

Visuo-motor imagery of specific manual actions: A multi-variate pattern analysis fMRI study

Nikolaas N. Oosterhof^{a,b,c,*}, Steven P. Tipper^a, Paul E. Downing^a

^a School of Psychology, Bangor University, Bangor, United Kingdom

^b Centro Interdipartimentale Mente/Cervello (CIMeC), Trento University, TN, Italy

^c Department of Psychological & Brain Sciences, Dartmouth College, Hanover, NH, USA

ARTICLE INFO

Article history:

Accepted 24 June 2012

Available online 2 July 2012

Keywords:

Action representations

Imagery

Functional magnetic resonance imaging

Multi-variate pattern analysis

Human “mirror neuron system”

ABSTRACT

An important human capacity is the ability to imagine performing an action, and its consequences, without actually executing it. Here we seek neural representations of specific manual actions that are common across visuo-motor performance and imagery.

Participants were scanned with fMRI while they performed and observed themselves performing two different manual actions during some trials, and imagined performing and observing themselves performing the same actions during other trials. We used multi-variate pattern analysis to identify areas where representations of specific actions generalize across imagined and performed actions. The left anterior parietal cortex showed this property. In this region, we also found that activity patterns for imagined actions generalize better to performed actions than vice versa, and we provide simulation results that can explain this asymmetry. The present results are the first demonstration of action-specific representations that are similar irrespective of whether actions are actively performed or covertly imagined. Further, they demonstrate concretely how the apparent cross-modal visuo-motor coding of actions identified in studies of a human “mirror neuron system” could, at least partially, reflect imagery.

© 2012 Elsevier Inc. All rights reserved.

Introduction

Humans are social beings with a highly developed brain that enables them to interact with others and their environment in complex ways not seen in other animals. The ability to predict the consequences of our actions is crucial for such interactions, not only while performing actions—explained by ‘forward’ models (Kawato, 1999; Wolpert et al., 2003)—but also when imagining the outcome of our actions without actually executing them (Wolpert et al., 2003).

Although imagery of actions may be performed without any overt behavior, there is a long line of evidence showing that imagined actions and overt actions may share a cognitive mechanism. For example, Shepard and Metzler (1971) found that a mental object matching task showed a linear increase of reaction times as a function of rotation angle. Mental imagery of actions has also attracted interest in sport psychology, and mental practice is commonly reported in elite athletes (Hall and Rodgers, 1990). Several studies have shown that not only physical practice can improve performance but that imagery of practice, although generally less powerful, can improve sport performance as well (Hinshaw, 1991; Feltz and Landers, 1983).

Apart from improving existing skills, it has also been suggested that imagery is important in acquiring new action skills (Annett, 1995). For example, Mackay (1981) demonstrated that sentence production improves not only when producing these out loud but also silently. Beneficial effects of imagery have also been observed in typing (Wohldmann et al., 2007), music (Brodsky et al., 2008), dance (Golmer et al., 2008) and surgery (Hall, 2002). Analogous mechanisms may be involved when learning from observation, i.e. actions performed by others could be represented during observation, and then such representations could be re-activated, through mental imagery, during action execution (e.g., Sheffield, 1961; Cross et al., 2009; Calvo-Merino et al., 2005; Buccino et al., 2004; Cross et al., 2006).

Inspired by the discovery of ‘mirror neurons’ in macaques—neurons that fire both when the monkey executes an action or observes the action—in premotor and parietal cortex (di Pellegrino et al., 1992; Gallese et al., 1996), functional magnetic resonance imaging (fMRI) studies have investigated the neural correlates of such putative action representations shared across different modalities with a prime focus on imitation, observation, and execution. These studies have consistently found that several areas in the frontal and parietal cortex show an increased response for the imitation, observation and execution of actions, a result that often has been interpreted as evidence for human mirror neurons (e.g., Molenberghs et al., 2011; Rizzolatti and Fabbri-Destro, 2008; Gazzola and Keysers, 2008; but

* Corresponding author at: Centro Interdipartimentale Mente/Cervello (CIMeC), Palazzo Fedrigotti, Corso Bettini 31, Rovereto I-38068, TN, Italy.

E-mail address: nikolaas.oosterhof@unitn.it (N.N. Oosterhof).

see Brass and Heyes, 2005; Welberg, 2008; Dinstein et al., 2008b; Hickok, 2009; Iacoboni and Dapretto, 2006). Fewer studies have investigated the role of imagery, although there is evidence that imagined actions engage similar areas as observed or executed actions (Filimon et al., 2007; Lotze et al., 1999; Ehrsson, 2003). Crucially, the large majority of studies did not investigate the representations of specific actions (Gazzola and Keysers, 2008) and—given the limited spatial resolution of fMRI—do not rule out that observed and executed actions are subserved by different but spatially overlapping neural populations (Dinstein et al., 2008b). The few studies that have investigated action-specific representations yielded mixed results and interpretations, with some arguing for different neural populations for observed and executed actions (Dinstein et al., 2008a; Lingnau et al., 2009) and others for evidence for a cross-modal visuo-motor population (Chong et al., 2008; Kilner et al., 2009; Oosterhof et al., 2010, 2012).

To complicate this debate further, a possible interpretation of ostensible visuo-motor coding of observed and executed actions is mental imagery. For example, when participants observe actions they may also imagine executing the action without any overt behavior. Alternatively, participants may imagine the observation of their own actions while they execute unseen actions. Such imagery effects may result in engagement of unimodal action-specific representations that are similar both during performed and executed actions, and therefore would appear (incorrectly) to reflect cross-modal coding. Indeed, neuroimaging studies have provided evidence for shared representations of imagined and performed or observed movements of specific body parts in somatosensory (Stippich et al., 2002) and occipito-temporal cortices (Orlov et al., 2010). However, it is unclear whether mental imagery of specific actions can explain the apparent cross-modal coding of visuo-motor representations identified in previous studies (Chong et al., 2008; Kilner et al., 2009; Oosterhof et al., 2010; Oosterhof et al., 2012).

We sought to investigate this specific possibility empirically, and more generally to extend our understanding of the neural representations behind action imagery. We asked participants to perform two manual actions (while they could see their own right (dominant) hand) during certain trials, and to imagine these actions without overt behavior during other trials, while they were scanned with fMRI. We used multi-variate pattern analysis (MVPA; Edelman et al., 1998; Haxby et al., 2001; Norman et al., 2006; Haynes and Rees, 2006) to distinguish between the neural patterns associated with two manual actions. To anticipate our findings, in the crucial analysis comparing patterns of activity produced during performing vs. imagining actions, we found that the left anterior parietal cortex represents specific actions in a manner shared across both conditions. These findings are the first evidence that neural coding of specific imagined actions is similar to overtly performed actions and raise questions about the interpretation of studies investigating cross-modal visuo-motor coding of actions.

Methods

Participants

12 right-handed, healthy adult volunteers were recruited from the Bangor University community. All participants had normal or corrected-to-normal vision. Participants satisfied all requirements in volunteer screening and gave informed consent. Procedures were approved by the Ethics Committee of the School of Psychology at Bangor University. Participation was compensated at £15.

Design and procedure

The same setup was used as in Oosterhof et al. (2010, Experiment 2) in which participants manipulated a cup-shaped object that they could

see through a forward-looking mirror that was attached to the scanner coil. In the ‘perform’ condition, participants opened their eyes and performed either a ‘lift’ or a ‘slap’ action while they saw their hand and the object through the forward looking mirror. In the ‘imagery’ condition, participants closed their eyes and imagined both the motor and visual aspects of performing a ‘lift’ or a ‘slap’ (as in the perform condition), but without actually moving their hand (or any other body part) and without seeing their hand or the object. Thus, the design was $2 \{ \text{perform, imagery} \} \times 2 \{ \text{lift, slap} \}$ with 4 conditions in total.

Instructions consisted of a combination of spoken instructions (generated by the Apple Mac OS 10.6 “say” speech synthesizer program) and sinusoid tones that increased or decreased linearly in frequency between 400 and 1000 Hz during a 400 ms period. In perform trials the word ‘open’ was spoken starting at trial onset (instructing participants to open their eyes), followed by a sinusoid tone starting at one second after trial onset. Increasing or decreasing frequencies indicated whether participants should perform a lift or a slap, with the pairing (increasing tone for lift and decreasing for slap; or vice versa) counterbalanced across participants. In imagery trials the word ‘close’ was spoken at trial onset (instructing participants to close their eyes) followed by either the spoken word ‘lift’ or ‘slap’ starting one second after trial onset. Each trial lasted for four seconds in total. We used different types of action instructions (sinusoid tones versus speech) for perform and imagery trials so that any shared representations across perform and imagery trials could not be due to similarities of auditory stimuli. Furthermore we used speech for the imagery condition for all participants because we reasoned that (1) representations of imagined actions would be weaker than that of performed actions and (2) spoken words of the two actions would facilitate action imagery more than sinusoid tones.

