Abstract

There is considerable evidence that there are anatomically and functionally distinct pathways for action and object recognition. However, little is known about how information about action and objects is integrated. This study provides fMRI evidence for task-based selection of brain regions associated with action and object processing, and on how the congruency between the action and the object modulates neural response. Participants viewed videos of objects used in congruent or incongruent actions and attended either to the action or the object in a one-back procedure. Attending to the action led to increased responses in a fronto-parietal action-associated network. Attending to the object activated regions within a fronto-inferior temporal network. Stronger responses for congruent action–object clips occurred in bilateral parietal, inferior temporal, and putamen. Distinct cortical and thalamic regions were modulated by congruency in the different tasks. The results suggest that (i) selective attention to action and object information is mediated through separate networks, (ii) object–action congruency evokes responses in action planning regions, and (iii) the selective activation of nuclei within the thalamus provides a mechanism to integrate task goals in relation to the congruency of the perceptual information presented to the observer.

INTRODUCTION

The ability to manipulate objects and tools is crucial for everyday life. This ability depends on an interaction between the properties of the object and the required action. Object-related actions involve on-line guidance of our effectors in response to these objects, thus object affordances (the potency of the object for action) may be an inherent part of object perception (Tucker & Ellis, 1998; Gibson, 1979). This study is concerned with how we process action, objects, and their relations.

Traditionally, theories have stressed that the retrieval of an appropriate action for an object is guided by access to semantic knowledge based on an object’s associations and its abstract function (e.g., Ochipa, Rothi, & Heilman, 1992; Roy & Square, 1985). For instance, a cup activates the action of drinking through access to semantic knowledge based on our prior associations with how cups are used and what they are used for. However, there is increasing neuropsychological and experimental evidence that access to action information can be evoked by the visual properties of objects in a relatively direct way, without the necessary involvement of semantic (associative) memory (see Humphreys et al., 2010; Humphreys & Riddoch, 2003, for reviews; also see Barsalou, 1999, for a similar view derived from a different literature). This evidence provides the backdrop for this study, in which we had participants make 1-back judgments on images of objects being manipulated by congruent or incongruent actions. We examine first the distinct neural correlates of action and object-related processing invoked by attention to actions and to objects. Subsequently, we examine the neural basis of the interaction between action and object pathways driven by the congruency of the action–object pairing.

Neuropsychological research suggests that knowledge of actions (how to manipulate objects) is dissociated from knowledge of objects (what the object is). Ferreira and colleagues (Ferreira, Giusiano, Ceccaldi, & Poncet, 1997) report a patient who was impaired at naming objects but unimpaired at naming actions and who was able to name the objects from gesture information. Additional patients with damage to the left occipito-temporal cortex, who have the clinical presentation of optic aphasia, are able to gesture the correct action to objects they cannot recognize, even when access to associative semantic information is impaired (Hillis & Caramazza, 1995; Riddoch & Humphreys, 1987). In such cases, vision is not simply used for on-line control of prehensile actions but also to access information about the category of action that can be performed. The results indicate that responding to action-related associations to objects can be dissociated from semantic recognition processes.

In contrast, there are other patients who have an intact access to semantic knowledge of objects, showing intact recognition and naming. These patients can produce correct action in response to the object’s name, but they...
Nevertheless fail to correctly act when objects are presented visually (e.g., Pilgrim & Humphreys, 1991; Riddoch, Humphreys, & Price, 1989; DeRenzi, Faglioni, & Sorgato, 1982). In these cases, visual access to semantics is preserved but visual access to action-related responses is blocked, though they are able to access action information via the indirect semantic route (e.g., from the object’s name).

Behavioral experiments with healthy participants also provide evidence for direct action-related and indirect semantic-related responses to objects (e.g., Yoon, Humphreys, & Riddoch, 2010; Yoon & Humphreys, 2005, 2007; Chainay & Humphreys, 2002). Decisions about which action to perform on an object are faster when made on pictures of objects than on object names, whereas decisions about the semantic context associated with objects are not affected by the mode of presentation (Chainay & Humphreys, 2002). Also in contrast with semantic categorical decisions, action decisions are not affected by semantic priming but are affected by the orientation of the handle to viewers (Yoon & Humphreys, 2007) and by the correct relative positions of paired objects within an observer’s egocentric reference frame (Yoon et al., 2010).

The evidence indicates that access to action and semantic information about objects can dissociate. The interrelations between objects and action, however, can affect responses based on the action and semantic recognition routes (Yoon & Humphreys, 2005). Notably, using stimuli similar to those employed in our study, Yoon and Humphreys found that both action and semantic decisions to objects are facilitated if stimuli are presented with a congruent object grip and if the object is used appropriately compared with when the grip or the action is inappropriate to the object.

