Exp Brain Res
DOI 10.1007/s00221-013-3659-1

RESEARCH ARTICLE

Supramodal and modality-sensitive representations of perceived action categories in the human brain

Richard Ramsey · Emily S. Cross · Antonia F. de C. Hamilton

Received: 5 February 2013 / Accepted: 24 July 2013

Abstract Seeing Suzie bite an apple or reading the sentence 'Suzie munched the apple' both convey a similar idea. But is there a common neural basis for action comprehension when generated through video or text? The current study used functional magnetic resonance imaging to address this question. Participants observed videos or read sentences that described two categories of actions: eating and cleaning. A conjunction analysis of video and sentence stimuli revealed that cleaning actions (compared to eating actions) showed a greater response in dorsal frontoparietal regions, as well as within the medial fusiform gyrus. These findings reveal supramodal representations of perceived actions in the human brain, which are specific to action categories and independent of input modality (video or written words). In addition, some brain regions associated with cleaning and eating actions showed an interaction with modality, which was manifested as a greater sensitivity for video compared with sentence stimuli. Together, this pattern of results demonstrates both supramodal and modality-sensitive representations of action categories in the human brain, a finding with implications for how we understand other people's actions from video and written sources.

Keywords Language · Reading · Motor system · Action observation · Social cognition

Electronic supplementary material The online version of this article (doi:10.1007/s00221-013-3659-1) contains supplementary material, which is available to authorized users.

R. Ramsey (✉) · E. S. Cross
Wales Institute of Cognitive Neuroscience, School of Psychology, Bangor University, Bangor, Gwynedd, UK
e-mail: r.ramsey@bangor.ac.uk

E. S. Cross
Behavioural Science Institute, Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, Nijmegen, The Netherlands

A. F. de C. Hamilton
School of Psychology, University of Nottingham, University Park, Nottingham, UK

A. F. de C. Hamilton
Institute of Cognitive Neuroscience, University College London, London, UK

Introduction

Comprehension of other people's actions provides a basic input for social learning. For example, one could learn that in some cultures, locusts are considered edible by either watching a documentary or reading an article about eating locusts. As such, written words and video can convey the same socially relevant information. Recently, there has been an increased interest in studying the neural processes underpinning visual and linguistic action comprehension (Aziz-Zadeh and Damasio 2008). Here, we used fMRI to identify brain regions that responded to specific categories of actions, such as eating or cleaning, independent of the input modality through which the actions were perceived (video or written words).

Previous human neuroimaging studies have demonstrated that similar brain systems process aspects of conceptual knowledge independent of perceptual modality. For example, object features, such as colour, form and motion, are processed by similar parts of visual cortex, independent of whether the stimuli are written words, spoken words or pictures (Martin 2007). Reading words or seeing pictures related to food engages brain regions associated with reward, taste and pleasure, such as prefrontal cortex and

Published online: 21 August 2013

ventral striatum (Killgore et al. 2003; Simmons et al. 2005; Goldberg et al. 2006a, b; Barrós-Loscertales et al. 2011). Similarly, the comprehension of sentences has been shown to rely on frontal and temporal cortical regions, independent of visual or auditory input modality (Braze et al. 2011). Together, these findings suggest that some aspects of knowledge are represented in a supramodal manner and can be accessed through multiple modes of input.

In a similar vein, neuroimaging studies provide evidence suggesting that actions can be represented independent of input modality. Specifically, the representation of action has been shown to be organised in a somatotopic manner for observed actions (Buccino et al. 2001), action words (Hauk et al. 2004) and sentences that describe actions (Aziz-Zadeh et al. 2006; Buccino et al. 2005). For example, Aziz-Zadeh et al. (2006) showed participant's videos of actions performed with the hand, foot or mouth, as well as sentences describing similar actions. The blood oxygen level-dependent (BOLD) responses for video and sentence stimuli demonstrated overlap in an effector-specific manner, which suggested that the effector used to perform actions is represented in an input modality-independent manner. In addition, inferior frontal and parietal brain areas have been shown to respond to both video and sentences describing hands' actions (Baumgaertner et al. 2007). These previous action perception studies have focussed on a kinematic representation of action, which is tied to the limb used to perform that action. In addition, region of interest analyses have focussed on sensorimotor brain regions and the putative human mirror neuron system (MNS), which includes inferior frontal and parietal cortices (Hauk et al. 2004; Aziz-Zadeh et al. 2006). However, when action comprehension is situated in more realistic social contexts, many more action features can be informative to understanding another person's behaviour, such as the type of object acted upon and what the agent intends to achieve. As such, supramodal responses may extend beyond the human MNS (Peelen et al. 2010; Spunt and Lieberman 2012). Additionally, not all processes engaged during action comprehension may be supramodal; some brain regions may show a preference for the mode of input (Spunt and Lieberman 2012). To date, however, interactions between input modality and action features have been underexplored.

Here, we tested whether actions from different categories are represented independently from the modality through which they are perceived as well as which brain regions show a preference for the mode of input through which action information is gained. We deliberately chose to contrast two action categories that varied across a number of characteristics, in order to increase differences between them and make the categories as distinct as possible. Thus, we defined two broad action categories: eating and cleaning. Actions differed in terms of the direction of motion (toward vs. away from one's body), object (food vs. non-food), goal/intention (eat vs. clean), reward (gustatory reward vs. no reward) and body parts (hand and mouth vs. hand). Videos and sentences describing eating and cleaning actions were shown to participants (Fig. 1a), and we directly compared BOLD responses to actions from these different categories (Fig. 1b).

Two predictions can be distinguished. Brain regions responding in a category-specific manner to both video and sentence stimuli, as revealed in a conjunction analysis, would support a supramodal representation of action category information. In other words, these brain regions would show the same preference for action category independent of the input modality through which this information is perceived. We hypothesise that supramodal responses may be observed in a circuit of brain regions encompassing frontal, parietal and temporal brain regions, which have been collectively termed the action observation network (Cross et al. 2009) and have previously been shown to respond to video and sentence depictions of actions (e.g. Aziz-Zadeh et al. 2006). In contrast, statistical interactions between action category (eat vs. clean) and modality (video vs. sentence) would support modality-sensitive action category representations; action category sensitivity would differ between video and sentence stimuli (Fig. 1b). Brain regions showing an interaction might be specifically tuned to a particular input modality and thus demonstrate a greater sensitivity to action category information received through video than written words and vice versa. For example, inferior frontal and middle temporal cortices could be more sensitive to reading written words than video as these regions are consistently engaged when reading sentences (Vigneau et al. 2006, 2011). In contrast, brain regions in occipitotemporal cortex could show greater sensitivity for video than sentence stimuli as these regions are engaged during the perception of human form and motion (Hein and Knight 2008; Downing and Peelen 2011). Thus, our study will allow us to localise brain systems that demonstrate supramodal as well as modality-sensitive responses to action categories.