Each participant was scanned during a single session with 6 to 8 functional scans and an anatomical scan if such a scan was not available from another scanning session. Each functional scan consisted of four ‘chunks’ of 16 trials (with a duration of 64 s) each, where each of the four conditions was presented four times in random order with the constraint that each condition preceded each condition equally often (i.e. the conditions were first-order counterbalanced; cf. Aguirre, 2007). Each chunk was preceded and followed by a 16 s baseline block that started with the auditory instruction ‘open’ followed by ‘relax for now’. Participants practiced performing and imagining the actions first outside the scanner and then a second time during a practice functional scan inside the scanner (for which the data was discarded). The practice scan was also used to ensure that the auditory stimuli could be heard well despite the scanner noise.

Participants were instructed similarly as in the study by Oosterhof et al. (2010) and as follows: to rest their right hand on the table, on the right-hand side of the object (from their perspective); to only move their right hand during perform trials; not to touch the object except during a perform trial; to keep their left hand and arm under the table, out of view; and after a “close” instruction, to keep their eyes closed until they were instructed to open them again and to imagine both executing and observing the instructed action without moving their hand or arm. Compliance of performing hand actions and opening and closing the eyes was monitored using an MRI compatible camera and eye tracker, respectively.

Data acquisition

The data was acquired using a 3 T Philips MRI scanner with a SENSE phased-array head coil. For functional imaging, a single shot T_2^* -weighted, gradient echo planar imaging sequence was used to achieve near-whole cerebrum coverage with the following parameters: repetition time (TR) 2500 ms; echo time (TE) 35 ms; flip angle 90°; 39 slices acquired in interleaved ascending order; no slice gap; field of view (FOV) $240 \times 240 \text{ mm}^2$; matrix size 96×96 ; $2.5 \times 2.5 \times 2.5 \text{ mm}^3$ voxels; anterior–posterior phase-encoding; SENSE factor 2. Slices were tilted

approximately 20° from the anterior commissure–posterior commissure axis in the frontal–superior direction. For participants with large brains the parietal lobe was fully covered at the expense of reduced coverage of the anterior–inferior part of the temporal lobes. Seven dummy scans were acquired before each functional run to reduce possible effects of T₁ saturation. Parameters for T₁-weighted anatomical scans were: matrix size 288 × 232; 1 × 1 × 1 mm³ voxels; TR 8.4 ms, TE 3.8 ms; flip angle = 8°.

Preprocessing

Data was processed using AFNI (Cox, 1996), SUMA (Saad et al., 2004), Matlab (the Mathworks Ltd., Cambridge, UK), the Surfing (Oosterhof et al., 2011a), LibSVM (Chang and Lin, 2011) and Fast Marching (Peyre, 2008) toolboxes, and custom written scripts in a similar fashion as described in earlier work (Oosterhof et al., 2011b).

Briefly, anatomical surfaces representing the outer (pial) and inner (white matter) boundaries of the cortical gray matter, and an inflated surface used for visualization, were reconstructed using the anatomical scans and Freesurfer's recon-all script (Fischl et al., 1999). These surfaces were aligned to a template brain based on cortical folding patterns, which improves inter-subject alignment. An intermediate surface was computed by taking the node-wise average of the pial and white matter surfaces. Surfaces were resampled to have a standard mesh topology with AFNI SUMA's Maplcohedron (200,000 triangles; 100,002 nodes) and subsequently aligned to the first functional volume of the first functional run using the surfing toolbox which uses AFNI's 3dAllineate (Saad et al., 2009). For visualization of group results, both the anatomical volume and the surfaces were transformed to Talairach (Talairach and Tournoux, 1988) space (using the estimated affine transformation from recon-all) and averaged. For cluster-based analysis, the group average area for each node was computed separately by taking the average across the intermediate surfaces across participants. For information mapping (Kriegeskorte et al., 2006; Oosterhof et al., 2011b) voxels were selected for each center node by constructing a searchlight disk on the intermediate surface based on geodesic distance, selecting the corresponding voxels in between or on the pial and white matter boundaries, and dynamically increasing the searchlight radius until approximately 200 voxels were selected.

Using AFNI (Cox, 1996), for each participant and each functional run separately, data was despiked (3dDespike), time-slice corrected (3dTshift), and motion corrected (3dvolreg) with trilinear resampling relative to the first functional volume of the first functional run. For surface mapping, the resulting data was projected onto the surface (3dVol2Surf) and spatially smoothed on the intermediate surface with a 5 mm full-width half-maximum (FWHM) Gaussian kernel. Percent signal change was computed at each spatial location (voxel or node) by dividing the signal in each time point by one percent of the average signal across the run.

Response estimates

The BOLD responses to the different conditions were estimated using the General Linear Model as implemented in AFNI's 3dDeconvolve. Predictors were based on a boxcar function (positive 1 to 4 s after trial onset) convolved with 3dDeconvolve's 'BLOCK' canonical hemodynamic response function. Two types of design matrices were used, one for univariate activation mapping and the other for MVPA. For activation mapping, the design matrices were based on the full session (all runs together) and contained predictors for perform and imagery trials, with no distinction between slap and lift actions. For information mapping, the design matrices were constructed for each run separately and contained predictors for each chunk (see above) and each action, yielding 8 action predictors per run. All design matrices contained predictors of no interest based

on head motion estimates (3 translation and 3 rotation parameters) and Legendre polynomials (up to and including third degree) to remove low-frequency drifts.

Univariate whole-brain activation mapping

To identify areas that showed an increased response during perform or imagery trials (compared to baseline periods), a group map was computed by testing the β -response estimates from individual participants' perform trials against zero with a one-sample *t*-test. A group map for imagery trials was computed similarly. Areas that showed an increased response during both perform and imagery trials were identified based on a conjunction map, where to each node separately, the minimum (or maximum) of the *t* values obtained from the perform and imagery maps was assigned if both values were positive (or negative, respectively). The value zero was assigned if the perform and imagery values had different signs.

Multivariate pattern analysis

Voxels were selected according to various criteria for MVPA (see below). Based on the *t*-values in these voxels, unimodal 'perform' pattern classification was conducted with a support vector machine classifier (Chang and Lin, 2011) by take-one-'perform action'-chunk-out cross-validation to distinguish between slap and lift actions. Classification accuracies were converted to *z* scores, where *z* = 0 (or *z* > 0) corresponds to at chance (or above chance, respectively) classification accuracy, and tested against zero (chance) using a one-tailed *t*-test. Unimodal 'imagery' patterns were analyzed similarly. For cross-modal perform-imagery MVPA, the train and test sets were in different modalities (train in perform, test on imagery; and vice versa) and results averaged. From each individual pattern the average was subtracted to normalize the response across patterns irrespective of main response differences, so that any action discrimination effects could not be due to global univariate activation differences across the two actions.

Multivariate region-of-interest analysis

Regions of interest were defined around center nodes using two criteria: functional and anatomical. Based on the functional criteria, group center nodes were selected under the constraints that (1) the *t*-value corresponded to an uncorrected significance level of *p* < .001 in the univariate conjunction analysis, and (2) no other node within 2.5 cm showed a higher *t*-value (cf., Fedorenko et al., 2010; Oosterhof et al., 2012). We note that although the same data was used for defining ROIs and subsequent MVPA in these ROIs, our approach is not susceptible to 'circular inference' (also known as 'double dipping') problems (Vul et al., 2009; Kriegeskorte et al., 2009) because the regressors used in the univariate analysis contained, by construction, no information about which specific action was performed or imagined during each trial. Based on the anatomical criteria, nodes of regions of interest found in a previous study (bilateral ventral and dorsal premotor cortex, anterior parietal cortex, and occipito-temporal cortex; Oosterhof et al., 2012) were taken as group center nodes.

Voxels were selected in individual participants as follows. First, individual center nodes were those with the maximal conjunction *t*-value within a 15 mm distance from the group center node. Second, a 10 mm radius sphere was centered around the voxel containing the individual center node. Third, a subset of voxels with the highest *t*-values in the individual's conjunction was selected for MVPA. Fourth, MVPA was conducted by taking 50 samples of random subsets from these voxels, MVPA was conducted with each sample, and the *z*-scores averaged across samples. Both the percentage of voxels initially selected ('ROI percentage') from the sphere and the percentage of voxels selected in random subsets ('subset percentage') varied from 10 to 90% in steps of 10%. This approach allowed for a (qualitative)

assessment of the reliability of the MVPA results with respect to variations in voxel selection parameters.