We note that the behavioral differences observed for accessing action-related versus semantic contextual knowledge occur even when explicit actions are not made by the participants to the objects. It is nevertheless possible that action-related effects are contingent on the activation of motor actions to objects modulated through the so-called mirror neuron system, commonly involved in action production and recognition. There is evidence that viewing actions evokes a simulated response in motor cortex (for a recent review, Rizzolatti & Sinigaglia, 2010). The neural areas comprising the action observation network (Grafton, 2009) include the bilateral STS, the inferior parietal lobule (IPL), the inferior frontal gyrus (IFG), and the premotor cortex (PM). In addition, the SMA (Dayan et al., 2007; Hamilton & Grafton, 2007), BG, and cerebellum (Blakemore, Frith, & Wolpert, 2001; Wolpert, Miall, & Kawato, 1998) have also been implicated in action simulation. Interestingly, Humphreys et al. (2010) report increased activity in dorsal PM when participants view objects gripped in a congruent relative to an incongruent manner, suggesting that motor-based simulation may also be sensitive to the interaction between action and object information.

Access to action-related information (how/where to grasp objects) and to semantic knowledge (what) has been linked to the dorsal and ventral visual streams (Milner & Goodale, 1995; Ungerleider & Mishkin, 1982). For example, Shmuelof and Zohary (2005) reported that, when participants see objects being manipulated, then recognition of the action is associated with activity in dorsal areas, including the anterior intraparietal sulcus (aIPS); while in contrast, recognition of the object is linked to activity within ventral regions including the fusiform gyrus (e.g., Grill-Spector, 2003; Chao, Hazby, & Martin, 1999; Grill-Spector, Kushnir, Edelman, Ittzak, & Malach, 1998). Consistent with this, the aIPS shows adaptation of activity when the same grasp information is viewed repeatedly, whereas adaptation in the fusiform gyrus depends on the repetition of the identity and form of the object (Shmuelof & Zohary, 2005). The aIPS appears insensitive to the long-term familiarity of the grasp, including whether a grasp is congruent with the correct action (Valyear & Culham, 2009). In other studies, observation of object manipulation has been associated with bilateral activation of the PM along with the IPS (e.g., Tunik, Rice, Hamilton, & Grafton, 2007; Baccino et al., 2001, for reviews), whereas the IFG is bilaterally involved when participants view static images of an object being grasped by a hand irrespective of whether it is held correctly or incorrectly for object use (Johnson-Frey et al., 2003). Valyear and Culham (2009), however, argue that grasp information is also encoded by the ventral stream regions. Using an ROI approach, they report that regions sensitive to objects, body parts, and motions in the occipital–temporal cortex are sensitive to whether stimuli are correctly grasped for actions.

In addition to cases where participants see actions performed on objects, objects that are inherently associated with actions, tools, and manipulable objects are reported to elicit responses in the dorsal action associated network. Specifically, the left anterior supramarginal gyrus (SMG; the rostral part of IPL) is suggested to code motor programs for object use (Peeters et al., 2009). The left anterior SMG, along with the posterior SMG/angular gyrus, is associated with planning the use of familiar objects (Johnson-Frey, Newman-Norlund, & Grafton, 2005). Damage to the SMG (e.g., Sunderland, Wilkins, & Dineen, 2011; Randerath, Goldenberg, Spijkers, Li, & Hermston, 2010), along with virtual lesions of this region (following TMS), is linked to poor use of tools. Observation of manipulable objects has also been shown to elicit activity in additional regions including the left inferior frontal lobe, the posterior middle temporal gyrus, and the posterior parietal cortex (e.g., Beauchamp, Lee, Hazby, & Martin, 2002; Devlin et al., 2002; Grèzes & Decety, 2002). However, it is unclear how responses in these regions are affected by attention and the congruency of action–object relations.

In the current study, we investigated the neural correlates of action and object processing and their interaction in response to video clips of objects being used in a congruent or incongruent manner. To selectively evoke action versus object (semantic) processing, we manipulated the attended property of the combined stimuli (similar to
Yoon & Humphreys, 2005). On identical video stimuli, participants were asked to perform a 1-back task on the action or on the object. To further facilitate the involvement of the semantic route and to ensure that the comparison across objects could not be based on simple visual features, the object-related 1-back task required object recognition of two different exemplars of an object category (e.g., two different cups). The congruency of the action was manipulated to gain a better understanding of how action and object information is integrated. We manipulated the congruency of the action to the objects. Thus, actions could be congruent with the identity of the object (hitting with a hammer) or incongruent (twisting a hammer). We hypothesized that responses would be elicited in the dorsal action network when participants selectively attended to the action, whereas the ventral object network would be activated by attending to the objects. We further hypothesized that the involvement of these networks would be greater for congruent than for incongruent object–action relations, as congruent action–object relations lend themselves more easily to motor simulations and facilitate action and object recognition (Yoon & Humphreys, 2005).