Materials and methods

Twenty-four participants gave informed consent to take part in the study in accordance with the local ethics committee at the Sir Peter Mansfield Magnetic Resonance Imaging Centre, University of Nottingham. One participant was removed from further analysis due to excessive head movement. The final sample of 23 comprised 13 females (mean age = 21.6). During fMRI scanning, participants observed silent videos of a female actor performing simple object-directed actions or read sentences describing similar action

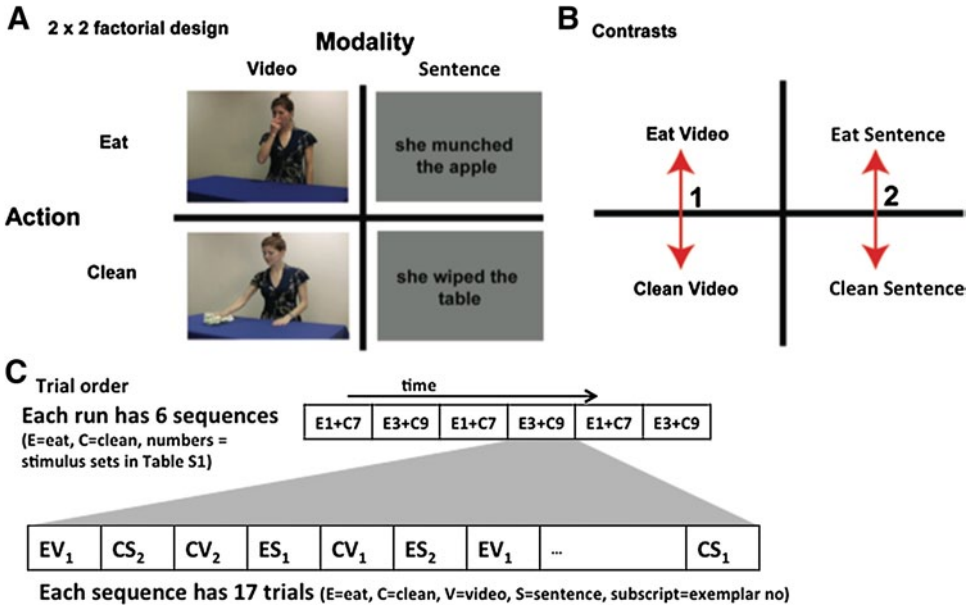

Fig. 1 Experimental design and contrasts. **a** Participants observed videos of an actor performing simple object-directed actions or read sentences describing similar action scenarios. The actions were from two categories: eating and cleaning. Thus, stimuli filled a 2 (action category: eat and clean) × 2 (modality: video and sentence) factorial design. **b** The simple effects of video (1) and sentence stimuli (2) were first calculated separately. Subsequently, through conjunction analyses, we tested for supramodal representations of action category, that is, brain regions that show the same preference for contrasts 1 and 2. Furthermore, through statistical interactions, we tested for modality-sensitive representations of action category, that is, brain regions that show different responses for contrasts 1 and 2. **c** Functional runs contained 6 sequences of 17 trials with sequences alternating between two different pairs of eating and cleaning actions. For example, in one functional run, odd sequences may involve apple eating and ball tossing stimuli, whereas even sequences may involve beer drinking and block stacking stimuli. In subsequent functional runs, different pairs of eating and cleaning action stimuli were used such that across the three runs, participants experienced all 12 action stimuli (see Supplementary Table S1). While the same pairs of actions were used across participants, the order that these pairs of stimuli were presented was randomised. Within each sequence, two video exemplars and two sentence exemplars of the eating and cleaning actions were shown in a pseudorandom order with the constraint that the identical trial was never shown twice in a row

scenarios. The actions were from two categories: eating and cleaning. For video stimuli, the actor performed 12 actions, 6 involving eating/drinking (e.g. eating an apple) and 6 involving cleaning/tidying (e.g. wiping a table). Two exemplars of each action were filmed, which had different low-level stimulus features without changing the meaning of the observed action. To do so, movies were filmed from two perspectives (45° from straight), and each perspective used a different exemplar of the same object. For example, one 'eat apple' movie involved eating a green apple from a left perspective, whereas a second 'eat apple' movie involved eating a red apple from a right perspective (Supplementary Table S1). This produced 24 movies in total, each of which was 2 s long and 640 pixels wide by 480 pixels high. Sentences described the same 12 action scenarios. For the sentences, two exemplars were generated that varied the verb while keeping the same general meaning of the sentence. Verbs were matched for frequency and imagability using the MRC Psycholinguisitc Database 2.0 (http://www.psych.rl.ac.uk/MRC_Psych_Db.html). Examples of sentences in the eating condition include 'she munched the apple' and 'she gnawed the apple', whereas cleaning examples include 'she wiped the surface' and 'she dusted the surface' (Supplementary Table S1). All stimuli were presented with Cogent running under Matlab 6.5.

Three functional runs were performed, each containing 6 sets of 17 trials (102 trials per functional run). Randomisation of stimuli was carried out at two levels: at the set level (choice of action categories) and at the exemplar level (choice of the individual stimulus). One set of actions comprises one cleaning action plus one eating action (i.e. two rows from Supplementary Table S1). Within one functional run, sets were alternated. For example, in one functional run, odd sets may involve apple eating and ball tossing stimuli, whereas even sets may involve beer drinking and block stacking stimuli (see Fig. 1c). In subsequent functional runs, different sets of eating and cleaning action stimuli were used such that across the three runs, participants experienced all 12 action categories (for a complete list of categories used, see Supplementary Table S1). The order of the action sets was randomised across runs and across participants.

The second level of randomisation concerns the presentation of the exemplars within each set. There are 8 exemplars to choose from (2 eating videos, 2 eating text, 2 cleaning videos and 2 cleaning text), and these were shown in a pseudorandom order with the constraint that the same exemplar was not shown in succession. The number of times each exemplar was shown varied slightly across sequences and participants, with an average of 76.5 (SD = 6.65, minimum 57) trials contributing to each experimental condition in the analysis. This unusual randomisation enabled an alternative data analysis, which is not presented here. For the present analysis, two critical conditions are met: trials in our contrasts of interest can be analysed independently and without order effects, and the number of trials that contributed to each of our four experimental conditions was sufficient to perform our analyses.

Stimuli were presented for 2 s and separated by a blank screen (1 s). There was no jittering of trials, but the full randomisation of stimulus order still allows us to separate out the BOLD response to each item in our design. Following each sequence, participants answered a question about the content of the last trial and were then able to rest (10 s). The content of questions could not be predicted, and these were included to maintain alertness. Six questions were used in total, which required a yes/no response (e.g. Did she eat or drink? Did she move the object away from her? Did she like the object?). Responses were made using an index finger key press on one of the two button boxes; left finger responses corresponded to 'yes' and right finger responses to 'no'.

Scanning was performed in a 3T Phillips Achieva scanner using an 8 channel-phased array head coil with 38 slices per TR (3 mm thickness); TR: 2,500 ms; TE: 40 ms; flip angle: 80°; FOV: 24 cm, matrix: 80 × 80. One hundred and fifty-five brain images were stored on each functional run. Data were realigned, unwarped, corrected for slice timing, normalised to the MNI template with a resolution of 3 × 3 × 3 mm and spatially smoothed (8 mm) using SPM8 software. A design matrix was fitted for each participant with regressors for each of the five stimulus types (EV, ES, CV, CS and Question; E = Eat, C = Clean, V = Video, S = Sentence). Each trial was modelled as a boxcar with the duration of that movie convolved with the standard hemodynamic response function.

Stimuli filled a 2 (action category: eat and clean) × 2 (modality: video and sentence) factorial design (Fig. 1a). First, we examined four simple effects within the factorial design separately. This involved two action category contrasts for video stimuli (EV > CV and CV > EV) and two for sentence stimuli (ES > CS and CS > ES) (Fig. 1b). Contrast images for all participants were taken to the second level for random effects analysis. Correction for multiple comparisons was performed at the cluster level (Friston et al. 1994), using a voxel-level threshold of $p < 0.005$ and 10 voxels and a family-wise error (FWE) cluster-level correction of $p < 0.05$. For each contrast, brain regions surviving the voxel-level threshold ($p < 0.005$ and 10 voxels) are reported in tables with regions surviving the FWE cluster correction highlighted in bold font. In Figs. 2 and 3, results are illustrated on a template brain, with plots showing parameter estimates (SPM betas) for each condition of interest within our design matrix. Brain regions were localised using a human brain atlas (Duvernoy 1999) in combination with an online search tool (http://sumsdb.wustl.edu/sums/) and the SPM Anatomy toolbox (Eickhoff et al. 2005).