Multivariate region-of-interest simulations

To investigate the potential contributions of different neural populations coding for specific actions either in a unimodal-perform, unimodal-imagine, or cross-modal sense, we simulated data in a region of interest (100 voxels, 16 chunks with each of the four actions) as follows. First, we chose values a and b (between 1/10 and 10) for the relative contribution of unimodal-imagine and cross-modal patterns relative to the contribution of unimodal-perform. The latter had a relative contribution of 1; lower (or higher) values for a and b mean that more voxels represent actions either in a unimodal-imagine or cross-modal manner. Second, given the values for a and b , the 100 voxels were partitioned into three hypothetical ‘neural populations’ as follows: $100/(1+a+b)$ voxels were assigned to the unimodal-perform population, $100 \cdot a/(1+a+b)$ to the unimodal-imagine population, and $100 \cdot b/(1+a+b)$ to the cross-modal imagery population (these numbers were rounded to the nearest integer). For example, values of $a=0.5$, $b=8.5$ would yield 10 unimodal-perform voxels, 5 unimodal-imagine voxels, and 85 cross-modal voxels. Third, action-specific patterns were generated for each of the four actions so that in each neural population, the voxels coding for their respective modality differed by a value of 1 (arbitrary units) in their respective response between the two actions. For example, given a voxel coding for unimodal-perform actions, its simulated response values would be 1 higher in the perform-lift than in the perform-slap trials, but these values would not differ across the imagine-lift and imagine-slap trials. Similarly, a voxel coding cross-modally would have response values that were 1 higher in both the perform-lift and imagine-lift trials compared to perform-slap and imagine-slap trials. Fourth, random Gaussian noise was added to the simulated response. Similar to earlier work (Oosterhof et al., 2010, Supplementary Fig. 6), we assumed that the imagined trials would elicit a noisier response than the performed trials. Therefore we added Gaussian noise with .1 arbitrary unit² variance to the perform-trials and .01 arbitrary unit² variance to the imagined trials. Fifth, we varied a and b independently between 0.1 and 10, and for each combination an SVM classifier was trained and tested on this

data using cross-validation with a 1000 iterations and classification accuracies were averaged.

Multivariate whole-brain information mapping

Based on the 200 voxels selected around each node (see above), MVPA was conducted as described in the previous section and the corresponding z-scores assigned to each node. A threshold-free cluster enhancement (TFCE) map (Smith and Nichols, 2009) was computed based on the node-wise average z-score across participants using recommended values of $h_0=0$, $E=0.5$, $H=2$ and $dh=0.1$. Whole-hemisphere corrected significance of TFCE scores was performed using a bootstrap procedure. A null TFCE map was based on sampling with replacement 12 (the number of participants) individual participant’s maps, negating each z-score for each sample with a probability of .5 and computing the mean across these samples. This procedure was repeated a thousand times and p-values for each TFCE value in the original map were computed by dividing the number of times that value exceeded the maximum of each TFCE null map by the number of permutations (a thousand).

Results

Univariate activation mapping

The conjunction group analysis (Table 1, Fig. 1) revealed several regions with an increased response during both the perform and imagery conditions. Seven local maxima were identified in the left hemisphere and four in the right hemisphere, including bilateral planum temporale, left anterior parietal cortex, left supplementary motor area, and several clusters in frontal cortex bilaterally. The most robust activation was found in the bilateral planum temporale, probably due to auditory stimulus presentation.

Multivariate region-of-interest analysis

The local maxima from the univariate analyses were used as centers of regions of interest around which voxels were selected for MVPA (see Methods). In a first analysis (Table 1, Fig. 2a), MVPA was based, in each region-of-interest (ROI), on the top 50% (based on

Table 1
Regions of interest.

Abbr.	Area	Univariate peak coordinates			Peak t	Perform			Imagine			Cross-modal		
		x	y	z		Acc/%	t	p	Acc/%	t	p	Acc/%	t	p
IPTa	Left anterior planum temporale	−60 (4)	−16 (7)	4 (3)	12.6	55.3	3.1	0.0048	51.5	1.6	0.0646	51.2	1.7	0.0621
IPTp	Left posterior planum temporale	−63 (4)	−33 (8)	9 (4)	16.4	57.7	4.5	0.0004	53.9	3.8	0.0015	49.7	−0.4	0.6558
ITPJ	Left temporo-parietal junction	−56 (7)	−43 (5)	23 (7)	15.9	56.7	4.4	0.0005	52.0	2.0	0.0385	49.0	−1.5	0.9143
ISMA	Left supplementary motor area	−7 (1)	6 (5)	58 (7)	14.2	56.6	6.6	0	49.1	−0.7	0.7631	51.4	2.0	0.0374
IPMv	Left ventral premotor cortex	−52 (7)	7 (5)	5 (3)	8.8	57.0	5.1	0.0002	50.5	0.6	0.2776	50.3	0.4	0.3466
IPMd	Left dorsal premotor cortex	−38 (8)	−1 (5)	51 (3)	14.5	56.2	3.7	0.0019	49.5	−0.5	0.683	50.2	0.5	0.3146
IPCa	Left anterior parietal cortex	−44 (6)	−40 (5)	48 (4)	11.7	61.1	5.5	0.0001	50.6	0.4	0.3514	52.3	3.2	0.0042
IPrCGI	Left lateral pre-central gyrus	−59 (5)	2 (4)	18 (8)	9.9	56.8	5.7	0.0001	52.3	1.7	0.0542	50.8	1.2	0.1272
rPTa	Right anterior planum temporale	66 (4)	−14 (4)	4 (3)	12.1	55.5	3.1	0.0054	51.9	1.7	0.0562	49.8	−0.4	0.6519
rPTp	Right posterior planum temporale	57 (5)	−30 (5)	13 (4)	10.9	54.7	4.4	0.0005	51.7	1.1	0.1416	50.6	0.8	0.2258
rINS	Right insula	37 (3)	22 (7)	6 (5)	6.5	56.0	3.8	0.0015	51.0	1.3	0.1083	50.2	0.4	0.3592
rSTG	Right superior temporal gyrus	63 (4)	−35 (6)	9 (5)	15.6	52.7	2.4	0.0166	51.6	1.7	0.0563	49.6	−0.5	0.7017
IPMv	Left ventral premotor cortex	−53 (4)	−0 (4)	43 (6)	11.7	55.6	5.0	0.0002	49.5	−0.4	0.6548	50.2	0.3	0.3665
IPMd	Left dorsal premotor cortex	−33 (6)	−5 (3)	49 (3)	14.0	54.8	3.5	0.0026	50.3	0.2	0.4406	49.9	−0.3	0.5996
IPCa	Left anterior parietal cortex	−41 (6)	−38 (4)	49 (6)	7.6	63.0	7.8	0	51.1	0.9	0.1845	52.1	2.3	0.02
IOT	Left occipito-temporal cortex	−50 (4)	−72 (7)	2 (4)	6.8	58.3	12.9	0	53.8	1.5	0.0794	50.9	1.3	0.1031
rPMv	Right ventral premotor cortex	56 (3)	0 (4)	43 (4)	10.9	54.8	3.4	0.0032	51.0	0.8	0.2087	49.7	−0.5	0.6858
rPMd	Right dorsal premotor cortex	43 (9)	−5 (2)	54 (6)	7.8	56.9	8.8	0	52.0	3.6	0.0022	49.5	−0.5	0.6909
rPCa	Right anterior parietal cortex	34 (3)	−43 (4)	51 (8)	7.4	56.8	5.1	0.0002	50.3	0.3	0.3998	50.6	0.9	0.1993
rOT	Right occipito-temporal cortex	53 (7)	−65 (8)	−2 (5)	4.6	56.3	3.9	0.0013	49.8	−0.1	0.5499	50.9	1.6	0.0725

Talairach coordinates, average univariate conjunction peak value, and classification accuracies for regions of interest are shown in Fig. 6. Mean coordinates are reported with standard deviations in parentheses. The first eleven areas (above the horizontal line) are defined functionally, the other eight areas (below the horizontal line) are defined anatomically (see Methods). Accuracies are reported as percentages (50% is chance); p-values are not corrected for multiple comparisons.