A final question of interest concerned the interaction between task and congruency. What brain areas enable information about the congruency of the action to modulate the object network and are different regions recruited to enable the congruency of the object to modulate the action network? Here we expected to find regions that take information about both objects and actions and respond differentially according to whether this information is congruent or incongruent. We speculated that such regions will be interconnected with the action observation and object-associated networks.

**METHODS**

**Participants**

Seventeen (11 women) right-handed healthy adults (averaged age = 24.5 years, SD = 6 years) participated for course credit or cash (£15). None had a prior history of neurological or psychiatric symptoms, and all had normal or corrected-to-normal vision. All participants signed an informed consent after the procedures were explained to them. The study was approved by the local ethics committee.

**Stimuli**

There were 120 movie clips depicting an animated action performed with an object (Yoon & Humphreys, 2005). Sixty clips depicted an action congruent to the typical way the object is used (e.g., hitting with a hammer), and 60 depicted an incongruent action (e.g., wiping with a hammer; see Figure 1 for still examples extracted from one clip). Each clip lasted about 1.425 sec. Actions are typically associated with a particular grip and the way one grasps an object affect the way it can be used (Johnson-Frey, 2003, 2004; Johnson & Grafton, 2003); therefore, we kept the action–grip relation compatible across different objects. This meant that, in the incongruent condition, both the action and the grip were incongruent with the object on some trials. To account for this variability, the clips were coded according to whether the type of grip was the same (0) or different (1) across the congruent and incongruent displays. In addition, the actions differed in the number of hands (one or two) used, and this was coded as well for each trial. This information was used later as a covariate of no interest in the fMRI analysis (see below).

**Experimental Design**

A full factorial within-participant design was used, with the following factors: Task (object, action) and Stimulus (congruent, incongruent). We used a 1-back task and manipulated the relevant dimension attended. Thus, participants performed a 1-back task focusing on the objects (ignoring the action being performed) or focusing on the actions (the way the hand moved; in this case, ignoring the object). Note that identical stimuli were used for the two tasks. To ensure that the object task could not
be performed based on low-level visual cues only, repetitions were of different exemplars of the same object (e.g., two different hummers). The task was manipulated across blocks, with six blocks per task, and the order of the blocks was randomized. Each block started by presenting the words “focus on hand movements” or “focus on objects” for 3 sec. In addition, an interstimulus fixation shape was used to remind participants of the relevant dimension (“^” for action and “o” for object). Stimulus congruency was manipulated as an event. There were 8–12 events in each block appeared in a random order. Each event started with a 500-msec fixation stimulus, which was followed by a 1425 msec clip depicting actions performed on objects. Events were separated by an ISI of 2–6 sec. There were in total 30 events for each of the condition (e.g., four conditions). The repetition events were relatively rare (16.7%) and were modeled separately in the fMRI design. Importantly, the trials of interests did not require a response; hence, any responses in motor-associated regions cannot be simply attributed to a hand movement. The experiment was realized using Cogent toolbox in Matlab (www.fil.ion.ucl.ac.uk/~cogent). The experiment was run in two fMRI sessions.

fMRI Data Acquisition

We used a Phillips 3-T Achieva placed in Birmingham University Imaging Center to acquire BOLD, contrast-weighted EPIs. Thirty-eight frontal–temporal oblique slices, 2-mm thickness with a 1.25-mm gap, were acquired in ascending order, with an in-plane resolution of 2 × 2 mm, 80° flip angle, 35 msec echo time, and 2400 msec slice repetition time. Images were acquired using an eight-channel phase array coil with a sense factor of 2.

fMRI Data Analysis

fMRI data was analyzed with SPM8 (www.fil.ion.ucl.ac.uk/~spm). Preprocessing of the data included correction of head motion (Ashburner & Friston, 2003a) and distortion by head motion interactions (spatial realignment and unwrapped; Andersson, Hutton, Ashburner, Tumer, & Friston, 2001), temporal realignment, transforming the data to Montreal Neurological Institute (MNI) space (Ashburner & Friston, 2003b), reslicing to 3 × 3 × 3 mm voxels, and smoothing using a Gaussian kernel with FWHD of 9 × 9 × 9 to account for residual intersubject differences and to adhere to the continuity assumption of random field theory (Worsley & Friston, 1995).

Summary statistics (Penny, Holmes, & Friston, 2003) using the general linear framework were performed to test the reliability of the effects across participants (Kiebel & Holmes, 2003). We first estimated the effect of each condition at the individual level. For each participant, we generated a model that included the onsets of an event of each condition (2 × 2; Task × Congruency); in addition, we modeled as covariates the number of hands used to grip each object and the similarity of the incongruent grip to the grip given on congruent trials. The repetition events were modeled separately to ensure that our results were not affected by the requirement to make a motor response to these trials. To account for the delay in the hemodynamic function, these regressors were convolved with the canonical hemodynamic response function. Finally, the six realignment parameters, the session effect, and the harmonic modeling of slow fluctuations in the signal with a cutoff of 1/128 Hz (typically associated with biological and scanner noise) were also included in each individual’s model. The effect size of each condition was estimated for each subject.

