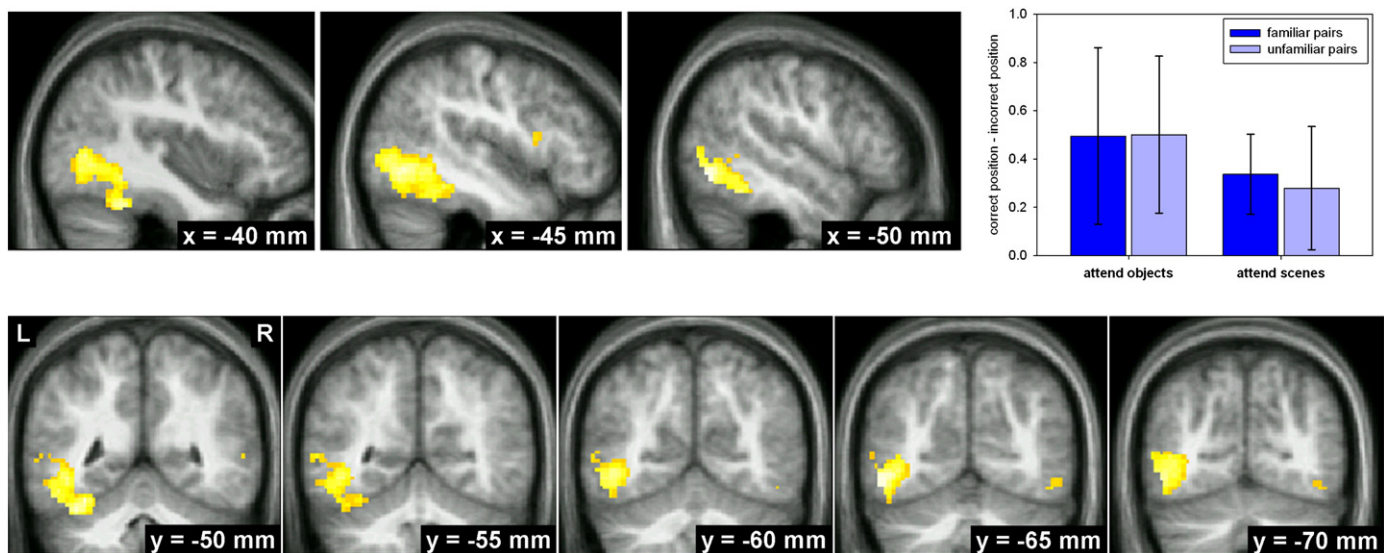
Regions showing increased activation when participants attended to objects versus scenes,  $p < 0.001$  uncorrected, extent  $> 5$  voxels. IFG: inferior frontal gyrus.

Region	MNI coordinates			z score	Cluster size
	x	y	z		
Left lateral occipital complex	-50	-68	-8	5.16	691
Right lateral occipital complex	48	-65	-12	3.52	35
	38	-88	-10	3.39	6
Right superior temporal sulcus	55	-42	8	4.42	122
Left IFG (pars opercularis)	-45	10	12	3.37	8
Caudate	8	-2	-8	3.77	14

## Discussion

The neuroimaging data indicate that the action relationship between two objects is processed in ventral-stream areas typically engaged in processing individual objects. In particular, positioning the objects in correct co-locations for action led to enhanced activation in bilateral lateral occipital complex and left fusiform gyrus. These data provide evidence that the ventral stream not only processes the identity and properties of single objects, but is also responsive to the relationship between objects. Strikingly, the action relationship influenced ventral-stream processing even when attention was directed elsewhere in the scene, consistent with the hypothesis that the action properties of objects are coded pre-attentively, at least with up to four stimuli present (Riddoch et al., 2003).

The left fusiform gyrus was specifically engaged when participants attended to familiar, correctly-positioned objects relative to familiar, incorrectly-positioned objects. A similar region of the fusiform gyrus has been found to be more engaged when participants view objects relative to 3D shapes (Creem-Regehr and Lee, 2005) and has been linked to the structural processing of objects (Liu et al., 2008). A slightly more medial region has also been found to be active when participants make judgements about how to perform actions to unfamiliar objects, but not when making action decisions to words or familiar objects (Phillips et al., 2002). This region appears to be involved in processing the structural properties of objects in relation to performing actions. Our data further suggest that recruitment of this area may be mediated by effects of the action familiarity/plausibility between object pairings. This effect mirrors the data found in extinction cases, where perceptual awareness of two objects increases when the objects form familiar/plausible pairs and are positioned correctly for action (Riddoch et al., 2003, 2006).