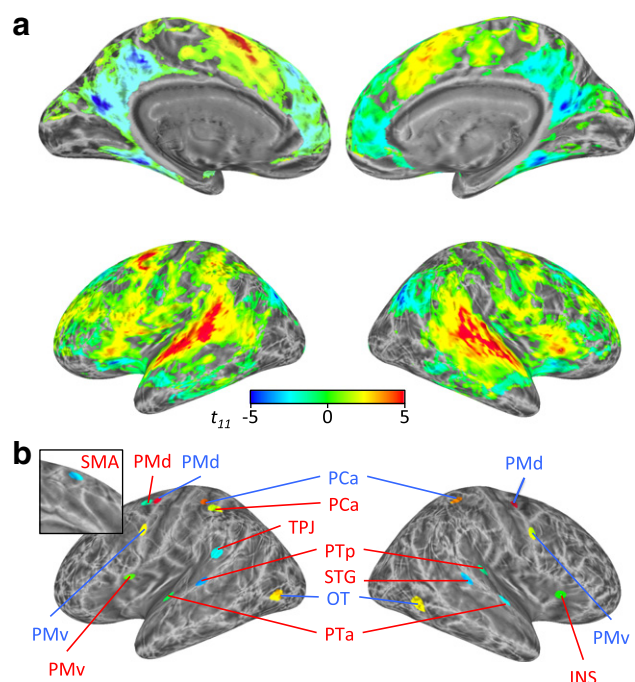


Fig. 1. Univariate conjunction group analysis and ROI definitions. (a) Conjunction group map of areas showing increased (red) or decreased (blue) activation compared to baseline for both performed and imagined actions. (b) Regions-of-interest based on local maxima in the conjunction analysis ((a); red font) and coordinates from a previous study (Oosterhof et al., 2012; blue font). See Table 1 for coordinates, statistics and abbreviations.

the conjunction response) of the voxels around a sphere with 10 mm radius. MVPA was based on the average z scores of 50 random samples that each contained 50% of the voxels. This analysis showed robust unimodal discrimination between lifts and slaps in the perform condition in all areas ($\max(p) = .016$ (uncorrected), all ROIs survive False Discovery Rate (FDR) correction). In the unimodal imagery condition action discrimination was weaker, with the strongest response in left posterior planum temporale [IPTp] ($p = .002$, surviving FDR correction), followed by left temporal parietal junction [ITPJ] ($p = .038$, not surviving FDR correction).

In the cross-modal perform-imagery analysis, action discrimination was observed in left anterior parietal cortex [IPTa] ($p = .004$, surviving FDR correction), followed by the left supplementary motor area [ISMA] ($p = .037$, not surviving FDR correction). A post-hoc analysis on the left anterior parietal cortex region showed an asymmetry in cross-modal classification: training on imagined trials and testing on performed trials showed better action discrimination ($t_{11} = 3.89$, $p = .001$) than training on performed trials and testing on imagined trials ($t_{11} = 1.87$, $p = .044$), and this difference was significant ($t_{11} = 4.10$, $p = 0.002$ (two-tailed)).

In an earlier study, with different participants, we identified other areas—bilateral ventral and dorsal premotor cortex, anterior parietal cortex, and occipito-temporal cortex—that showed increased activity (compared to baseline) for both performed and observed actions (Oosterhof et al., 2012). MVPA around these areas (Fig. 2b) showed similar effects: robust action discrimination in the uni-modal perform condition and weaker discrimination in the uni-modal imagery condition (with a possible exception for right dorsal premotor cortex). The cross-modal analysis based on these areas showed similar but weaker evidence for action discrimination in left anterior parietal cortex ($p = .02$, uncorrected) and no evidence for such action discrimination in other areas.

To assess the robustness of the MVPA results with respect to voxel selection parameters we varied both the percentage of voxels selected

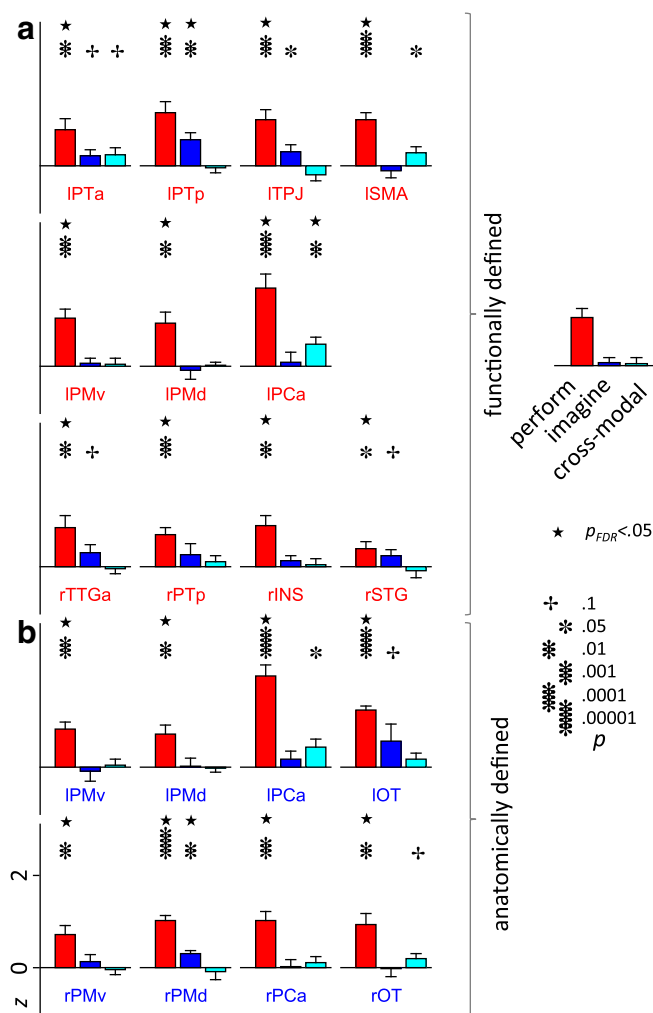


Fig. 2. Region-of-interest pattern classification accuracies. (a) Pattern classification z-scores (0=chance) for action discrimination in 12 regions-of-interest (ROIs; see Fig. 1, Table 1) localized from the univariate group analysis (see Fig. 1) for unimodal performed, unimodal-imagined, and cross-modal MVPA. p-Values are not corrected for multiple comparisons; bars indicated with a '★' (star) survive False Discovery Rate correction for 11 ROIs. (b) Similar classification scores based on coordinates from an earlier study (Oosterhof et al., 2012). Conventions are as in (a).

based on the largest perform-imagery conjunction response and the percentage of voxels in random subsets for MVPA (see Methods). Results are shown in heat maps for the cross-modal analysis for both the mean accuracy z scores (Fig. 3) and t scores (Fig. 4). Note that the center square in an ROI's heatmap represents the corresponding cross-modal bar in Fig. 2. We do not attempt a full quantitative analysis of the effects of voxel selection—not at least because of chance capitalization considerations—but on a qualitative level the most robust response, irrespective of specific voxel selection parameters, was observed in left anterior parietal cortex.

To increase our understanding of the imagine-perform and perform-imagery cross-modal classification accuracies asymmetry observed in the left anterior parietal region (see above), in the region-of-interest simulation (Fig. 5) we assessed the effect of different contributions from three different, hypothetical populations representing specific actions either in the performed modality, in the imagined modality, or cross-modally. As expected – by construction of the amount of noise added to the generated patterns – within-modality classification accuracies were higher for performed (Fig. 5a) than for imagined (Fig. 5d) actions. More interestingly, and consistent with the classification accuracy asymmetry observed in the left anterior parietal cluster, classification accuracies were greater

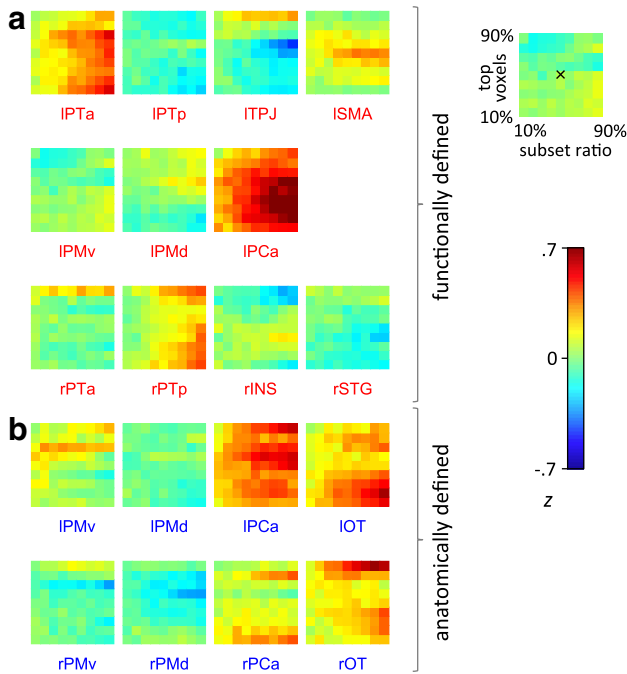


Fig. 3. Region-of-interest cross-modal perform–imagine pattern classification accuracy z scores as a function of voxel selection parameters. Heat maps for each region-of-interest (see Fig. 1, Table 1), defined (a) functionally or (b) anatomically, representing cross-modal perform–imagine multi-variate pattern analysis action discrimination z -scores as a function of the percentage of maximally responsive voxels selected from the univariate conjunction analyses (rows) and the percentage of voxels taken from these of random subsets (columns); see inset. The value at the center of each heat map, indicated with an 'x' in the inset, corresponds to the cross-modal bars in each of the region-of-interest plots displayed in Fig. 2.