To enable generalization of the results, a random effect analysis was conducted, treating participants as a random variable. In this second analysis, maps depicting the effect size per condition per participant were used as the independent variables; dependency between conditions was assumed. We used a mixed peak and threshold approach (Poline, Worsley, Evans, & Friston, 1997), with Z > 3.6 required for a peak and at least 20 voxels showing a Z > 2.69. We focused on results that survive family-wise correction, but for completeness, we mention in the tables and text all clusters that also survived the above threshold. Anatomical labeling was assigned using the Anatomical Automatic Labeling toolbox (Tzourio-Mazoyer et al., 2002) and the Duvernoy Human Brain Atlas (Duvernoy, Cabanis, & Vannson, 1991). The descriptive bar plots in the figures are based on the estimated effect size (beta values) computed in the general linear model. These values were extracted from a 3-mm sphere centered on the group peaks.

RESULTS

Behaviorally, observers found the object task easier than the action task. Responses were also more accurate during the object than the action task (95.07 ± 1.93 and 90.76 ± 1.78, respectively; F(1, 16) = 9.3, p < .01), although note that overall accuracy in both tasks was high. Similarly, accurate responses to repeated objects were carried out more rapidly than those to repeated actions (1038.8 ± 41.29 and 1360.5 ± 40 msec, respectively; F(1, 16) = 24.4, p < .001). The differences in RT may partly be affected by the nature of the stimuli that define actions and objects, because actions but not objects depend on a signal that emerges over time.

More interestingly, task affected the brain response, despite the tasks being performed on identical stimuli (Figure 2; Table 1). Attending to the objects (vs. the action) elicited a larger response in the superior medial frontal gyrus (SmFG) and, at a lower threshold, the object task activated the left inferior temporal gyrus (ITG, MNI: −45 −43 −8; peak Z = 3.52; only 14 voxels with Z > 2.69). As the reliability of the ITG cluster was below our threshold, this result should be treated with caution. We nevertheless report this activation, as it fits with our original predictions, that is, that attention to objects
would involve regions that are part of the ventral visual stream (see Introduction). Furthermore, the ITG was also affected by stimulus congruency (see below, Table 2), which may have attenuated the effect of the task. In contrast to the object attention task, attending to the action (rather than the object) led to an increased response in the action observation network. In the parietal cortex, this included the bilateral SMG extending to the post-central sulcus (CS), plus the bilateral IPS extending to the precuneus; in frontal regions, there was bilateral activation in the pre-CS (below threshold on the left as well) and the right IFG.

Table 1. Main Effects of Task: Action versus Objects

<table>
<thead>
<tr>
<th>Anatomy</th>
<th>Cluster Size</th>
<th>Peak-Z</th>
<th>x y z (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Main Effect Task: Object &gt; Action</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SmFG</td>
<td>159*</td>
<td>3.53</td>
<td>-9 47 43</td>
</tr>
<tr>
<td>ITG</td>
<td>14</td>
<td>3.52</td>
<td>-45 -43 -8</td>
</tr>
<tr>
<td><strong>Main Effect Task: Action &gt; Object</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SMG</td>
<td>246*</td>
<td>4.73*</td>
<td>-60 -40 31</td>
</tr>
<tr>
<td>Post-CS/SMG</td>
<td>88</td>
<td>3.46</td>
<td>54 -22 34</td>
</tr>
<tr>
<td>IPS, SPG</td>
<td>99</td>
<td>4.28</td>
<td>12 -70 55</td>
</tr>
<tr>
<td>IPS, SPG</td>
<td>71</td>
<td>3.98</td>
<td>-12 -67 52</td>
</tr>
<tr>
<td>IFG</td>
<td>168*</td>
<td>3.84</td>
<td>57 14 4</td>
</tr>
<tr>
<td>Pre-CS</td>
<td>24</td>
<td>3.27</td>
<td>-21 -1 58</td>
</tr>
<tr>
<td>Pre-CS</td>
<td>50</td>
<td>3.77</td>
<td>30 -4 46</td>
</tr>
</tbody>
</table>

Cluster size reliability: #voxels > 150, FWE-corr p < .05; #voxels > 80, uncorr p < .01; #voxels > 50, uncorr p < .05.

SPG = superior parietal gyrus.

*At peak/cluster level p (FWE-corrected) < .05.
A reverse pattern was observed in the lateral-dorsal (LD) part of the thalamus; in this case, there were stronger congruency effects (Cong > Inc) when participants attended to objects. Greater congruency effects linked to the object task were also observed in the bilateral temporal poles, the right superior frontal sulcus, the cerebellum and the pons.