Subsequently, we tested for supramodal representations of action category, that is, brain regions that show the same preference for action category independent of the input modality through which they were perceived. To do so, two conjunction analyses of video and sentence stimuli were performed: (1) Clean > Eat [(CV > EV) ∩ (CS > ES)] and (2) Eat > Clean [(EV > CV) ∩ (ES > CS)]. Conjunction analyses were performed by visualising independent contrasts on the same brain template using our voxel-level threshold, $p < 0.005$ and 10 voxels. If a cluster shows visible overlap between independent contrasts, then the voxels showing overlap have the same pattern in all contrasts at the designated threshold. Thus, each conjunction involved two statistically independent contrasts. This conjunction method requires that both comparisons in the conjunction are individually significant (Nichols et al. 2005), and means that brain regions identified in the conjunction show the same pattern of response to category information for both video and sentence stimuli. For example, brain regions emerging from the first conjunction analysis would demonstrate stronger responses to cleaning actions compared with eating actions for both video and sentence stimuli. Likewise, brain regions emerging from the second conjunction analysis would demonstrate stronger responses to eating actions compared with cleaning actions for both video and sentence stimuli.

In addition to the conjunction analyses, we also tested for brain regions that were sensitive to the modality of presentation. We accomplished this by calculating statistical interactions between action category (eat vs. clean) and modality (video vs. sentence). Brain regions showing a significant interaction between action category and modality would suggest that action category preference is different between video and sentence stimuli or is at least modulated by modality. The first interaction tested the relative activation of Clean X Video [(CV > EV) > (CS > ES)]. The second interaction tested the relative activation of Eat X Video [(EV > CV) > (ES > CS)]. Brain regions showing the first interaction would demonstrate a greater sensitivity to Clean > Eat actions when perceived through video compared with a sentence. Brain regions showing the

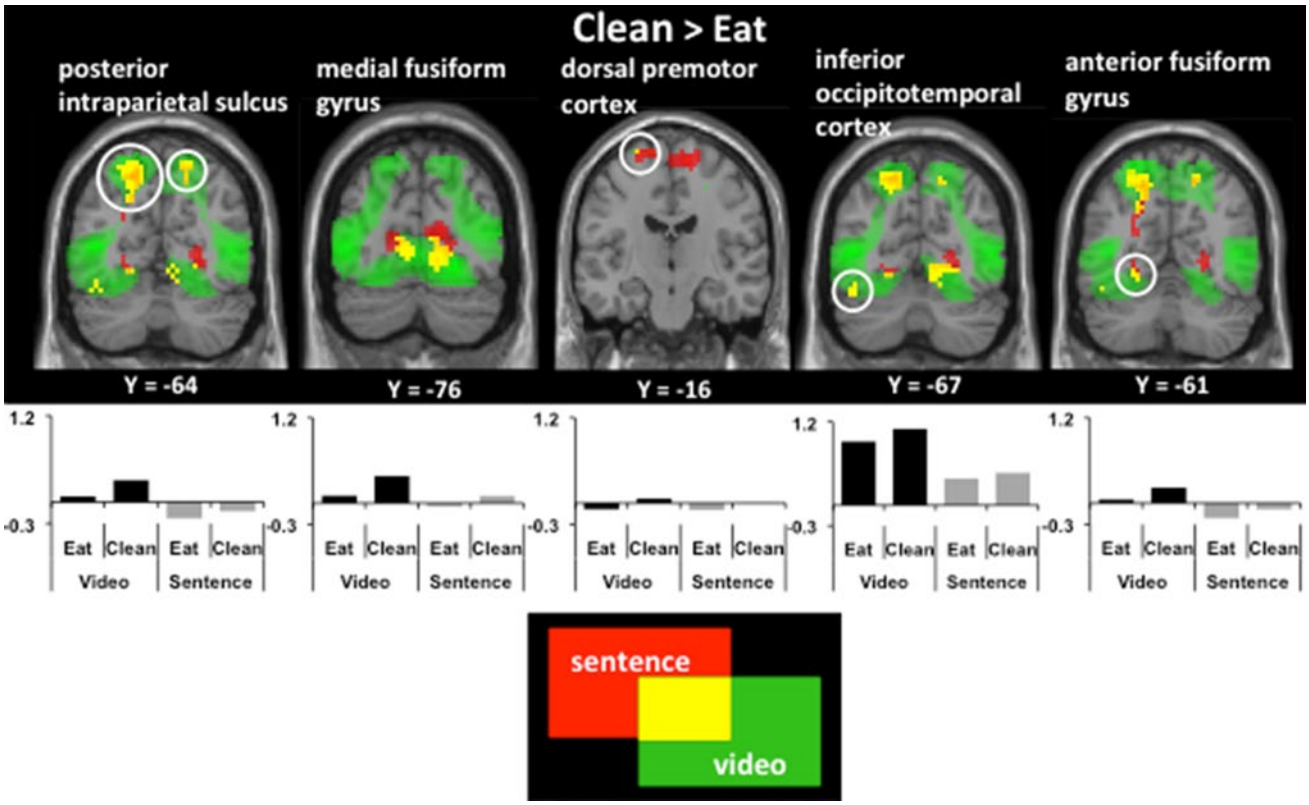

Fig. 2 Brain regions showing the conjunction of video and sentence stimuli. Significantly greater activity was seen for cleaning actions compared with eating actions for video (*black bars*) and sentence stimuli (*grey bars*) in posterior intraparietal sulcus, medial fusiform gyrus, dorsal premotor cortex, inferior occipitotemporal cortex and anterior fusiform gyrus. Responses to video stimuli are displayed in *green*, sentence in *red* and overlap in *yellow* (whole-brain threshold: $p < 0.005$, $K = 10$). Parameter estimates (SPM betas) are plotted for each region (*note*: left hemisphere responses are plotted when activity is bilateral)

second interaction would demonstrate a greater sensitivity to Eat > Clean actions when perceived through video compared with a sentence. Such patterns of BOLD response would suggest that action category information is not entirely supramodal or amodal, but rather the representation of action categories is dependent, in part, on the perceptual mode of input.

Results

Simple effects and conjunctions

We first consider brain regions more sensitive to cleaning than eating actions. A number of brain regions showed the simple effect of cleaning within the video modality (CV > EV), including dorsal premotor cortex (PMd), posterior intraparietal sulcus (pIPS), fusiform gyrus, lateral occipitotemporal cortex (OT) and thalamus (Table 1). A similar set of brain regions showed the simple effect within the sentence modality (CS > ES), including PMd, pIPS, fusiform gyrus and lateral OT (Table 1). In addition, several other brain regions were also sensitive to CS > ES, including left middle frontal gyrus, dorsomedial prefrontal cortex, inferior frontal gyrus and inferior parietal lobule. The conjunction analysis of Clean > Eat for both video and sentence stimuli identified this pattern in bilateral pIPS and bilateral medial fusiform gyrus, as well as left PMd and left lateral OT (Table 1; Fig. 2).