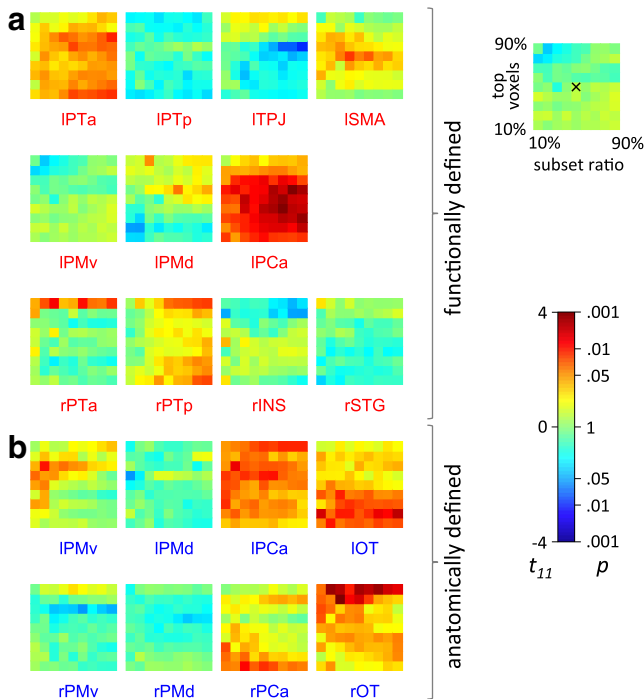


Fig. 4. Region-of-interest cross-modal perform–imagine pattern classification accuracy t -scores as a function of voxel selection parameters. Heat maps for each region-of-interest, defined (a) functionally or (b) anatomically, are as in Fig. 3 but representing t - rather than z -scores.

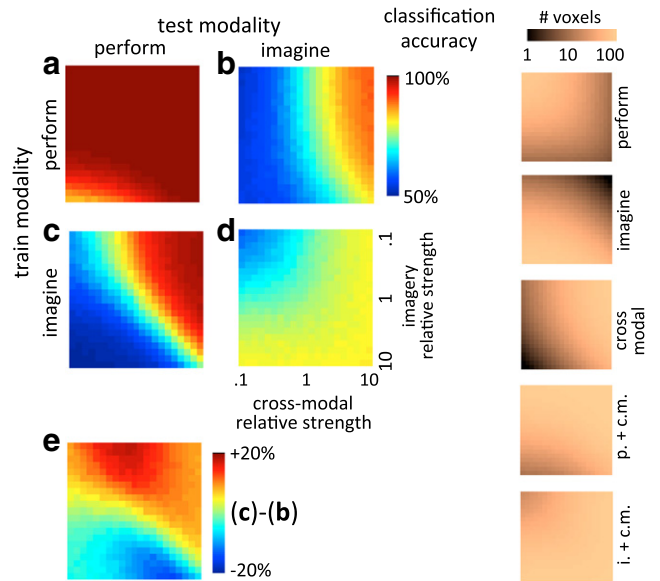


Fig. 5. Region-of-interest simulation of cross-modal classification accuracies for different mixtures of different neural populations. Heat maps of action-specific classification accuracies when the train–test set of trials are either (a) perform–perform, (b) perform–imagine, (c) imagine–perform, or (d) imagine–imagine. Patterns consisted of a mixture of simulated neural populations representing performed, imagined, and cross-modal performed–imagined actions (see Methods). The relative contribution of each of these three populations varies as a function of the location in the heat map: the horizontal and vertical axes show the strength of patterns (in terms of number of voxels) representing actions cross-modally and imagined, respectively, relative to the strength of patterns representing performed actions (strength = 1). The number of voxels in each simulated population, and the sum of the number of voxels in the cross-modal and each of the other two populations, is shown on the right hand side. Simulated patterns had more noise added for imagined trials than performed trials. (e) The difference between maps (c) and (b), illustrating that different mixtures of the three types of simulated populations can yield asymmetries between perform–imagine and imagine–perform cross-modal classification accuracies. Abbreviations: p. + c.m., perform and cross-modal; i. + c.m., imagine and cross-modal.

when the classifier was trained on imagined actions and tested on performed actions (Fig. 5c) than in the reverse case (Fig. 5b) when the simulated imagery population was stronger (contained more voxels). The difference between these cases was maximal when the contribution of unimodal imagery patterns was relatively low (fewer voxels) compared to unimodal performed or cross-modal patterns, and minimal (negative) when the contribution of unimodal performed patterns was relatively strong (more voxels) compared to unimodal imagined and cross-modal patterns. We note that additional analyses where performed and imagined actions were of similar strength (i.e. similar amounts of noise were added), showed quantitatively similar results as Fig. 5e although the effects were weaker, which suggests that asymmetries in classification accuracies are not only due to different noise levels across performed and imagined actions but can be amplified by such differences.

Multivariate whole-brain information mapping

Consistent with the ROI analyses, the unimodal perform information map (Fig. 6a) showed large clusters surviving multiple-comparison correction in and around the visual, auditory, motor and somatosensory areas. In the unimodal imagery information map (Fig. 6b), regions in the bilateral auditory cortex (planum temporale) and also a cluster in left frontal cortex survived multiple-comparison correction. In the critical cross-modal case (Fig. 6c), only a cluster in left anterior parietal cortex ($x, y, z = -52, -34, 40$, area 770 mm^2) survived multiple-comparison correction. A post-hoc analysis directly comparing accuracies in two different cross-modal train–test arrangements (imagine–perform and perform–imagine) found no significant clusters, although

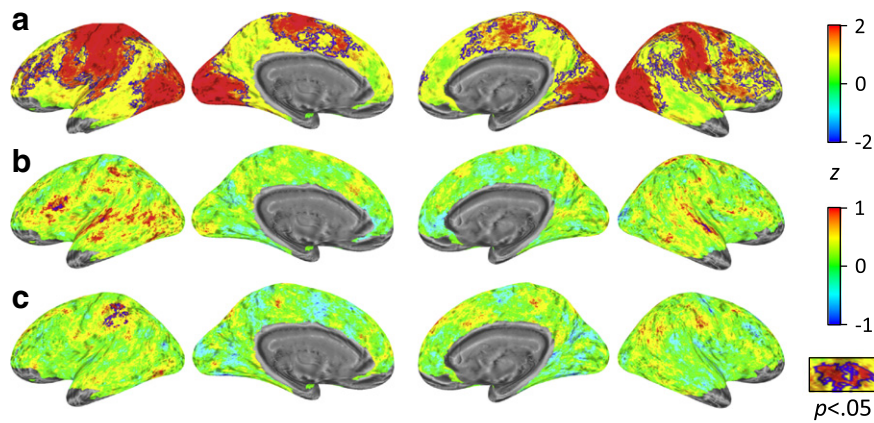


Fig. 6. MVPA group analysis. Whole-brain group analysis showing classification z-scores for (a) unimodal perform, (b) unimodal imagery, and (c) cross-modal perform-imagery discrimination between lift and slap actions. Nodes surviving Threshold-Free Cluster Enhancement multiple-comparison correction (see [Methods](#)) are surrounded by blue.

the strongest cluster across the brain ($\min(p \text{ corrected}) = 0.14$) was located at a similar location in the left anterior parietal cortex, consistent with the ROI results reported above.

Discussion

Using fMRI MVPA we investigated how the human brain represents performed and imagined actions. In the unimodal perform analysis, when participants performed two distinct object-related actions (lifts and slaps) while observing their hand and the object that was manipulated, we found that spatially distributed patterns dissociated the two actions across large portions of the cortex that included auditory, visual, somatosensory, and motor areas. The involvement of these areas is not surprising, given that the two actions were cued with different auditory cues, performing the actions required action-specific motor planning and execution, and observing the hand and object while performing the actions yielded visual inputs that differed between the actions.

In the unimodal imagery analysis we found that the planum temporale, bilaterally, dissociated the two actions. Because this area contains the auditory cortices, this is most likely due to the difference between the sounds of the words that cued the action that was imagined. The only other area that also showed discrimination between actions in the whole-brain analyses was in the left frontal cortex near Broca's language area. One explanation is that this area is involved in processing the spoken action instructions. Another explanation is that this area is involved in representing manual actions (e.g., [Fogassi and Ferreri, 2007](#); [Heiser et al., 2003](#)). With the current paradigm we are not able to dissociate these two explanations and refrain from further speculation.

Most importantly, the cross-modal perform-imagery analysis found that the left anterior parietal cortex represented specific actions similarly across the perform and imagery conditions. Because in the imagery condition participants closed their eyes, did not move, and received auditory stimuli (words) that were different than those presented in the perform condition (sinusoid tones), these effects cannot be explained by trivial stimulus properties shared across the perform and imagery condition such as motor planning or execution, or visual or auditory input. We note that although the discrimination in the 'imagine' condition was relatively weak and did not reach significance on itself, we have shown earlier ([Oosterhof et al., 2010](#), Supplementary Fig. 6) and in the current work ([Fig. 5](#)) that this not does preclude the possibility of detecting cross-modal information if the other ('perform', in this case) modality shows strong action-specific representations.