To ensure that the neuroanatomical dissociation effects observed in the thalamus were not a by-product of the group analysis (e.g., averaging across different brains and smoothing the data), we analyzed the data at the individual subject level, assessing now the reliability of the effects across trials within each subject. We used unsmoothed data in this analysis. All participants showed a clear neuroanatomical dissociation within the thalamus when attending to actions versus objects that depended also on object–action congruency (see Table 4). Within the left thalamus 12 subjects showed dissociated responses between two different nuclei, one associated with congruent stimuli when actions were attended and the other when objects were attended (compared with the corresponding incongruent conditions); in two subjects, the left thalamus responded only to congruent stimuli when objects were attended; and in three subjects, congruency effects were observed only when actions were attended. Within the right thalamus, nine participants

<table>
<thead>
<tr>
<th>Table 2. Stimulus Effects: Congruent versus Incongruent</th>
<th>Cluster Size</th>
<th>Peak-Z</th>
<th>x y z (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Main Effect Cong: Cong &gt; Incong</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SFG, SMA, BA 6</td>
<td>66</td>
<td>4.2</td>
<td>−12 29 61</td>
</tr>
<tr>
<td>IPL</td>
<td>52</td>
<td>4.05</td>
<td>−51 −55 52</td>
</tr>
<tr>
<td>IPL</td>
<td>49</td>
<td>3.87</td>
<td>54 −55 49</td>
</tr>
<tr>
<td>Putamen/GP</td>
<td>103</td>
<td>3.69</td>
<td>24 −1 1</td>
</tr>
<tr>
<td>Putamen/GP</td>
<td>39</td>
<td>3.18</td>
<td>−24 −7 1</td>
</tr>
<tr>
<td>ITG/FFG</td>
<td>43</td>
<td>3.68</td>
<td>−48 −43 −17</td>
</tr>
<tr>
<td>STS</td>
<td>80</td>
<td>3.57</td>
<td>63 −22 −2</td>
</tr>
<tr>
<td><strong>Main Effect Cong: Incong &gt; Cong</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mMOG, lingual G, Cal BA 17, 18</td>
<td>98*</td>
<td>4.1</td>
<td>−15 −97 −2</td>
</tr>
<tr>
<td>Cal, lingual G BA 17</td>
<td>83*</td>
<td>3.67</td>
<td>9 −91 −11</td>
</tr>
</tbody>
</table>

Cluster size reliability: #voxels > 80, uncorr-p < .01; #voxels > 50, uncorr-p < .05. SFG, superior frontal gyrus; GP, globus palidus; FFG, fusiform gyrus; MOG, medial middle occipital gyrus; Cal, Calcarine sulcus; Cong, Congruent; Inc, Incongruent.

*At cluster level p(FDR-corrected) < .05.

Figure 3. SPM blobs overlaid on a T1 single subject template (side columns) and a lateral view of a rendered brain. The SPM outputs were thresholded at p < .005, uncorrected, with at least 40 voxels. Foci showing a larger response to incongruent stimuli (Inc) are depicted in blue; foci showing larger responses to the congruent stimuli (Con) are depicted in yellow. The bars show the averaged effect size extracted from a 3-mm sphere around the group maxima. Error bars depict SEMs. R, right; L, left; 1, calcarine sulcus (Cal); 2, STS; 3, SmFG; 4, putamen; 5, IPL angular gyrus; 6, fusiform gyrus (FFG)/ITG.
showed congruency responses in two different nuclei that depended on the task, three showed congruency effects only when objects were attended, and three showed congruency responses only when actions were attended. This analysis demonstrated that the neuroanatomical dissociation between different thalamic nuclei, reflecting combined effects of task and stimulus congruency, were reliable also at the individual subject level.

We next examined the cortical connections of the thalamic nuclei that were affected by both the task and stimulus congruency. This was done using the Thalamic Connectivity Atlas developed by FMRIB (www.fmrib.ox.ac.uk/connect/). We found that the VPL region (MNI: $-15 -19 -2$), showing a larger congruency effect in the action task, had a high probability of connection to cortical regions involved in action planning and motor control, including the primary motor cortex (.43), the sensory cortex (.42), the PM (.43), and, more weakly, the posterior parietal cortex (.09). In contrast the LD (MNI: $-15 -19 19$), which showed a greater congruency effect in the object task, had a high probability of connectivity to regions associated with object processing; this included the temporal cortex (.36), ventral pFC (.29), posterior occipital cortex (.04), and also, more weakly, the regions associated with motor responses (the posterior parietal cortex [.14] and the PM [.04]).