Second, we consider brain regions more sensitive to eating than cleaning actions. A number of brain regions showed the simple effect of eating within the video modality (EV > CV), including prefrontal and cingulate cortices as well as midbrain and occipital regions (Table 2). No brain regions showed the simple effect within the sentence modality (ES > CS). The conjunction analysis of Eat > Clean for both video and sentence stimuli was not possible because the sentence contrast did not produce any suprathreshold clusters.

Interaction analyses

Statistical interactions between action category (eat vs. clean) and modality (video vs. sentence) were calculated. The

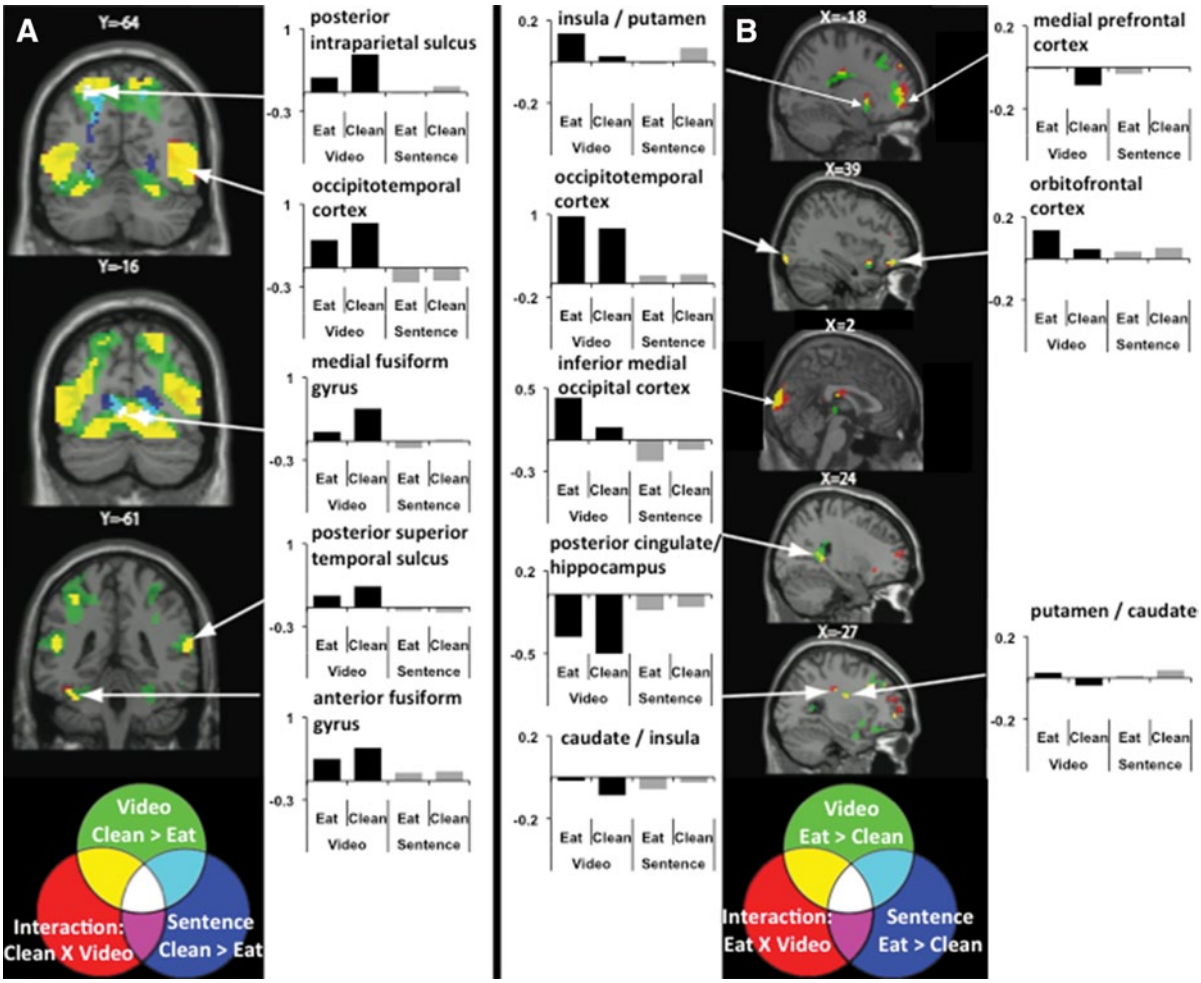

Fig. 3 Brain regions showing statistical interactions between action category (eat and clean) and modality (video and sentence). **a** The first interaction tested the relative activation of Clean X Video [(CV > EV) > (CS > ES)] and revealed activity in occipitotemporal cortex, fusiform gyrus, posterior intraparietal sulcus (pIPS) and posterior superior temporal sulcus. Each of these brain regions overlapped with the simple effect of CV > EV (*yellow*), which suggests that the first interaction is driven by the significant simple effect of CV > EV. This pattern of results is illustrated by the parameter estimates (SPM betas): the difference between cleaning and eating actions is greater for video stimuli (*black bars*) than for sentence stimuli (*grey bars*). In addition, two of these regions, pIPS and medial fusiform gyrus, also overlapped with the simple effect of CT > ET (*white*). This suggests that pIPS and medial fusiform gyrus respond to C > E for both video and textual modalities, but also show a relatively greater preference for video compared to sentence stimuli. The parameter estimates for pIPS and medial fusiform gyrus illustrate this pattern of results. **b** The second interaction tested the relative activation for Eat X Video [(EV > CV) > (ES > CS)] and revealed widespread responses in prefrontal, cingulate and occipital cortices as well as several midbrain regions (*red*). Many of these regions overlapped with the simple effect of EV > CV (*yellow*), but not with any other contrast. This result suggests that the second interaction is driven by the significant simple effect of EV > CV. This pattern of results is illustrated by the parameter estimates (SPM betas): the difference between eating and cleaning actions is greater for video stimuli (*black bars*) than for sentence stimuli (*grey bars*). Contrasts are displayed using the same whole-brain threshold: $p < 0.005$, $K = 10$

first analysis, which tested the interaction of Clean X Video [(CV > EV) > (CT > ET)], revealed activity in OT, fusiform gyrus, pIPS and pSTS (Table 3). This demonstrates that the difference in BOLD response between cleaning and eating actions is greater for video stimuli than for sentence stimuli in these brain areas (Fig. 3a). Each of these brain regions overlapped with the simple effect of CV > EV (Fig. 3a; yellow). This means that the interaction (Clean X Video) is driven by the significant simple effect of CV > EV. Furthermore, pIPS and medial fusiform gyrus showed a three-way overlap, which included the simple effect of Clean > Eat for video (CV > EV), the simple effect of Clean > Eat for sentences (CT > ET) and the interaction between Clean X Video. This means that pIPS and medial fusiform gyrus respond to Clean > Eat for both video and sentence modalities, but that this selectivity is stronger for video stimuli. The parameter estimates (SPM betas) for pIPS and medial fusiform gyrus illustrate this pattern of results (Fig. 3a).