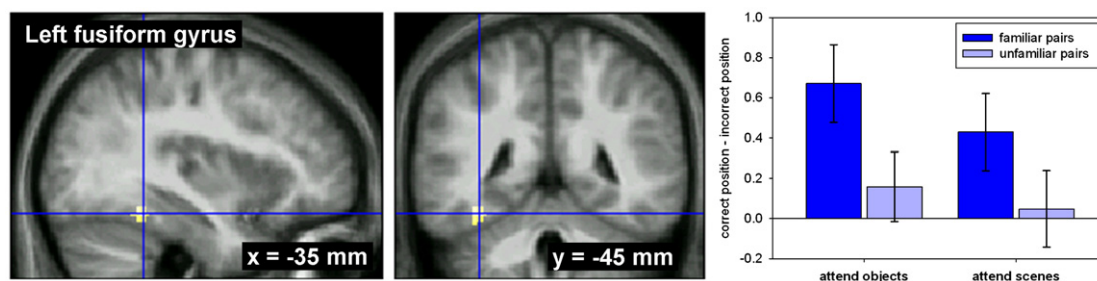


**Fig. 4.** Enhanced activation in left (−50, −68, −8 mm) and right (48, −65, −12 mm) lateral occipital complex (LOC) when attending to objects rather than scenes. The left-hemisphere cluster includes a subpeak within fusiform gyrus at −38, −48, −25 mm. The plot is based on estimated activation within a 5-mm sphere centred on the peak coordinates in left LOC, and shows the increase in activation when objects are correctly versus incorrectly positioned for action. Darker bars show the position-related increase in activation when the objects form a familiar pair; lighter bars show the position-related increase when the objects form an unfamiliar pair. Data are shown for when the objects are explicitly attended (attend objects) and when the objects are unattended (attend scenes). Error bars show standard errors.

Prior work indicates that ventral temporo-occipital cortex is relatively insensitive to changes to the orientation of graspable objects (James et al., 2002; Rice et al., 2007; Valyear et al., 2007). It is notable, then, that effects of positioning objects for action were observed in LOC, but these effects arose even when the objects formed unfamiliar and less plausible pairs. This suggests that the activation is not contingent on familiar pairings of objects but on an affordance for action offered by the active object even in the minimal interactive context provided by its correctly positioned partner. The rating data (see Methods) indicate that participants perceive an action relationship between objects even when the objects would not typically be used together: unfamiliar pairs of objects were considered more likely to be positioned for action when correctly positioned than incorrectly positioned. This is consistent with Gibson's (1979) original premise that affordances are flexible and tied as much to the viewer's goals as to the standard purpose of an object. Thus, while a jug and a glass are commonly viewed together, the pouring action associated with the jug can be aimed at a range of 'passive' objects, including a plant, basin or dessert. Similarly, a screwdriver could be used to pry the lid off a tin of paint or punch holes in the lid of a box. These visually dissimilar 'passive' objects would share a common location relative to a jug oriented to pour or relative to the pointed end of the screwdriver, providing a context that may be coded irrespective of the identities of the exemplars present. In addition, our activation within the LOC region was quite extensive and may partially reflect activation linked

to other properties of the objects (such as MT+ activation to implied motion), not just identity coding. Consistent with this last point, Riddoch et al. (2010) have recently reported that patients with parietal lesions can show reduced extinction based on implied motion between stimuli. Increased activation for linked objects in LOC could be important for this effect on extinction. The LOC region is also engaged when participants make action decisions about non-objects (Phillips et al., 2002), and nearby regions of the posterior middle temporal cortex are engaged when participants retrieve action-related information (Liu et al., 2008; Martin et al., 1995). Both findings support the argument that the LOC can respond to action possibilities even with objects that are unfamiliar as a pair.