**DISCUSSION**

Our study investigated the processing of object–action relations by manipulating attention to displays showing

### Table 3. Task-by-Stimulus Interaction

<table>
<thead>
<tr>
<th>Anatomy</th>
<th>Cluster Size</th>
<th>Peak-Z</th>
<th>x y z (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>INT: Act, Cong &gt; Inc &amp; Obj, Inc &lt; Cong</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thalamus (VPL, Pul)</td>
<td>29</td>
<td>4.22</td>
<td>$18 -22 -8$</td>
</tr>
<tr>
<td>Lingual gyrus/MOG</td>
<td>137**</td>
<td>3.93</td>
<td>$-21 -88 -20$</td>
</tr>
<tr>
<td>Calcarine (MOG)</td>
<td>137**</td>
<td>3.8</td>
<td>$18 -97 1$</td>
</tr>
<tr>
<td><strong>INT: Obj, Cong &gt; Inc &amp; Obj: Inc &gt; Cong</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Temporal pole</td>
<td>207*</td>
<td>4.68*</td>
<td>$-39 8 -26$</td>
</tr>
<tr>
<td>MTL</td>
<td>117</td>
<td>4.26</td>
<td>$12 -1 -23$</td>
</tr>
<tr>
<td>Thalamus (LD)</td>
<td>102</td>
<td>4.12</td>
<td>$-15 -19 19$</td>
</tr>
<tr>
<td>SFS</td>
<td>122</td>
<td>3.93</td>
<td>$21 2 37$</td>
</tr>
<tr>
<td>Cerebellum (Lobule VI)</td>
<td>49</td>
<td>3.78</td>
<td>$6 -67 -26$</td>
</tr>
<tr>
<td>Pones</td>
<td>53</td>
<td>3.68</td>
<td>$0 -37 -23$</td>
</tr>
</tbody>
</table>

Cluster size reliability: #voxels > 150, FEW-corr $p < .05$; #voxels > 80, uncorr-$p < .01$; #voxels > 50, uncorr-$p < .05$. Cong, Congruent; Inc, Incongruent; Act, Action; Obj, Object; VPL, VPL thalamic nucleus; Pul, pulvinar; MOG, middle occipital gyrus; MTL, medial-temporal lobe including amygdala and the hippocampus structures; LD, LD thalamic nucleus; SFS, superior frontal sulcus.

*At peak/cluster level $p$(FWE-corrected) < .05.

**At cluster level $p$(FWE-corrected) = .08; $p$(FDR-corrected) = .05.
actions being used with a congruent or incongruent grip being applied. We observed effects of task, object–action congruency, and an interaction of task and congruency within action and object-associated regions of the brain and also in regions that may coordinate effects of task and congruency. We consider the factors of attention, congruency, and their interaction in turn.

Manipulating attention through task demands activated foci respectively associated with the dorsal and ventral streams, despite the stimuli being identical across the two tasks. Relative to the object task, the action task was associated with activity in the action–observation network; overlapping the dorsal pathway. Specifically, the bilateral SMG, the bilateral IPS, the bilateral pre-CS (below threshold on the left), and the right IFG showed an increased response when participants attended to actions compared with when they attended to objects. Although many studies have shown that observing actions elicit responses in the action observation network (for reviews, Rizzolatti & Sinigaglia, 2010; Grafton, 2009; Decety, 1996), here we show that attending to the action being performed with a stimulus is important for eliciting responses in this network; when the actions were not attended, responses in this network were much reduced (Figure 2).

In contrast with the action task, the object task was associated with activity in the SmFG and, at a lower threshold, with activation of the left ITG. The (weaker) activity in the left ITG is potentially linked to object processing itself, given that this region forms part of the classic ventral visual stream subserving object recognition (Milner & Goodale, 1995). Recruitment of this region is consistent with the N-back task on objects relying in part on activation in the object recognition system. The reason for increased activity in the SmFG is less clear, and future research is needed to understand its role here—for example, is the SmFG recruited to inhibit action information to allow access for object processing? Overall attention to the objects had weaker and less reliable effects on neural responses than attending to the actions. We can only speculate on the reasons for this. First, we note that the object-oriented task was easier than the action task. Typically, task difficulty positively correlates with the extent of the neural responses. Second, the ITG responded also when object–action relationships were congruent independent of the task (Tables 2 and 3; Figure 2). This may suggest that object processing as opposed to action processing is affected not only by task instructions but also by congruency. For example, congruent action–object relations may automatically elicit object-related processing, irrespective of the attention condition. Further research is needed to clarify this effect.

Taken together, the results are in accordance with previously reported behavioral results (Yoon & Humphreys, 2005), suggesting that action and object decisions involve separate networks. Action decision relies on a direct action observation fronto-parietal network, including motor planning regions. In contrast, N-back object decisions at least partly recruit regions within the ventral temporal visual stream.