Table 1 Brain regions showing Clean > Eat for video and sentence stimuli separately

Region	Number of voxels	*T*	Montreal Neurological Institute coordinates		
			x	*y*	*z*
Clean_Video > *Eat_Video*					
Right posterior fusiform gyrus extending into left fusiform gyrus and occipitotemporal cortex	**6,959**	**9.67**	**24**	**−79**	**−8**
			−45	**−67**	**7**
			−42	**−76**	**1**
Thalamus	34	6.01	18	−28	4
Left dorsal premotor cortex	**239**	**4.66**	**−36**	**−7**	**36**
			−24	**−7**	**64**
			−24	**−13**	**73**
Right dorsal premotor cortex	26	3.77	27	−13	52
			24	−10	61
Right inferior occipitotemporal cortex	11	3.33	54	−46	−20
Clean_Sentence > *Eat_Sentence*					
Left posterior intraparietal sulcus	130	5.37	−18	−64	58
			−21	−61	37
			−24	−61	22
Left dorsal premotor cortex extending into right dorsal premotor cortex	**280**	**4.99**	**−12**	**−16**	**73**
			12	**−16**	**67**
			−9	**−4**	**73**
Inferior medial occipital extending into right fusiform gyrus	**673**	**4.70**	**3**	**−88**	**16**
			15	**−79**	**13**
			−9	**−88**	**13**
Right inferior frontal gyrus/insula	26	4.07	27	23	−8
Left dorsomedial prefrontal cortex	13	4.02	−9	8	58
Left anterior fusiform gyrus	31	3.91	−24	−58	−2
			−21	−67	−5
Left middle frontal gyrus	15	3.52	−48	35	28
			−45	29	37
Right posterior intraparietal sulcus	17	3.42	18	−64	58
			18	−64	49
Left inferior parietal lobule	20	3.40	−60	−49	37
Left lateral fusiform gyrus/OT	11	3.25	−45	−67	−17
Right inferior parietal lobule	10	3.19	33	−31	34
Right inferior frontal gyrus/insula	10	3.12	48	2	−2

Only regions surviving a voxel-level threshold of $p < 0.005$ uncorrected and 10 voxels are reported. Subpeaks more than 8 mm from the main peak in each cluster are listed

Bold indicates regions that survive the family-wise error cluster-corrected threshold at $p < 0.05$

OT occipitotemporal cortex

The second interaction, which tested the relative activation for Eat X Video [(EV > CV) > (ES > CS)], revealed widespread responses in prefrontal, cingulate and occipital cortices, as well as several midbrain regions (Table 3). In these brain areas, the difference in BOLD response between eating and cleaning actions is greater for video stimuli than for sentence stimuli (Fig. 3b). Many of these regions overlapped with the simple effect of EV > CV (Fig. 3b; yellow), but not with any other effects. This means that the second interaction (Eat X Video) is driven by the significant simple effect of EV > CV. This pattern of results is illustrated by the parameter estimates (SPM betas) shown in Fig. 3b.

Discussion

Our results reveal both supramodal and modality-sensitive responses to action information across the human brain. For both video and sentence stimuli, PMd, pIPS, medial fusiform gyrus and OT show a greater response for cleaning

Table 2 Brain regions showing Eat > Clean for video and sentence stimuli separately

Region	Number of voxels	*T*	Montreal Neurological Institute coordinates		
			x	*y*	*z*
Eat_Video > Clean_Video					
Right posterior cingulate	140	5.23	21	−46	10
			30	−52	10
			24	−34	28
Left middle cingulate/caudate	139	5.16	−18	−25	31
			−30	−7	25
			−18	−34	22
Medial inferior occipital gyrus	127	4.96	6	−100	13
			−6	−103	7
Right lateral occipitotemporal cortex	26	4.95	36	−94	−11
			27	−97	−5
Hippocampus/caudate	46	4.67	12	−1	22
			9	−16	25
			0	−22	19
Right lateral orbitofrontal cortex	21	4.62	36	41	−14
			27	35	−17
Medial prefrontal cortex extending into medial orbitofrontal cortex	**194**	**4.52**	**−15**	**47**	**−2**
			−18	**44**	**10**
			−6	**32**	**−8**
Thalamus	32	4.44	0	−22	1
			−6	−7	4
Middle cingulate extending into hippocampus/caudate	30	4.20	−9	−1	19
			−18	14	19
Left lateral prefrontal cortex	119	4.13	−24	23	37
			−15	20	43
			−24	35	40
Medial prefrontal cortex	21	4.12	−15	50	43
Right precentral cortex	19	4.05	30	−22	43
Left amygdala	25	3.97	−30	5	−23
Left ventral striatum	50	3.84	−24	11	−11
Right temporal pole	20	3.84	36	11	−17
Left orbitofrontal cortex	29	3.71	−24	32	−17
			−18	38	−14
Left posterior cingulate	11	3.27	−24	−52	10
Medial prefrontal cortex	11	3.09	6	56	37
Eat_Sentence > Clean_Sentence					
No suprathreshold clusters					

Only regions surviving a voxel-level threshold of $p < 0.005$ uncorrected and 10 voxels are reported. Subpeaks more than 8 mm from the main peak in each cluster are listed

Bold indicates regions that survive the family-wise error cluster-corrected threshold at $p < 0.05$

than eating actions. These regions show a category-specific but input modality-independent representation of other people's action. In contrast, several brain regions showed a bias towards video stimuli compared with sentences. Below we discuss the implications of these findings in relation to previous literature and suggest new directions for future research.

Supramodal representations of perceived action categories

The present findings demonstrate supramodal neural responses during the comprehension of actions perceived through video and sentences. Previous neuroimaging research has shown an effector-specific action

Table 3 Brain regions revealed by the interaction of action (eat vs. clean) and modality (video vs. sentence)

Region	Number of voxels	*T*	Montreal Neurological Institute coordinates		
			x	*y*	*z*
Interaction 1—Clean X Video [(Clean_Video > Eat_Video) > (Clean_Sentence > Eat_Sentence)]					
Right lateral occipitotemporal cortex	**965**	**8.74**	**51**	**−64**	**4**
			42	**−58**	**7**
			39	**−76**	**10**
Left lateral occipitotemporal cortex	**892**	**8.14**	**−45**	**−79**	**1**
			−39	**−64**	**7**
			−39	**−88**	**10**
Right posterior fusiform gyrus extending into left fusiform gyrus	**717**	**6.68**	**18**	**−82**	**−17**
			15	**−88**	**−11**
			−30	**−64**	**−17**
Left posterior intraparietal sulcus	**361**	**5.54**	**−21**	**−58**	**67**
			−15	**−58**	**61**
			−36	**−46**	**61**
Left anterior fusiform gyrus	22	4.20	−36	−40	−29
			−39	−40	−20
Left posterior superior temporal sulcus	83	4.02	−57	−34	16
			−51	−40	13
Right posterior superior temporal sulcus	49	3.64	66	−40	16
			57	−37	10
Interaction 2—Eat X Video [(Eat_Video > Clean_Video) > (Eat_Sentence > Clean_Sentence)]					
Left caudate	45	6.35	−21	−22	34
			−33	−22	25
Inferior medial occipital cortex	**269**	**5.60**	**−3**	**−100**	**10**
			6	**−97**	**22**
			6	**−91**	**31**
Right occipitotemporal cortex	20	4.86	36	−94	−11
Right lateral orbitofrontal cortex	21	4.86	39	41	−14
Thalamus	45	4.66	3	−16	22
			−9	−4	19
Left lateral prefrontal cortex	**170**	**4.57**	**−18**	**50**	**10**
			−18	**50**	**−5**
			−27	**59**	**7**
Left putamen/hippocampus	17	4.16	−30	−7	22
Medial prefrontal cortex	13	3.96	15	47	−2
Right prefrontal cortex	31	3.95	24	53	13
			21	65	13
Right posterior cingulate cortex	44	3.92	15	−46	10
			24	−49	13
Thalamus	17	3.92	−3	−7	7
Left anterior cingulate cortex	25	3.79	−27	20	34
			−39	29	31
Left insula/putamen	43	3.67	−24	11	−8
			−18	5	1
Right anterior insula	11	3.65	24	26	−8
Left lateral prefrontal cortex	39	3.40	−24	44	34
			−18	50	40