Attending to objects rather than scenes led to extensive activation in temporo-occipital cortex, particularly in the left hemisphere. Although the analyses reported here focused on the peak activation, potentially this region encompassed functionally distinct cortical regions which responded differently to the action properties of the stimuli. Of particular interest is whether the motion-sensitive regions MT and MST were responsive to the implied motion of the action-related stimuli (Kourtzi et al., 2002; Senior et al., 1999). Future studies should include functional localizers to define object- and motion-processing regions within the temporo-occipital cortex. A recent study by Valyear and Culham (2010) incorporated such localisers and found a pattern of results highly consistent with those presented here. Participants viewed video clips of tools (e.g. trowel, whisk) being



**Fig. 5.** Region engaged when attending to correctly-positioned related objects versus incorrectly-positioned related objects ( $p < 0.001$  uncorrected, extent  $> 5$  mm). The plot shows the increase in estimated activation in left fusiform gyrus when objects were correctly versus incorrectly positioned for action, based on a 5-mm radius spherical ROI located at the peak of the activation (−35, −45, −18). Error bars show standard errors.

grasped with a typical or atypical grasp. Typical grasps produced increased activation in ventral stream regions, including MT+ and lateral occipital cortex (LO), but produced no change in frontal or parietal regions, despite these regions showing greater activation to grasping than to reaching. Valyear and Culham (2010) concluded that the ventral stream activation was associated with learned semantic or contextual associations, specifically stored knowledge of object-specific actions. Our findings indicate a similar ventral-stream response to the appropriateness of object-object (rather than hand-object) positioning, providing support for Valyear and Culham's conclusions.

Viewing graspable objects typically leads to activation in visuo-motor (premotor and parietal) regions of the cortex (Chao and Martin, 2000; Grèzes and Decety, 2002). In the present study, the inferior frontal gyrus (pars opercularis) showed increased activation when participants attended to objects versus scenes, but was not sensitive to the action relationship between the objects. At a slightly reduced threshold ( $p < 0.005$ ) a region of the left inferior parietal lobe was more responsive when attention was directed towards objects than to scenes (17 voxels, peak at -60, -42, 32), but activity in this region was also insensitive to the action relationship between the objects ( $F_s < 1$ ).

This lack of visuomotor response to the positioning of the objects highlights that the findings in LOC and fusiform gyrus reflect a visual response to the possibilities of action, i.e. one that is not driven by activation in motor regions of the cortex. This finding is somewhat surprising in light of a recent study by Mahon et al. (2007), where connectivity was reported between dorsal-stream motor areas and the ventral-stream fusiform gyrus when participants viewed pictures of tools. Mahon et al. concluded that the category-specific response to tools in the medial fusiform gyrus (Chao et al., 1999; Noppeney et al., 2006) is mediated by the motor response in the parietal lobe and middle temporal gyrus. Weisberg et al. (2007) also found that training in the functional use of novel objects resulted in increased activation in motor areas (left middle temporal gyrus, premotor and parietal regions), along with more focal and selective activation in the fusiform gyrus. Here, however, we have no evidence of a dorsal-stream response to the interaction between two objects and therefore no suggestion that the visual response is influenced by motor-related activation in these regions. Our lack of effect might partly be due to our using tools in all cases, if the prior results reflected activation to tools irrespective of their context.

In addition, it is possible that the design of the current study did not emphasise the difference in the strength of any visuomotor response across the experimental conditions. Fig. 2 illustrates objects that were positioned correctly and incorrectly for action along both the upward (active object top right) and downward (active object top left) diagonals of the display. The ease of the visuomotor response to the objects may vary with the positions of the stimuli in the display (the pliers may be more easily grasped by a right-handed viewer when incorrectly oriented along the downward diagonal rather than correctly oriented along the upward diagonal). Any context-related increase in the strength of the visuomotor response to objects positioned for action could be offset by a position-related decrease in this response. Alternatively, the task requirements may have emphasised the relationship between the objects but minimised processing of the action potential of individual objects. Boronat et al. (2005) found that the left inferior parietal lobe was more active when participants made judgements about how an object would be manipulated than about its function. Further studies are needed to investigate the role of context and object orientation on affordance-related activation in these visuomotor regions.

