Title: Grasping with a twist: Dissociating action goals from motor actions in human frontoparietal circuits

Running Title: Decoding of isolated and task-directed grasping

Authors: Guy Rens¹,²,³, Teresa D. Figley⁴, Jason P. Gallivan⁵, Yuqi Liu⁶,⁷, Jody C. Culham¹,⁴

¹Department of Psychology, University of Western Ontario, London, Ontario, Canada
²Laboratorium voor Neuro- en Psychofysiologie, Department of Neurosciences, KU Leuven, Belgium
³Leuven Brain Institute, KU Leuven, Belgium
⁴Graduate Program in Neuroscience, University of Western Ontario, London, Ontario, Canada
⁵Departments of Psychology & Biomedical and Molecular Sciences, Centre for Neuroscience Studies, Queen’s University, Kingston, Ontario, Canada
⁶Department of Neuroscience, Georgetown University Medical Center, Washington, DC, 20057, USA
⁷Institute of Neuroscience, CAS Center for Excellence in Brain Sciences and Intelligence Technology, Chinese Academy of Sciences, Shanghai, China

*Corresponding Author:
Guy Rens
The Brain and Mind Institute
University of Western Ontario
Ontario N6A 3K7, Canada
grens@uwo.ca

Acknowledgements: This work was supported by grants from the Canadian Institutes of Health Research Operating Grant (MOP84293 to J.C.C.) and Natural Sciences and Engineering Research Council of Canada (to J.C.C.) CIHR-Doctoral Research Award 215131 (T.D.F).

Conflict of interest: The authors declare no conflicts of interest.
Abstract

In daily life, prehension is typically not the end goal of hand-object interactions but a precursor for manipulation. Nevertheless, functional MRI (fMRI) studies investigating manual manipulation have primarily relied on prehension as the end goal of an action. Here, we used slow event-related fMRI to investigate differences in neural activation patterns between prehension in isolation and prehension for object manipulation. Eighteen participants were instructed either to simply grasp the handle of a rotatable dial (isolated prehension) or to grasp and turn it (prehension for object manipulation). We used representational similarity analysis to investigate whether the experimental conditions could be discriminated from each other based on differences in task-related brain activation patterns. We also used temporal multivoxel pattern analysis to examine the evolution of regional activation patterns over time. Importantly, we were able to differentiate isolated prehension and prehension for manipulation from activation patterns in the early visual cortex, the caudal intraparietal sulcus, and the superior parietal lobule. Our findings indicate that object manipulation extends beyond the putative cortical grasping network (anterior intraparietal sulcus, premotor and motor cortices) to include the superior parietal lobule and early visual cortex.

Significance statement

A simple act such as turning an oven dial requires not only that the central nervous system encode the initial state (starting dial orientation) of the object but also the appropriate posture to grasp it in order to achieve the desired end state (final dial orientation) and the motor commands to achieve that state. Using advanced temporal neuroimaging analysis techniques, we reveal how such actions unfold over time and how they differ between object manipulation (turning a dial) vs. grasping alone. We find that a combination of brain areas implicated in visual processing and sensorimotor integration can distinguish between the complex and simple tasks during planning, with neural patterns that approximate those during the actual execution of the action.
Introduction

The hand is central in physical interactions with our environment. Typical hand-object interactions consist of sequential phases, starting with reaching and ending with object manipulation (Castiello, 2005). Electrophysiological studies in macaques have implicated a frontoparietal network in hand-object interactions (for a review see Gerbella et al., 2017). Previous functional MRI (fMRI) research (for a review see Errante et al., 2021) has identified a similar network in humans. This network comprises a dorsomedial pathway, consisting of the superior parietal occipital cortex (SPOC, corresponding to V6/V6A) and the dorsal premotor cortex (PMd), and a dorsolateral pathway, consisting of anterior intraparietal sulcus (aIPS) and ventral premotor cortex (PMv).

While past human neuroimaging studies revealed the neural substrates of grasping, most treated prehension as the end goal, not as a step towards meaningful hand-object interactions. In contrast, real-world hand-object interactions typically involve grasping only as a prelude to subsequent actions such as manipulating or moving objects. Importantly, previous studies have shown that the final action goal shapes prehension: Initial prehension strategies are affected by the goal in the ‘end-state comfort’ effect (Rosenbaum et al., 1990) and in actions like tool use (Comalli et al., 2016). Moreover, the end goal affects brain responses during action planning and prehension (Fogassi et al., 2005; Gallivan et al., 2016b).

The purpose of the current study was to investigate how isolated and sequential actions unfold differently on a moment-to-moment basis. The temporal unfolding of brain activation during actions has been largely overlooked (or only studied with EEG, without localizing the specific brain regions involved, e.g., Guo et al., 2019). Most studies that have used multivoxel pattern analysis (MVPA) to investigate hand actions have averaged data within time bins for planning and execution (e.g., Gallivan et al., 2011). Notably, several studies have revealed the temporal unfolding of MVPA grasping representations for sequential timepoints (Ariani et al., 2018a; Gallivan et al., 2013). In addition, one study examined isolated (grasping an object) vs. sequential (grasping to move an object to one of two locations) actions, showing that activation patterns could be discriminated across the grasping network even during action planning (Gallivan et al., 2016).