Table 3 continued

Region	Number of voxels	*T*	Montreal Neurological Institute coordinates		
			x	*y*	*z*
Right lateral prefrontal cortex/inferior frontal gyrus	19	3.40	39	38	16
			45	44	10
Medial superior frontal gyrus	12	3.34	−3	32	58
			−9	38	55
Right insula	12	3.31	36	11	−14

Only regions surviving a voxel-level threshold of $p < 0.005$ uncorrected and 10 voxels are reported. Subpeaks more than 8 mm from the main peak in each cluster are listed

Bold indicates regions that survive the family-wise error cluster-corrected threshold at $p < 0.05$

representation when perceiving actions through video, reading text or hearing sentences (Buccino et al. 2001, 2005; Aziz-Zadeh et al. 2006; Hauk et al. 2004). Across these studies, somatotopic mapping of hand, foot and mouth actions was found in premotor, primary motor and inferior parietal cortices independent of the mode of input. We extend these results by showing that when action categories are made more distinct, a distributed set of brain regions, which are not restricted to sensorimotor cortical areas, show supramodal responses. More specifically, PMd, pIPS, medial fusiform gyrus and OT were sensitive to cleaning compared with eating actions for both video and sentence stimuli. Such supramodal responses are consistent with a recent fMRI study that reported supramodal responses to action features that are not effector-specific (Spunt and Lieberman 2012). In their study, Spunt and Lieberman (2012) showed that for both video and text stimuli, considering 'how' actions were performed engaged a premotor–parietal brain circuit, whereas considering 'why' actions were performed engaged medial prefrontal, posterior cingulate and temporoparietal cortical regions. Together, these studies suggest that despite distinctly different visual properties, video and sentence stimuli provide access to a common representation of action information. More broadly, this result is consistent with recent reports in non-action domains, such as the perception of emotions (Peelen et al. 2010), appetitive stimuli (Simmons et al. 2005; Goldberg et al. 2006a, b; Killgore et al. 2003), objects (Martin 2007) and sentence comprehension (Braze et al. 2011). In sum, our findings support the notion that similar brain systems process specific categories of knowledge independent of the modality through which they are perceived.

Our main focus in this study was to establish supramodal action representations. In our experimental design, we deliberately defined two distinct action categories that varied across a number of characteristics. The cleaning and eating actions vary in terms of the direction of motion, the object of the action, the actor's goal/intention, the gustatory reward and the body parts primarily used. We do not attempt to distinguish which of these factors drives our results, because further research is needed to separate the influence of each factor. Rather, we consider how our findings relate to previous studies of these different action characteristics.

One possible factor driving different BOLD responses to stimuli featuring cleaning and eating is the perceived or implied direction of motion. Eating stimuli involved actions directed toward the actor (away from the camera), while cleaning stimuli involved actions directed away from the actor (toward the camera). Widespread areas of visual cortex are sensitive to motion direction (Shipp and Zeki 1995; Andersen 1997; Kamitani and Tong 2006; Brewer et al. 2005), and specifically optic flow towards an observer (Wunderlich et al. 2002). Thus, the stronger responses in medial fusiform gyrus and OT to cleaning actions could be due to the direction of motion of the actor's hand (and the object) moving towards the observer, and thus becoming larger in their visual field, during these stimuli. Similarly, a premotor–parietal circuit has been shown to respond to objects approaching one's body in monkeys and humans (Graziano and Cooke 2006; Makin et al. 2008). Findings in our study are consistent with this idea. PMd and pIPS responded more to the perception of cleaning actions, which moved objects toward the observer, than eating actions, which moved objects away from the observer.

A second possible factor is the type of object associated with eating or cleaning (i.e. food objects vs. non-food objects). Compared to food objects, non-food objects have more functional and manipulable associations, similar to tools, and it is known that processing of tools engages premotor, parietal and temporal brain regions (Martin 2007; Beauchamp and Martin 2007). In contrast, food stimuli activate brain regions associated with gustatory reward, taste and pleasure, such as prefrontal and cingulate cortices as well as the ventral striatum (Simmons et al. 2005; Killgore et al. 2003; Goldberg et al. 2006a, b). For example, pictures of edible objects have been shown to activate

prefrontal and midbrain areas, whereas pictures of non-edible objects activate occipitotemporal areas (Killgore et al. 2003). Our results are consistent with these findings: for video stimuli only, observing eating compared with cleaning actions activated medial prefrontal and cingulate cortices, whereas for both video and sentence stimuli, observation of cleaning compared with eating actions activated PMd, pIPS, fusiform gyrus and OT. This could suggest that different actions are categorised in the brain according to the properties of the object acted upon or, in the case of reading a sentence, the action that a referenced object affords (Binkofski et al. 2004; Carota et al. 2012). We note that explanations of our data in terms of the direction of motion and in terms of object properties are not mutually exclusive. It is likely that both factors play a role, and further research would be needed to distinguish between them.

Modality-sensitive representations of perceived action categories

In addition to supramodal responses, several brain regions showed greater (action) category-selectivity for video stimuli than for sentence stimuli as revealed in the interaction analyses. Specifically, for cleaning compared with eating actions, there was a Clean X Video interaction and a simple effect of CV > EV in pIPS, fusiform gyrus, pSTS and OT. For eating compared with cleaning actions, there was a Eat X Video interaction and a simple effect of EV > CV in prefrontal, cingulate and occipital cortices as well as several midbrain regions. This means that these brain regions are more (action) category-selective when information is presented in video than when it is presented as a sentence. Two brain regions, pIPS and medial fusiform gyrus, showed greater responses for cleaning than eating actions in both sentence and video modalities, but also an interaction with stronger (action) category-selectivity for video stimuli. Together, this pattern of results suggests that brain responses that distinguish between categories of other people's actions are not always independent of the input modality; distinctions are often stronger for video stimuli.

There are two possible reasons for these modality-sensitive responses. First, these brain regions might be specifically tuned to a particular input modality, for example, receiving information from systems for action observation, but not for language. Second, these brain regions might receive both sources of information, but respond more to the video stimuli because these are more detailed and compelling than sentence stimuli, and thus engage attention and cognitive resources to a greater extent. The possibility that video stimuli are more engaging than text stimuli is a problem for any fMRI study comparing video and word stimuli (Aziz-Zadeh et al. 2006), not just the current study, and cannot easily be avoided. Multivoxel pattern analysis fMRI methods (Haynes and Rees 2006; Peelen et al. 2010) might provide a way to define how well brain regions select for different stimulus categories independent of absolute BOLD level, but this approach was not suitable for the current data set.

Broader implications

Our data have implications for current theories of how the brain responds to other people's actions in social contexts. Research on action perception from video stimuli over the last decade has focused largely on the human MNS, comprising areas of cortex mainly within inferior frontal gyrus and inferior parietal lobule (Rizzolatti and Craighero 2004; Gallese et al. 2004). These regions respond when participants observe other people's actions in situations that are often devoid of social context, such as observing a hand grasping an object (Buccino et al. 2001; Hamilton and Grafton 2006; Gazzola and Keysers 2009; Kilner et al. 2009; Caspers et al. 2010; Grèzes and Decety 2001). In contrast, when actions in video format are embedded in more complex social contexts, a broader brain network is engaged. This network includes regions associated with mental state reasoning and also with body part processing (de Lange et al. 2008; Ramsey and Hamilton 2010; 2012; Brass et al. 2007; Grèzes et al. 2004a, b; Spunt et al. 2011; Marsh and Hamilton 2011). The current study supports the idea that the comprehension of action categories engages brain regions that extend beyond the human MNS (Keysers and Gazzola 2007; Uddin et al. 2007; Adolphs 2009).