In both the lateral occipital complex and the fusiform gyrus, positioning the objects for action led to an increase in activation regardless of the focus of attention. This is consistent with behavioural research showing an effect of object affordances even when the object

is irrelevant to the task (Phillips and Ward, 2002) and with the early and automatic processing of action relations found in patients with extinction (Riddoch et al., 2003). In the behavioural (reaction-time) data there was no indication of interference between the responses to the attended and unattended stimuli. However, it should be noted that the competing (scene) task was of relatively low perceptual load, and so it is not certain that the same results would be found with a highly-demanding secondary task.

This neuroimaging study provides a first insight into the neural processes involved in processing the action relationship between two objects. It indicates that the action relationship is coded in ventral stream regions involved in processing action properties of objects, and that this information is coded even when attention is directed towards competing stimuli (scenes rather than objects). The effect of action on visual coding can occur even when visuomotor responses are not differentially affected by the positioning of the stimuli.

**Appendix A. Object stimuli. Objects were viewed in their original pairs and also re-paired across categories (i.e. the corkscrew was viewed with both the wine bottle and the paint tray). Different exemplars of wine bottles were used for sets 1 and 5.**

	Kitchen objects		DIY objects	
	Active	Passive	Active	Passive
1	corkscrew	wine bottle	paint roller	paint tray
2	jug	tumbler	pliers	nail
3	spoon	sundae	spanner	nut
4	soup spoon	dish	wire cutters	barbed wire
5	wine bottle	wine glass	screwdriver	screw
6	teaspoon	boiled egg	saw	log
7	chopsticks	bowl	wire strippers	plug wires
8	juicer	orange	paint brush	paint tin

## Appendix B. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.neuroimage.2010.05.044.

## References

- Biederman, I., 1972. Perceiving real world scenes. *Science* 177, 77–80.
- Boronat, C.B., et al., 2005. Distinctions between manipulation and function knowledge of objects: Evidence from functional magnetic resonance imaging. *Cognitive Brain Research* 23, 361–373.
- Chao, L.L., Martin, A., 2000. Representation of manipulable man-made objects in the dorsal stream. *Neuroimage* 12, 478–484.
- Chao, L.L., Haxby, J.V., Martin, A., 1999. Attribute-based neural substrates in temporal cortex for perceiving and knowing about objects. *Nature Neuroscience* 2, 913–919.
- Creem-Regehr, S.H., Lee, J.N., 2005. Neural representations of graspable objects: Are tools special? *Cognitive Brain Research* 22, 457–469.
- Dale, A.M., 1999. Optimal experimental design for event-related fMRI. *Human Brain Mapping* 8, 109–114.
- Desimone, R., Duncan, J., 1995. Neural mechanisms of selective visual attention. *Annual Review of Neuroscience* 18, 193–222.
- Gibson, J.J., 1979. *The Ecological Approach to Visual Perception*. Lawrence Erlbaum, Hillsdale, NJ.
- Gilchrist, G., Humphreys, G.W., Riddoch, M.J., 1996. Grouping and extinction: evidence for low-level modulation of visual selection. *Cognitive Neuropsychology* 13, 1223–1249.
- Grèzes, J., Decety, J., 2002. Does visual perception of object afford action? Evidence from a neuroimaging study. *Neuropsychologia* 40, 212–222.
- Grill-Spector, K., 2003. The neural basis of object perception. *Current Opinion in Neurobiology* 13, 159–166.
- Humphreys, G.W., Riddoch, M.J., 2006. Features, objects, action: The cognitive neuropsychology of visual object processing, 1984–2004. *Cognitive Neuropsychology* 23, 156–183.
- Jacques, C., Rossion, B., 2004. Concurrent processing reveals competition between visual representations of faces. *NeuroReport* 15, 2417–2421.
- James, T.W., Humphrey, G.K., Gati, J.S., Menon, R.S., Goodale, M.A., 2002. Differential effects of viewpoint on object-driven activation in dorsal and ventral streams. *Neuron* 35, 793–801.