More specifically, our results showed higher cross-modal classification accuracies when the classifier was trained on imagined trials and tested on performed trials than in the reverse case. We found similar effects in simulations that assumed varying degrees of contributions (in terms of relative numbers of voxels) of three different neural populations that represent specific actions only when performed, only when imagined, or across modalities. These findings can be explained by considering which information a classifier uses when it is trained on a series of patterns. If the pattern information for performed actions is relatively strong (due to visual input, motor planning and execution, etc.) compared to imagined and cross-modal patterns, then training a classifier on performed actions causes the decision boundary of the classifier to be based mainly on patterns evoked by these aspects of action performance that do not generalize to imagined actions. Hence classification accuracies of these imagined actions will be relatively low. On the other hand, training a classifier on imagined actions will cause the decision boundary to be based on a mixture of imagined and cross-modal information, and generalization to performed actions will be relatively high.

The preceding analysis suggests that patterns evoked during imagined actions are less affected by brain responses induced by modality-specific effects and therefore more likely to access relatively 'pure' representations of actions that generalize across modalities. Note in contrast that in previous work ([Oosterhof et al., 2010](#), Supplementary Figs. 2 and 4) we did not find a similar asymmetry between training and testing on trials with actions that were observed (without concurrent action execution) or performed (without visual input). We stress that although our simulations can be used as a model to increase the understanding of the potential contribution of pattern information from different neural populations, it is not necessarily the case that such a model represents brain function accurately at either the voxel or at the neural level, and inferences at a neural population level based on such a model should remain speculative.

What is the nature of these action-specific representations that are common across action imagery and performance? While there is an extensive literature on imagery in general, for the purpose of interpreting our findings we distinguish three, not mutually exclusive, possibilities: visual, motor, and amodal. According to a visual interpretation, participants represented the imagined actions as visual "pictures" through top-down control, and the resulting brain activity patterns were similar to those when participants actually saw their own hand perform the actions. Indeed, several studies have demonstrated that imagined and observed pictures share a common representation in early and high-level visual cortex. For example, imaging and seeing the letters 'O' and 'X' elicit similar response patterns in lateral occipital cortex ([Stokes et al., 2009](#)), different categories of

objects elicit imagined-seen cross-modal patterns in ventral temporal cortex (Reddy et al., 2010), and different locations of objects reveal analogous cross-modal representations in early visual cortex (Cichy et al., 2011). (Note however that the preceding studies all concerned static stimuli and images in contrast to the dynamic content tested here).

Alternatively, according to a motor interpretation, participants represented the imagined actions as executing motor actions without actually moving (Guillot and Collet, 2005; Decety and Grezes, 2006; Johnson, 2000; Ramsey et al., 2010), and the resulting brain activity patterns were similar to when participants executed the corresponding action. This interpretation is consistent with findings that executed and imagined hand movements elicit increased activation in primary and secondary motor cortex of the contralateral hemisphere (Lotze et al., 1999), which has been interpreted as possibly reflecting similar neural substrates for motor execution and imagery. Also consistent with this interpretation is a study that showed evidence for a somatotopic organization in primary, supplementary and pre-motor cortex when participants moved or imagined moving their hands, feet, and tongues (Ehrsson, 2003).

A third alternative is an amodal representation, where actions are represented neither visually nor motorically but on a more abstract level (Pylyshyn, 2003). For example, findings that listening to action verbs showed increased activation in Broca's area compared to non-action verbs have been interpreted as evidence for involvement in abstract action representation of this area (Tettamanti et al., 2005). Other evidence comes from a study with congenitally blind participants, in which listening to action sounds, compared to environmental sounds, activated similar frontal and parietal areas as when participants with normal vision observed actions visually, which is consistent with findings supporting amodal action representations in these areas (Ricciardi et al., 2009).

As we noted earlier, these possibilities are not mutually exclusive and different brain areas may represent actions differently. This question is further complicated by the difference between the subjective phenomenological (as reported verbally by participants, for example) and the objective brain response (as measured with fMRI, for example) aspects of consciousness generally (Lamme, 2006) and action imagery specifically. Although imagery involves a subjective experience, behavioral experiments have shown that tasks that are thought to require either visual (Kosslyn et al., 1978) or motor (Johnson, 1982) imagery yield behavioral effects that are similar to overt visual or motor tasks, showing that subjective motor and visual imagery dissociate in objectively measurable effects. Brain imaging studies have also shown that different brain networks are recruited when participants are explicitly instructed to imagine performing or viewing actions (Pelgrims et al., 2009; Guillot et al., 2009; Sirigu and Duhamel, 2001; Decety, 1990; Kosslyn and Thompson, 1997). Altogether this suggests that the distinction between motor and visual imagery representations has objectively measurable correlates.

One might take the position that an action representation is visual if imagery of that action activates visual areas, motoric if it activates motor areas, and amodal if it activates other areas. Apart from simple interpretational challenges (Poldrack, 2006), such a position is problematic, however, for more substantial reasons. First, differences in activation within or across brain areas can be caused by trivial aspects of the experiment. For example, if one were to perform visual imagery and visual observation in an experiment where participants had their eyes closed and open, respectively, then the overall activation in 3-visual cortex may be decreased during the imagery condition because there is no visual input to the participants, yet it does not indicate that the visual cortex is not involved during visual imagery (Stokes et al., 2009; Reddy et al., 2010; Cichy et al., 2011).

Second, brain areas active during visual and motor imagery are not identical to those involved during overt action observation and execution (Hanakawa, 2002; Ganis et al., 2004). This is not surprising given

that visual imagery may require top-down cognitive control and the engagement of memory areas. Similarly, motor imagery may require inhibitory processes that prevent imagined planned actions being translated into actual movements. Therefore it is not straightforward to interpret the visual and motor aspects of the neural patterns associated with imagery of performed or observed actions.

Third, there is extensive evidence that individual neurons show a response to multiple modalities, ranging from single unit responses in rodents (Barth et al., 1995) and monkeys (Bignall and Imbert, 1969) to humans (Mukamel et al., 2010). Responses from populations of such neurons, as measured indirectly by fMRI cannot, by definition, be considered as unimodal only.

Relevant for this third point, especially with respect to action representations, is the finding of 'mirror neurons' in ventral premotor cortex in the macaque (di Pellegrino et al., 1992). These neurons have been shown to increase their firing rate when a macaque either performed an action or observed the experimenter performing the same action. Later studies showed neurons with similar properties in the macaque anterior parietal (Gallese et al., 1996) and primary motor cortex (Dushanova and Donoghue, 2010), and in human hippocampus and pre-supplementary motor cortex (Mukamel et al., 2010).

Imaging studies that considered similar action-specific coding across the visual and motor modalities have investigated to what extent areas in premotor, parietal and occipito-temporal cortices may represent specific actions similarly across the visual and motor domains (Kilner et al., 2009; Oosterhof et al., 2010, 2012; Lingnau et al., 2009; Dinstein et al., 2008a; Dinstein et al., 2007; Chong et al., 2008). Results and conclusions of these studies have been mixed—with some claiming evidence for shared visuo-motor representations (Chong et al., 2008; Kilner et al., 2009; Oosterhof et al., 2010, 2012) and others claiming no evidence for such representations (Dinstein et al., 2008a; Dinstein et al., 2007; Lingnau et al., 2009). The current results provide evidence that also support a unimodal representational account of the previous imaging findings: if participants imagined performing or observing actions while actively observing or performing actions (respectively), then the shared response for viewing and executing specific actions may be due to shared imagined-overt visual, motoric, or amodal coding.

We note that we only found evidence for imagined-overt cross-modal coding in the anterior parietal cortex and not in other areas such as the premotor and occipito-temporal cortices that were identified in similar previous studies considering visuo-motor action representations (Kilner et al., 2009; Oosterhof et al., 2010, 2012). The conclusion that these other areas are not representing specific actions during imagery cannot be drawn, however. First, statistical power to detect imagery effects in these areas may have been too weak. Second, the absence of an overt task may have affected the strength of imagery representations. In the present study participants were asked to imagine actions but not to perform any other task, while in earlier MVPA studies (Oosterhof et al., 2010, 2012) that investigated visuo-motor cross-modal coding, participants were required to respond after certain observed trials, which might have led to deeper encoding of the actions. We note that these considerations also prevent meaningful interpretations of direct comparisons between the present and other studies (e.g., Willems et al., 2010).

The current findings are the first demonstration that specific imagined actions are represented similarly to overtly performed and observed actions, and provide a potential mechanism for ostensible visuo-motor coding claimed in human action representation theories. The precise nature of these representations—visual, motor, or amodal—and how they can be modulated by task requirements is still elusive, however. Because participants in our study were instructed to both execute and view their actions (in the perform condition) or to imagine themselves executing and viewing the actions (in the imagery condition), the present data cannot dissociate these possibilities. One approach would be an experiment where participants are instructed to perform, view, or view and

perform, and imagine to perform, view, or view and perform, specific actions. The relative strength of several types of cross-modal coding (visual, motoric, and both, crossed with imagined and overtly performed actions; cf. Kriegeskorte et al., 2008) could help in characterizing the nature of action representations in different brain areas and across different modalities.