Similar conclusions can be drawn for studies of action comprehension from linguistic sources. There is increasing evidence that brain systems for language and action are linked (Aziz-Zadeh et al. 2006; Aziz-Zadeh and Damasio 2008; Pulvermuller 2005; Hauk et al. 2004; Rueschemeyer et al. 2010; Noppeney et al. 2005). In particular, reading about actions performed by specific body parts engages premotor and primary motor cortices in a similar way as observing or performing those actions (Aziz-Zadeh et al. 2006; Hauk et al. 2004). Thus, some authors have argued that the motor system supports semantic representations of words and sentences (Aziz-Zadeh and Damasio 2008; Gallese and Lakoff 2005; Boulenger et al. 2009; Glenberg et al. 2008). However, studies that examine the processing of action-text in broader contexts again find responses in a broader brain network. For example, when reading stories, brain activity in visual and motor cortices reflects that which would be expected if one was actually experiencing the situation rather than simply reading words (Speer et al. 2009). Furthermore, reading engages brain regions associated with mental state reasoning when a reader must consider why an action has been performed (Spunt et al. 2010) or must judge whether words could describe a person's

psychological state (Mitchell et al. 2002; 2005). Taken together, these findings suggest that the motor system alone is not sufficient, in all cases, to extract action information when reading sentences (Mahon and Caramazza 2008; Toni et al. 2008; Willems and Hagoort in press). Moreover, the findings are consistent with the claim that the motor system's role in language understanding is flexible and depends on the level of detail described in an action sentence (Willems and Hagoort in press).

Conclusion

Comprehension of other people's actions through different sources of input, such as observing their actions and reading textual descriptions, is an important part of social learning. Our results show that the perception of broad action categories, such as eating and cleaning, involve both supramodal and modality-sensitive neural representations. For both video and sentence stimuli, PMd, pIPS, medial fusiform gyrus and OT showed a stronger response for cleaning than eating actions. These supramodal responses demonstrate sensitivity to action–category information independent of the perceptual mode of input (i.e. video or sentence). In contrast, widespread modality-sensitive responses were shown across the cerebral cortex and demonstrate greater sensitivity to video than sentence stimuli. The results point towards a distributed neural representation of action category information, which extends beyond the human MNS. A challenge for future research is to delineate more precisely which aspects of action information are supramodal and which are modality-sensitive, as this will further inform theories of action comprehension and the neural basis of action perception.

References

Adolphs R (2009) The social brain: neural basis of social knowledge. Annu Rev Psychol 60(1):693–716. doi:10.1146/annurev.psych.60.110707.163514

Andersen RA (1997) Neural mechanisms of visual motion perception in primates. Neuron 18(6):865–872

Aziz-Zadeh L, Damasio A (2008) Embodied semantics for actions: findings from functional brain imaging. J Physiol Paris 102(1–3):35–39

Aziz-Zadeh L, Wilson SM, Rizzolatti G, Iacoboni M (2006) Congruent embodied representations for visually presented actions and linguistic phrases describing actions. Curr Biol 16(18):1818–1823

Barrós-Loscertales A, González J, Pulvermüller F, Ventura-Campos N, Bustamante JC, Costumero Vc, Parcet MaA, Ávila Cs (2011) Reading salt activates gustatory brain regions: fMRI evidence for semantic grounding in a novel sensory modality. Cereb Cortex. doi:10.1093/cercor/bhr324

Baumgaertner A, Buccino G, Lange R, McNamara A, Binkofski F (2007) Polymodal conceptual processing of human biological actions in the left inferior frontal lobe. Eur J Neurosci 25(3):881–889. doi:10.1111/j.1460-9568.2007.05346.x

Beauchamp MS, Martin A (2007) Grounding Object Concepts in Perception and Action: evidence from FMRI Studies of Tools. Cortex; a journal devoted to the study of the nervous system and behavior 43(3):461–468

Binkofski F, Buccino G, Zilles K, Fink GR (2004) Supramodal representation of objects and actions in the human inferior temporal and ventral premotor cortex. Cortex J Devot Study Nerv Syst Behav 40(1):159–161

Boulenger V, Hauk O, Pulvermuller F (2009) Grasping ideas with the motor system: semantic somatotopy in idiom comprehension. Cereb Cortex 19(8):1905–1914. doi:10.1093/cercor/bhn217

Brass M, Schmitt RM, Spengler S, Gergely G (2007) Investigating action understanding: inferential processes versus action simulation. Curr Biol 17(24):2117–2121

Braze D, Mencl WE, Tabor W, Pugh KR, Todd Constable R, Fulbright RK, Magnuson JS, Van Dyke JA, Shankweiler DP (2011) Unification of sentence processing via ear and eye: an fMRI study. Cortex J Devot Study Nerv Syst Behav 47(4):416–431

Brewer AA, Liu J, Wade AR, Wandell BA (2005) Visual field maps and stimulus selectivity in human ventral occipital cortex. Nat Neurosci 8(8):1102–1109

Buccino G, Binkofski F, Fink GR, Fadiga L, Fogassi L, Gallese V, Seitz RJ, Zilles K, Rizzolatti G, Freund HJ (2001) Action observation activates premotor and parietal areas in a somatotopic manner: an fMRI study. Eur J Neurosci 13(2):400–404

Buccino G, Riggio L, Melli G, Binkofski F, Gallese V, Rizzolatti G (2005) Listening to action-related sentences modulates the activity of the motor system: a combined TMS and behavioral study. Cognit Brain Res 24(3):355–363

Carota F, Moseley R, Pulverm√°ller F (2012) Body-part-specific representations of semantic noun categories. J Cognit Neurosci 24(6):1492–1509. doi:10.1162/jocn_a_00219

Caspers S, Zilles K, Laird AR, Eickhoff SB (2010) ALE meta-analysis of action observation and imitation in the human brain. NeuroImage 50(3):1148–1167

Cross ES, Kraemer DJM, Hamilton AFDC, Kelley WM, Grafton ST (2009) Sensitivity of the action observation network to physical and observational learning. Cereb Cortex 19(2):315–326

de Lange FP, Spronk M, Willems RM, Toni I, Bekkering H (2008) Complementary systems for understanding action intentions. Curr Biol 18(6):454–457

Downing PE, Peelen MV (2011) The role of occipitotemporal body-selective regions in person perception. Cogn Neurosci 2:1–24

Duvernoy HM (1999) The human brain: surface, blood supply, and three-dimensional sectional anatomy. Springer, New York

Eickhoff SB, Stephan KE, Mohlberg H, Grefkes C, Fink GR, Amunts K, et al (2005) A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. NeuroImage 25(4):1325–1335

Friston KJ, Worsley KJ, Frackowiak RSJ, Mazziotta JC, Evans AC (1994) Assessing the significance of focal activations using their spatial extent. Hum Brain Mapp 1(3):210–220

Gallese V, Lakoff G (2005) The brain's concepts: the role of the sensory-motor system in conceptual knowledge. Cognit Neuropsychol 22(3–4):455–479. doi:10.1080/02643290442000310

Gallese V, Keysers C, Rizzolatti G (2004) A unifying view of the basis of social cognition. Trends Cognit Sci 8(9):396–403