A matching between the action (and grasp) and the object (on congruent trials) increased the responses of both the action observation and object-associated networks. Congruent object–action stimuli generated larger responses (relative to incongruent stimuli) bilaterally in the IPL, the SmFG (most likely SMA), and the putamen including the globus pallidus, the STS, and (at a weaker threshold) the ITG. These results are in accordance with the previous literature. Activation in the IPL have been associated with responses to an action goal during action observation (e.g., Fogassi et al., 2005), the motor representation linked to familiar graspable tools (Vingerhoets, 2008), and coding the postural and spatial relationships between a hand and an object (Sunderland et al., 2011). The evidence for greater activation with congruent (familiar) action–object pairings could fit with each of these proposals, as the goal of an action is easier to compute in the congruent conditions, the motor response associated with the tool will be stronger and the postural and spatial relations of the hand to the specific object will be more familiar. We note that the effects in the IPL appears to be specific to the congruency of an action and an object and not to grasp typicality, as previous research has failed to observe effects of grip congruency in this region (Valyear & Culham, 2009). The reason for this conflicting result is unclear. One difference between the two experiments is the actual action that was executed upon the objects and as a consequence the intentions of the actor differed. In the current study, the action imitated

<table>
<thead>
<tr>
<th>Contrast</th>
<th>Thalamus</th>
<th>Cluster Size (Std)</th>
<th>Voxel Z</th>
<th>Voxel p (unc)</th>
<th>mean MNI (Std)</th>
<th>x</th>
<th>y</th>
<th>z</th>
</tr>
</thead>
<tbody>
<tr>
<td>Congruent object vs. Action</td>
<td>Left</td>
<td>49.4 (50.23)</td>
<td>2.3</td>
<td>.017</td>
<td>−14.4 (6.38)</td>
<td>−19.1 (6.6)</td>
<td>10.1 (5.23)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Right</td>
<td>59.5 (107.58)</td>
<td>2.31</td>
<td>.016</td>
<td>11.9 (5.83)</td>
<td>−19.9 (10.4)</td>
<td>5.8 (4.31)</td>
<td></td>
</tr>
<tr>
<td>Congruent action vs. object</td>
<td>Left</td>
<td>35.9 (54.01)</td>
<td>2.2</td>
<td>.02</td>
<td>−12.4 (5.67)</td>
<td>−20.7 (7.75)</td>
<td>3.4 (6.95)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Right</td>
<td>66.2 (120.81)</td>
<td>2.27</td>
<td>.02</td>
<td>13.5 (5.59)</td>
<td>−20.7 (9.43)</td>
<td>7.6 (5.92)</td>
<td></td>
</tr>
</tbody>
</table>

The table presents the averaged of the single subject analyses results observed in the thalamus: descriptive measures (cluster size, MNI coordinates) and statistical tests (Z and p values). The analyses were restricted to the thalamus using a mask generated with WFU Pickatlas.
the object’s functional use; in Valyear and Culham, the objects were grasped and then lifted in a usual or unusual way. It may be that the aIPS encode viable (congruent) actions, and in Valyear and Culham, action goals and grasping for lifting were always encoded as congruent (or the same action), whether the grasping was typical or atypical. This is in contrast with the current study in which the action and as a consequence the actor’s intention varied between the congruent and incongruent trials.

The SmFG (SMA) has been shown to respond to object–grip congruency (Bach, Peelen, & Tipper, 2010); this accords with our observation of the involvement of the SmFG (SMA) when there is a match between the action and the object. Similarly, the posterior STS region has been shown to be involved in action observation, for example, in visual processing of biological motion (for a review, Thompson & Parasuraman, 2012). Single-cell recording in the monkey shows that some neurons in the STS selectively respond when presented with actions executed on objects (Perrett et al., 1989). This area shows reduced activity when either the hand or the object is presented alone (Barraclough, Keith, Xiao, Oram, & Perrett, 2008). Similarly, congruent grip has been shown to modulate responses in regions that are in the vicinity of areas observed here (Valyear & Culham, 2009). In addition, the STS appear to be involved in understanding the intention of actions performed by others (Pelphrey, Morris, & McCarthy, 2004). Again intentions are likely to be easier to interpret when a congruent rather than an incongruent action is performed with an object. The congruency action–object effects that were observed in the left fusiform gyrus/ITG accords with congruent grip effects reported before (Valyear & Culham, 2009). The IPL, SmFG, ITG, and BG are implicated in associating motor responses to a given sensory stimulus (for reviews, Lazar & Vaadia, 2008; Halsband & Freund, 1993). This network has also been observed when participants are asked to imagine complex daily actions (such as eating a meal; Szameitat, Shen, & Sterr, 2007). These previous findings accords with the increased involvement of the network, observed here, for well-rehearsed and familiar actions made on appropriate objects (congruent trials), as opposed to novel actions made on the same objects (incongruent trials). The novel contribution of this study is that these effects were observed independent of the task demands (attend to the action or the object).

Interestingly, incongruent trials resulted in greater activation of the medial occipital gyrus and the calcarine sulcus compared with congruent trials. This pattern of results is most easily attributed to the greater visual processing being required to encode stimuli when the action and object are incongruent, with greater recruitment of early visual processes in this case. This interpretation accords with the predictive coding account of brain operations (e.g., Friston, 2005). Thus, this model predicts increases in early sensory processing for unfamiliar and unpredicted input, as was the case when the actions were incongruent with the objects.