The authors declare no conflict of interest.

Acknowledgments

This research was supported by the ESRC (grant to SPT and PED), and the Wales Institute of Cognitive Neuroscience. NNO was supported by a fellowship awarded by the Boehringer Ingelheim Fonds. We would like to thank Marius Peelen for helpful discussions, and Emily Cross, Angelika Lingnau, Nick Peatfield, Marius Peelen, Richard Ramsey, and two anonymous reviewers for helpful comments on an earlier draft of this manuscript.

References

- Aguirre, G.K., 2007. Continuous carry-over designs for fMRI. *Neuroimage* 35 (4), 1480–1494.
- Annett, J., 1995. Motor imagery: perception or action? *Neuropsychologia* 33 (11), 1395–1417.
- Barth, D.S., Goldberg, N., Brett, B., Di, S., 1995. The spatiotemporal organization of auditory, visual, and auditory–visual evoked potentials in rat cortex. *Brain Res.* 678, 177–190.
- Bignall, K.E., Imbert, M., 1969. Polysensory and cortico-cortical projections to frontal lobe of squirrel and rhesus monkeys. *Electroencephalogr. Clin. Neurophysiol.* 26, 206–215.
- Brass, M., Heyes, C., 2005. Imitation: is cognitive neuroscience solving the correspondence problem? *Trends Cogn. Sci.* 9 (10), 489–495.
- Brodsky, W., Kessler, Y., Rubinstein, B.-S., Ginsborg, J., Henik, A., 2008. The mental representation of music notation: notational adaptation. *J. Exp. Psychol. Hum. Percept. Perform.* 34 (2), 427–445.
- Buccino, G., Vogt, S., Ritzl, A., Fink, G.R., Zilles, K., Freund, H.J., et al., 2004. Neural circuits underlying imitation learning of hand actions: an event-related fMRI study. *Neuron* 42 (2), 323–334.
- Calvo-Merino, B., Glaser, D., Grezes, J., 2005. Action observation and acquired motor skills: an fMRI study with expert. *Cereb. Cortex* 15 (8), 1243–1249.
- Chang, C.-C., Lin, C.-J., 2011. LIBSVM: a library for support vector machines. *ACM Trans. Intell. Syst. Technol.* 2, 27:1–27:27. <http://www.csie.ntu.edu.tw/~cjlin/libsvm>.
- Chong, T.T.-J., Cunningham, R., Williams, M., Kanwisher, N., Mattingley, J., 2008. fMRI adaptation reveals mirror neurons in human inferior parietal cortex. *Curr. Biol.* 18 (20), 1576–1580.
- Cichy, R.M., Heinze, J., Haynes, J.D., 2011. Imagery and perception share cortical representations of content and location. *Cereb. Cortex*. <http://dx.doi.org/10.1093/cercor/bhr106>.
- Cox, R.W., 1996. AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Comput. Biomed. Res.* 29 (3), 162–173.
- Cross, E.S., Hamilton, A.F.C., Grafton, S.T., 2006. Building a motor simulation de novo: observation of dance by dancers. *Neuroimage* 31 (3), 1257–1267.
- Cross, E.S., Kraemer, D.J.M., Hamilton, A.F.C., Kelley, W.M., Grafton, S.T., 2009. Sensitivity of the action observation network to physical and observational learning. *Cereb. Cortex* 19 (2), 315–326.
- Decety, J., 1990. Brain structures participating in mental simulation of motor behavior: a neuropsychological interpretation. *Acta Psychol.*
- Decety, J., Grezes, J., 2006. The power of simulation: imagining one's own and other's behavior. *Brain Res.* 1079 (1), 4–14.
- di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., Rizzolatti, G., 1992. Understanding motor events – a neurophysiological study. *Exp. Brain Res.* 91 (1), 176–180.
- Dinstein, I., Hasson, U., Rubin, N., Heeger, D.J., 2007. Brain areas selective for both observed and executed movements. *J. Neurophysiol.* 98 (3), 1415–1427.
- Dinstein, I., Gardner, J.L., Jazayeri, M., Heeger, D.J., 2008a. Executed and observed movements have different distributed representations in human aIPs. *J. Neurosci.* 28 (44), 11231–11239.
- Dinstein, I., Thomas, C., Behrmann, M., Heeger, D.J., 2008b. A mirror up to nature. *Curr. Biol.* 18 (1), R13–R18.
- Dushanova, J., Donoghue, J., 2010. Neurons in primary motor cortex engaged during action observation. *Eur. J. Neurosci.* 31 (2), 386–398.
- Edelman, S., Grill-Spector, K., Kushnir, T., Malach, R., 1998. Toward direct visualization of the internal shape representation space by fMRI. *Psychobiology* 26 (4), 309–321.
- Ehrsson, H.H., 2003. Imagery of voluntary movement of fingers, toes, and tongue activates corresponding body-part-specific motor representations. *J. Neurophysiol.* 90 (5), 3304–3316.
- Fedorenko, E., Hsieh, P.J., Nieto-Castanon, A., Whitfield-Gabrieli, S., Kanwisher, N., 2010. New method for fMRI investigations of language: defining ROIs functionally in individual subjects. *J. Neurophysiol.* 104 (2), 1177–1194.
- Feltz, D.L., Landers, D.M., 1983. The effects of mental practice on motor skill learning and performance: a meta-analysis. *J. Sport Psychol.* 5, 25–57.
- Filimon, F., Nelson, J.D., Hagler, D.J., Sereno, M.I., 2007. Human cortical representations for reaching: mirror neurons for execution, observation, and imagery. *Neuroimage* 37 (4), 1315–1328.
- Fischl, B., Sereno, M.I., Tootell, R.B.H., Dale, A.M., 1999. High-resolution intersubject averaging and a coordinate system for the cortical surface. *Hum. Brain Mapp.* 8 (4), 272–284.
- Fogassi, L., Ferrari, P.F., 2007. Mirror neurons and the evolution of embodied language. *Curr. Dir. Psychol. Sci.* 16 (3), 136–141.
- Gallese, V., Fadiga, L., Fogassi, L., Rizzolatti, G., 1996. Action recognition in the premotor cortex. *Brain* 119, 593–609.
- Ganis, G., Thompson, W.L., Kosslyn, S.M., 2004. Brain areas underlying visual mental imagery and visual perception: an fMRI study. *Cogn. Brain Res.* 20 (2), 226–241.
- Gazzola, V., Keysers, C., 2008. 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.
- Golmer, E., Boulliette, A., Mertz, C., Keller, J., 2008. Effects of mental imagery styles on shoulder and hip rotations during preparation of pirouettes. *J. Mot. Behav.* 40 (4), 281–290.
- Guillot, A., Collet, C., 2005. Contribution from neurophysiological and psychological methods to the study of motor imagery. *Brain Res. Rev.* 50 (2), 387–397.
- Guillot, A., Collet, C., Nguyen, V.A., Malouin, F., Richards, C., Doyon, J., 2009. Brain activity during visual versus kinesthetic imagery: an fMRI study. *Hum. Brain Mapp.* 30 (7), 2157–2172.
- Hall, J.C., 2002. Imagery practice and the development of surgical skills. *Am. J. Surg.* 184 (5), 465–470.
- Hall, C.R., Rodgers, W.M., 1990. The use of imagery by athletes in selected sports. *Sport Psychologist* 4, 1–10.
- Hanakawa, T., 2002. Functional properties of brain areas associated with motor execution and imagery. *J. Neurophysiol.* 89 (2), 989–1002.
- Haxby, J.V., Gobbini, M.I., Furey, M.L., Ishai, A., Schouten, J.L., Pietrini, P., 2001. Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science* 293 (5539), 2425–2430.
- Haynes, J.-D., Rees, G., 2006. Decoding mental states from brain activity in humans. *Nat. Rev. Neurosci.* 7 (7), 523–534.
- Heiser, M., Iacoboni, M., Maeda, F., Marcus, J., Mazziotta, J.C., 2003. The essential role of Broca's area in imitation. *Eur. J. Neurosci.* 17 (5), 1123–1128.
- Hickok, G., 2009. Eight problems for the mirror neuron theory of action understanding in monkeys and humans. *J. Cogn. Neurosci.* 21 (7), 1229–1243.
- Hinshaw, K.E., 1991. The effects of mental practice on motor skill performance: critical evaluation and meta-analysis. *Imagin. Cogn. Pers.* 11 (1), 3–35.
- Iacoboni, M., Dapretto, M., 2006. The mirror neuron system and the consequences of its dysfunction. *Nat. Rev. Neurosci.* 7 (12), 942–951.
- Johnson, P., 1982. The functional equivalence of imagery and movement. *Q. J. Exp. Psychol. A* 34 (3), 349–365.
- Johnson, S., 2000. Thinking ahead: the case for motor imagery in prospective judgments of prehension. *Cognition* 74 (1), 33–70.
- Kawato, M., 1999. Internal models for motor control and trajectory planning. *Curr. Opin. Neurobiol.* 9 (6), 718–727.
- Kilner, J.M., Neal, A., Weiskopf, N., Friston, K.J., Frith, C.D., 2009. Evidence of mirror neurons in human inferior frontal gyrus. *J. Neurosci.* 29 (32), 10153–10159.
- Kosslyn, S.M., Thompson, W.L., 1997. Neural systems shared by visual imagery and visual perception: a positron emission tomography study. *Neuroimage* 6 (4), 320–334.
- Kosslyn, S.M., Ball, T.M., Reiser, B.J., 1978. Visual images preserve metric spatial information: evidence from studies of image scanning. *J. Exp. Psychol.* 4 (1), 47–60.
- Kriegeskorte, N., Goebel, R., Bandettini, P., 2006. Information-based functional brain mapping. *Proc. Natl. Acad. Sci. U. S. A.* 103 (10), 3863–3868.
- Kriegeskorte, N., Mur, M., Bandettini, P., 2008. Representational similarity analysis – connecting the branches of systems neuroscience. *Front. Syst. Neurosci.* 2:4. <http://dx.doi.org/10.3389/neuro.06.004.2008>.
- Kriegeskorte, N., Simmons, W.K., Bellgowan, P.S.F., Baker, C.I., 2009. Circular analysis in systems neuroscience: the dangers of double dipping. *Nat. Neurosci.* 12 (5), 535–540.
- Lamme, V.A.F., 2006. Towards a true neural stance on consciousness. *Trends Cogn. Sci.* 10 (11), 494–501.
- Lingnau, A., Gesierich, B., Caramazza, A., 2009. Asymmetric fMRI adaptation reveals no evidence for mirror neurons in humans. *Proc. Natl. Acad. Sci. U. S. A.* 106 (24), 9925–9930.
- Lotze, M., Montoya, P., Erb, M., Hulsmann, E., Flor, H., Klose, U., et al., 1999. Activation of cortical and cerebellar motor areas during executed and imagined hand movements: an fMRI study. *J. Cogn. Neurosci.* 11 (5), 491–501.
- Mackay, D.G., 1981. The problem of rehearsal or mental practice. *J. Mot. Behav.* 13 (4), 274–285.
- Molenberghs, P., Cunnington, R., Mattingley, J.B., 2011. Brain regions with mirror properties: a meta-analysis of 125 human fMRI studies. *Neurosci. Biobehav. Rev.* <http://dx.doi.org/10.1016/j.neubiorev.2011.07.004>.
- Mukamel, R., Ekstrom, A.D., Kaplan, J.T., Iacoboni, M., Fried, I., 2010. Single-neuron responses in humans during execution and observation of actions. *Curr. Biol.* 20 (8), 750–756.
- Norman, K., Polyn, S., Detre, G., Haxby, J.V., 2006. Beyond mind-reading: multi-voxel pattern analysis of fMRI data. *Trends Cogn. Sci.* 10 (9), 424–430.
- Oosterhof, N.N., Wiggett, A.J., Diedrichsen, J., Tipper, S.P., Downing, P.E., 2010. Surface-based information mapping reveals crossmodal vision–action representations in human parietal and occipitotemporal cortex. *J. Neurophysiol.* 104 (2), 1077–1089.
- Oosterhof, N. N., Wiestler, T., & Diedrichsen, J. (2011). Surfing: a Matlab toolbox for surface-based voxel selection. Available from <http://surfing.sourceforge.net>.