Gazzola V, Keysers C (2009) The observation and execution of actions share motor and somatosensory voxels in all tested subjects: single-subject analyses of unsmoothed fMRI data. Cereb Cortex 19(6):1239–1255. doi:10.1093/cercor/bhn181

Glenberg AM, Sato M, Cattaneo L, Riggio L, Palumbo D, Buccino G (2008) Processing abstract language modulates motor system activity. Q J Exp Psychol 61(6):905–919

Goldberg RF, Perfetti CA, Schneider W (2006a) Distinct and common cortical activations for multimodal semantic categories. Cognit Affect Behav Neurosci 6(3):214–222. doi:10.3758/cabn.6.3.214

Goldberg RF, Perfetti CA, Schneider W (2006b) Perceptual knowledge retrieval activates sensory brain regions. J Neurosci 26(18):4917–4921. doi:10.1523/jneurosci.5389-05.2006

Graziano MSA, Cooke DF (2006) Parieto-frontal interactions, personal space, and defensive behavior. Neuropsychologia 44(6):845–859

Grèzes J, Decety J (2001) Functional anatomy of execution, mental simulation, observation, and verb generation of actions: a meta-analysis. Hum Brain Mapp 12(1):1–19

Grèzes J, Frith C, Passingham RE (2004a) Brain mechanisms for inferring deceit in the actions of others. J Neurosci 24(24):5500–5505

Grèzes J, Frith CD, Passingham RE (2004b) Inferring false beliefs from the actions of oneself and others: an fMRI study. Neuroimage 21(2):744–750

Hamilton AF, Grafton ST (2006) Goal representation in human anterior intraparietal sulcus. J Neurosci 26(4):1133–1137

Hauk O, Johnsrude I, Pulvermüller F (2004) Somatotopic representation of action words in human motor and premotor cortex. Neuron 41(2):301–307

Haynes J-D, Rees G (2006) Decoding mental states from brain activity in humans. Nat Rev Neurosci 7(7):523–534

Hein G, Knight RT (2008) Superior temporal sulcus—it's my area: or is it? J Cogn Neurosci 20:2125–2136

Kamitani Y, Tong F (2006) Decoding seen and attended motion directions from activity in the human visual cortex. Curr Biol 16(11):1096–1102

Keysers C, Gazzola V (2007) Integrating simulation and theory of mind: from self to social cognition. Trends Cognit Sci 11(5):194–196

Killgore WDS, Young AD, Femia LA, Bogorodzki P, Rogowska J, Yurgelun-Todd DA (2003) Cortical and limbic activation during viewing of high-versus low-calorie foods. NeuroImage 19(4):1381–1394

Kilner JM, Neal A, Weiskopf N, Friston KJ, Frith CD (2009) Evidence of mirror neurons in human inferior frontal gyrus. J Neurosci 29(32):10153–10159. doi:10.1523/jneurosci.2668-09.2009

Mahon BZ, Caramazza A (2008) A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. J Physiol Paris 102(1–3):59–70

Makin TR, Holmes NP, Ehrsson HH (2008) On the other hand: dummy hands and peripersonal space. Behav Brain Res 191(1):1–10

Marsh LE, Hamilton AFdC (2011) Dissociation of mirroring and mentalising systems in autism. NeuroImage 56(3):1511–1519. doi:10.1016/j.neuroimage.2011.02.003

Martin A (2007) The representation of object concepts in the brain. Annu Rev Psychol 58(1):25–45. doi:10.1146/annurev.psych.57.102904.190143

Mitchell JP, Heatherton TF, Macrae CN (2002) Distinct neural systems subserve person and object knowledge. Proc Natl Acad Sci USA 99:15238

Mitchell JP, Banaji MR, Macrae CN (2005) General and specific contributions of the medial prefrontal cortex to knowledge about mental states. NeuroImage 28(4):757–762

Nichols T, Brett M, Andersson J, Wager T, Poline J-B (2005) Valid conjunction inference with the minimum statistic. NeuroImage 25(3):653–660

Noppeney U, Josephs O, Kiebel S, Friston KJ, Price CJ (2005) Action selectivity in parietal and temporal cortex. Cognit Brain Res 25(3):641–649

Peelen MV, Atkinson AP, Vuilleumier P (2010) Supramodal representations of perceived emotions in the human brain. J Neurosci 30(30):10127–10134. doi:10.1523/jneurosci.2161-10.2010

Pulvermuller F (2005) Brain mechanisms linking language and action. Nat Rev Neurosci 6(7):576–582

Ramsey R, Hamilton AF (2010) Understanding actors and object-goals in the human brain. NeuroImage 50:1142–1147

Ramsey R, Hamilton AF (2012) How does your own knowledge influence the perception of other people's actions in the human brain? Soc Cognit Affect Neurosci 7:242–251

Rizzolatti G, Craighero L (2004) The mirror-neuron system. Annu Rev Neurosci 27:169–192

Rueschemeyer S-A, Pfeiffer C, Bekkering H (2010) Body schematics: on the role of the body schema in embodied lexical-semantic representations. Neuropsychologia 48(3):774–781

Shipp S, Zeki S (1995) Segregation and convergence of specialised pathways in macaque monkey visual cortex. J Anat 187:547–562

Simmons WK, Martin A, Barsalou LW (2005) Pictures of appetizing foods activate gustatory cortices for taste and reward. Cereb Cortex 15(10):1602–1608. doi:10.1093/cercor/bhi038

Speer NK, Reynolds JR, Swallow KM, Zacks JM (2009) Reading stories activates neural representations of visual and motor experiences. Psychol Sci 20(8):989–999. doi:10.1111/j.1467-9280.2009.02397.x

Spunt RP, Lieberman MD (2012) Dissociating modality-specific and supramodal neural systems for action understanding. J Neurosci 32(10):3575–3583

Spunt RP, Falk EB, Lieberman MD (2010) Dissociable neural systems support retrieval of how and why action knowledge. Psychol Sci 21(11):1593–1598

Spunt RP, Lieberman, Satpute MD (2011) Identifying the what, why, and how of an observed action: an fMRI study of mentalizing and mechanizing during action observation. J Cognit Neurosci 23(1):63–74

Toni I, de Lange FP, Noordzij ML, Hagoort P (2008) Language beyond action. J Physiol Paris 102(1–3):71–79

Uddin LQ, Iacoboni M, Lange C, Keenan JP (2007) The self and social cognition: the role of cortical midline structures and mirror neurons. Trends Cognit Sci 11(4):153–157

Vigneau M, Beaucousin V, Hervé PY, Duffau H, Crivello F, Houdé O, et al (2006) Meta-analyzing left hemisphere language areas: phonology, semantics, and sentence processing. NeuroImage 30:1414–1432

Vigneau M, Beaucousin V, Hervé P-Y, Jobard G, Petit L, Crivello F, et al (2011) What is right-hemisphere contribution to phonological, lexico-semantic, and sentence processing?: insights from a meta-analysis. NeuroImage 54:577–593

Willems R, Hagoort P (in press) Cortical motor contributions to language understanding. In: Hermer L (ed) Contributions of primary sensory and motor cortices to higher cognitive processes. Research Signpost press, Scarborough, Ontario

Wunderlich G, Marshall JC, Amunts K, Weiss PH, Mohlberg H, Zafiris O, Zilles K, Fink GR (2002) The importance of seeing it coming: a functional magnetic resonance imaging study of motion-in-depth towards the human observer. Neuroscience 112(3):535–540