- Kanwisher, N., McDermott, J., Chun, M.M., 1997. The fusiform face area: A module in human extrastriate cortex specialized for face perception. *The Journal of Neuroscience* 17, 4302–4311.
- Kastner, S., et al., 2001. Modulation of sensory suppression: Implications for receptive field sizes in the human visual cortex. *Journal of Neurophysiology* 86, 1398–1411.
- Kourtzi, Z., Bühlhoff, H.H., Erb, M., Grodd, W., 2002. Object-selective responses in the human motion area MT/MST. *Nature Neuroscience* 5, 17–18.
- Kourtzi, Z., Kanwisher, N., 2001. Representation of perceived object shape by the human lateral occipital complex. *Science* 293, 1506–1509.
- Kriegeskorte, N., Simmons, W.K., Bellgowan, P.S.F., Baker, C.I., 2009. Circular analysis in systems neuroscience: The dangers of double dipping. *Nature Neuroscience* 12, 535–540.
- Liu, X., Steinmetz, N.A., Farley, A.B., Smith, C.D., Joseph, J.E., 2008. Mid-fusiform activation during object discrimination reflects the process of differentiating structural descriptions. *Journal of Cognitive Neuroscience* 20, 1711–1726.
- Mahon, B.Z., Millelville, S.C., Negri, G.A.L., Rumiat, R.I., Caramazza, A., Martin, A., 2007. Action-related properties shape object representations in the ventral stream. *Neuron* 55, 507–520.
- Martin, A., Haxby, J.V., Lalonde, F.M., Wiggs, C.L., Ungerleider, L.G., 1995. Discrete cortical regions associated with knowledge of color and knowledge of action. *Science* 270, 102–105.
- Nobre, A.C., Allison, T., McCarthy, G., 1994. Word recognition in the human inferior temporal lobe. *Nature* 372, 260–263.
- Noppeney, U., Price, C.J., Penny, W.D., Friston, K.J., 2006. Two distinct neural mechanisms for category-selective responses. *Cerebral Cortex* 16, 437–445.
- Op de Beeck, H., Bèatse, E., Wagemans, J., Sunaert, S., Van Hecke, P., 2000. The representation of shape in the context of visual object categorization tasks. *Neuroimage* 12, 28–40.
- Phillips, J.A., Humphreys, G.W., Noppeney, U., Price, C.J., 2002. The neural substrates of action retrieval: An examination of semantic and visual routes to action. *Visual Cognition* 9, 662–684.
- Phillips, J.C., Ward, R., 2002. S-R correspondence effects of irrelevant visual affordance: Time course and specificity of response activation. *Visual Cognition* 9, 540–558.
- Rice, N.J., Valyear, K.F., Goodale, M.A., Milner, A.D., Culham, J.C., 2007. Orientation sensitivity to graspable objects: An fMRI adaptation study. *Neuroimage* 36, T87–T93.
- Riddoch, M.J., Bodley-Smith, S., Humphreys, G.W., 2010. No direction home: Extinction is affected by implicit motion. *Cortex* 46, 678–684.
- Riddoch, M.J., Humphreys, G.W., Edwards, S., Baker, T., Willson, K., 2003. Seeing the action: Neuropsychological evidence for action-based effects on object selection. *Nature Neuroscience* 6, 82–89.
- Riddoch, M.J., et al., 2006. I can see what you are doing: Action familiarity and affordance promote recovery from extinction. *Cognitive Neuropsychology* 23, 583–605.
- Rolls, E.T., Tovee, M.J., 1995. The responses of single neurons in the temporal visual cortical areas of the macaque when more than one stimulus is present in the receptive-field. *Experimental Brain Research* 103, 409–420.
- Senior, C., Barnes, J., Giampietro, V., Simmons, A., Bullmore, E.T., Brammer, M., David, A. S., 1999. The functional neuroanatomy of implicit-motion perception or 'representational momentum'. *Current Biology* 10, 16–22.
- Spiridon, M., Fischl, B., Kanwisher, N., 2006. Location and spatial profile of category-specific regions in human extrastriate cortex. *Human Brain Mapping* 27, 77–89.
- Valyear, K.F., Cavina-Pratesi, C., Stiglick, A.J., Culham, J.C., 2007. Does tool-related fMRI activity within the intraparietal sulcus reflect the plan to grasp? *Neuroimage* 36, T94–T108.
- Valyear, K.F., Culham, J.C., 2010. Observing learned object-specific functional grasps preferentially activates the ventral stream. *Journal of Cognitive Neuroscience* 22, 970–984.
- Weisberg, J., van Turenout, M., Martin, A., 2007. A neural system for learning about object function. *Cerebral Cortex* 17, 513–521.