Along with the main effects of task and congruence, a set of brain regions were activated according to the particular combination of task and congruency. Most notably, there was a double dissociation in the selective activation of the subnuclei of the thalamus for congruent relative to incongruent stimuli in particular tasks; this was observed at the group level and confirmed by a follow-up analyses at the individual subject level. In the VPL nucleus, when participants attended the actions, there was greater activation for congruent relative to incongruent stimuli. In the LD nucleus, there was increased activity for congruent over incongruent trials but only for the object task. The VPL is most connected with regions of the dorsal cortex (particular to motor and premotor regions). The LD is mostly connected with the ventral cortex (temporal, ventral pFC, posterior-occipital cortex). The results suggest that, when attention was directed to the property associated with different cortical regions, there was gating of activation to connected thalamic nuclei so that the VPL was recruited during the action task and the LD during the object task.

These data are consistent with thalamic activity being affected by task constraints (attention to the task-relevant features of the image). Prior work has indicated that the thalamus is sensitive to attention manipulation of specific features of a stimulus, such as the sensitivity of the pulvinar to stimulus location (Snow, Allen, Rafal, & Humphreys, 2009; O’Connor, Fukui, Pinsk, & Kastner, 2002; LaBerge & Buchsbaum, 1990). Here we show that responses in the thalamus are also sensitive to more complex attentional manipulations reflecting the goals of the task (attend to the action or the object). Furthermore, it is interesting that the task-based thalamic dissociations here matched the structural connectivity between these nuclei and the task-relevant cortical regions. Notably, the VPL nucleus was linked to action and connects to motor and PM whereas the LD nucleus, linked to attention to objects, connects strongly to occipito-temporal cortices.

Our second piece of evidence was that thalamic activity was affected by the congruence of object–action relations, although the LD and VPL connect predominantly into ventral and dorsal cortex. There is mounting evidence for cortically based visual and motor responses to objects that are seen being used correctly, relative to when they are used incorrectly, which may reflect the visuomotor familiarity of the action (Roberts & Humphreys, 2010). For example, there is greater visual activity in the LOC, and greater activity putatively related to motor intention in dorsal PM, when objects are grasped correctly relative to when they are grasped incorrectly (Humphreys et al., 2010). We propose that feedback from these regions to the thalamus interacts with the task-based selection of specific thalamic nuclei, enhancing activity in the LD nucleus when the action is congruent with the relevant object task and enhancing activity in the VPL when the object is congruent with the relevant action task. This interaction, then, generates activity reflecting both congruence between objects and actions and task-based recruitment of thalamic nuclei. In this case, the selective
activation of thalamic nuclei provides a mechanism by which task-based constraints are integrated with signals reflecting stimulus congruency in the image.

This study was set up to test whether object action association is supported by a direct route to motor regions or an indirect route via semantic knowledge. We first want to say that our results are based on correlation analyses; hence, any interpretation about the necessary structures that supports a specific cognitive procedure is limited and speculative by nature. With that in mind, we believe that the congruency effects observed in both the action and object-related networks suggest that both networks represent objects with their corresponding actions. Because of the correlational nature of this study, we cannot comment whether both networks are essential for accurate execution of an action on an object, but we can say that each contains relevant information for this. Thus, it is plausible that processing in the dorsal and motor-associative cortex is sufficient for executing a correct action without relying on information from the ventral visual stream, as suggested by neuropsychological dissociations (see Introduction). On the other hand, we note that the anterior temporal pole, traditionally associated with semantic knowledge, was associated with stimulus congruency only when attention was directed to the object (Table 3). This suggests that semantic knowledge of an object includes action knowledge, but this information is accessed only when it is task relevant. Therefore, it may also be possible to guide a correct action based on knowledge represented in the ventral visual stream, but this process may depend on the involvement of attention.

In summary, the present data show that there is selective processing of action- and object-related properties of the same image in action- and object-related regions of cortex, according to the type of information weighted in the tasks, along with increased activity in different regions for congruent over incongruent action–object pairs. In addition, there is evidence that different thalamic subnuclei are recruited by the contrasting tasks, although their activity is then modulated by action–object congruence. The data suggest that the activity of thalamic nuclei is driven by the task as well as by differential stimulus signals reflecting object–action congruence.

Acknowledgments

This work was supported by grants from the BBSRC (G. W. H.), the Leverhulme Trust (P. R.), and European Union FP7 (Cogwatch, 288912).

Reprint requests should be sent to Eun Young Yoon or Pia Rotshtein, School of Psychology, University of Birmingham, Edgbaston, Birmingham B15 2TT, United Kingdom, or via e-mail: knic@daum.net, P.Rotshtein@bham.ac.uk.

REFERENCES