- Oosterhof, N.N., Wiestler, T., Downing, P.E., Diedrichsen, J., 2011b. A comparison of volume-based and surface-based multi-voxel pattern analysis. *Neuroimage* 56 (2), 593–600.
- Oosterhof, N.N., Tipper, S.P., Downing, P.E., 2012. Viewpoint (in) dependence of action representations: an MVPA study. *J. Cogn. Neurosci.* 24 (4), 975–989.
- Orlov, T., Makin, T.R., Zohary, E., 2010. Topographic representation of the human body in the occipitotemporal cortex. *Neuron* 68 (3), 586–600.
- Pelgrims, B., Andres, M., Olivier, E., 2009. Double dissociation between motor and visual imagery in the posterior parietal cortex. *Cereb. Cortex* 19 (10), 2298–2307.
- Peyre, G., 2008. Toolbox Fast Marching: A Toolbox for Fast Marching and Level Sets Computations. Available from <http://www.ceremade.dauphine.fr/peyre/matlab/fast-marching/content.html>.
- Poldrack, R.A., 2006. Can cognitive processes be inferred from neuroimaging data? *Trends Cogn. Sci.* 10 (2), 59–63.
- Pylyshyn, Z., 2003. Return of the mental image: are there really pictures in the brain? *Trends Cogn. Sci.* 7 (3), 113–118.
- Ramsey, R., Cumming, J., Eastough, D., Edwards, M., 2010. Incongruent imagery interferes with action initiation. *Brain Cogn.* 74 (3), 249–254.
- Reddy, L., Tsuchiya, N., Serre, T., 2010. Reading the mind's eye: decoding category information during mental imagery. *Neuroimage* 50 (2), 818–825.
- Ricciardi, E., Bonino, D., Sani, L., Vecchi, T., Guazzelli, M., Haxby, J.V., et al., 2009. Do we really need vision? How blind people “see” the actions of others. *J. Neurosci.* 29 (31), 9719–9724.
- Rizzolatti, G., Fabbri-Destro, M., 2008. The mirror system and its role in social cognition. *Curr. Opin. Neurobiol.* 18 (2), 179–184.
- Saad, Z.S., Reynolds, R.C., Argall, B., 2004. Suma: an interface for surface-based intra- and inter-subject analysis with AFNI. *IEEE International Symposium on Biomedical Imaging: from Nano to Macro*. IEEE, Arlington VA, pp. 1510–1513.
- Saad, Z.S., Glen, D.R., Chen, G., Beauchamp, M.S., Desai, R., Cox, R.W., 2009. A new method for improving functional-to-structural MRI alignment using local Pearson correlation. *Neuroimage* 44 (3), 839–848.
- Sheffield, F.D., 1961. Theoretical Considerations in the Learning of Complex Sequential Tasks from Demonstration and Practice. National Academy of Sciences – National Research Council, Washington DC, USA.
- Shepard, R.N., Metzler, J., 1971. Mental rotation of three-dimensional objects. *Science* 171 (3972), 701–703.
- Sirigu, A., Duhamel, J.R., 2001. Motor and visual imagery as two complementary but neurally dissociable mental processes. *J. Cogn. Neurosci.* 13 (7), 910–919.
- Smith, S.M., Nichols, T., 2009. Threshold-free cluster enhancement: addressing problems of smoothing, threshold dependence and localisation in cluster inference. *Neuroimage* 44, 83–98.
- Stippich, C., Ochmann, H., Sartor, K., 2002. Somatotopic mapping of the human primary sensorimotor cortex during motor imagery and motor execution by functional magnetic resonance imaging. *Neurosci. Lett.* 331 (1), 50–54.
- Stokes, M., Thompson, R., Cusack, R., Duncan, J., 2009. Top-down activation of shape-specific population codes in visual cortex during mental imagery. *J. Neurosci.* 29 (5), 1565–1572.
- Talairach, J., Tournoux, P., 1988. *Co-planar Stereotaxic Atlas of the Human Brain. 3-Dimensional Proportional System: An Approach to Cerebral Imaging*. Thieme, New York.
- Tettamanti, M., Buccino, G., Saccuman, M.C., Gallese, V., Danna, M., Scifo, P., et al., 2005. Listening to action-related sentences activates fronto-parietal motor circuits. *J. Cogn. Neurosci.* 17 (2), 273–281.
- Vul, E., Harris, C., Winkielman, P., Pashler, H., 2009. Puzzlingly high correlations in fMRI studies of emotion, personality, and social cognition. *Perspect. Psychol. Sci.* 4 (3), 274–290.
- Welberg, L., 2008. Mirror neurons: towards a clearer image. *Nat. Rev. Neurosci.* 9, 888–889.
- Willems, R., Toni, I., Hagoort, P., Casasanto, D., 2010. Neural dissociations between action verb understanding and motor imagery. *J. Cogn. Neurosci.* 22 (10), 2387–2400.
- Wohldmann, E., Healy, A., Bourne Jr., L., 2007. Pushing the limits of imagination: mental practice for learning sequences. *J. Exp. Psychol. Learn. Mem. Cogn.* 33 (1), 254.
- Wolpert, D.M., Doya, K., Kawato, M., 2003. A unifying computational frame-work for motor control and social interaction. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 358 (1431), 593–602.