Here we examined (univariate) activation levels and (multivariate) activation patterns (using representational similarity analysis, RSA (Kriegeskorte, 2008). Moreover, we also used a new methodological approach, temporal MVPA (tMVPA; Ramon et al., 2015; Vizioli et al., 2018) to examine the representation similarities across trials for the same and different points in time, separately for isolated vs. sequential actions.
We measured brain activation using fMRI while participants performed a motor task consisting of either simple grasping of a dial (with two possible initial orientations) or grasping followed by rotation of the dial (clockwise or counterclockwise), as one might turn an oven dial (See Figure 1). Based on earlier findings that brain regions within the grasping network plan the full action sequence, and not just the initial grasp (see Gallivan et al., 201), we expected that during the plan phase, as well as the execute phase, tMVPA would reveal representations of the task (grasp vs. turn). Given that visual orientation (Kamitani & Tong, 2005), surface/object orientation (Rice et al., 2007; Shikata et al., 2001; Valyear et al., 2006) and grip orientation (Monaco et al., 2011) are represented in early visual cortex, the caudal intraparietal sulcus (cIPS), and reach-selective cortex (SPOC), respectively, we predicted that the representation of orientation would be biased to the start orientation early in planning, with a greater emphasis on end orientation as the movement progressed. Our approach with tMVPA allowed us to examine, across different regions, how representations unfolded over time for isolated vs. sequential actions. Specifically, we expected that regions sensitive to the kinematics of executed actions (e.g., M1 and S1) would only show highly similar representations during action execution; whereas, regions involved in more abstract features of action planning (e.g., aIPS for coding object shape) would show similar representations across planning and execution.

**Methods**

**Participants**

Data from sixteen right-handed volunteers was utilized in the analysis (7 males, 9 females, mean age: 24.4 years). Participants were recruited from Western University (London, Ontario, Canada) and provided informed consent in accordance with procedures approved by the University’s Health Sciences Research Ethics Board. Data from an additional two subjects (one male and one female) was collected but excluded due to excessive motion artifacts.

**Setup and apparatus**

The experimental setup is illustrated in Figure 1A-B. Participants lay supine in a 3-Tesla MRI scanner with the head and head coil tilted approximately 30° to allow for direct viewing without mirrors of a manipulandum positioned above the participant’s hips. The manipulandum consisted of a black rotatable dial (9-cm diameter; Figure 1B) with a yellow rectangular handle (5-cm length x 1-cm width x 2-cm depth). The dial was mounted on a black surface. The black surface was positioned such that the dial was approximately perpendicular to the subject’s line of gaze and comfortably within reach of their right hand.
arm. Two yellow markers were put on the black surface (Figure 1B) indicating the start and end positions for turning the dial. A grey platform (not shown in Figure 1A) was positioned above the participant’s lower torso serving as the home/resting position for the right arm between trials. Participants’ upper arms were braced above their torsos and just above their elbows (Figure 1A) to limit movement of the shoulder, which can induce motion artifacts in fMRI signals. As such, participants could only rely on elbow flexion/extension and forearm rotation to perform the experimental task. Considering these constraints, the position and orientation of the dial were adjusted for each participant to optimize participant comfort during task performance and ensure the dial remained fully visible. The position of the yellow markers on the black surface was adjusted for each participant individually so that dial rotation would not exceed 80% of the participant’s maximum range of motion when turning the dial clockwise or counterclockwise.

During the experiment, the dial was illuminated from the front by a bright yellow Light Emitting Diode (LED) attached to flexible plastic stalks (Figure 1A; Loc-Line, Lockwood Products, Lake Oswego, OR). A dim red LED (masked by a 0.1° aperture) was positioned approximately 15° of visual angle above the rotatable dial and just behind it to provide a fixation point for participants (Figure 1A-B). Experimental timing (see below) and lighting were controlled with in-house MATLAB scripts (The MathWorks Inc., Natick, MA).

**Experiment design and timing**

**Behavioral task.** This experiment was a 2 (starting orientation: left or right) x 2 (action: grasp or turn) delayed movement paradigm (Figure 1C). For each trial, the dial would appear in one of the two yellow-marked starting positions. The grasp condition consisted of reaching towards the dial and squeezing it between the middle phalanges of the index and middle fingers (trial end position shown in Figure 1A; hand shape during grasp shown in Figure 1B). After grasp completion, participants returned their arm back to the home position. In the turn condition, participants performed the same reach-to-grasp action but would then subsequently rotate the dial clockwise or counterclockwise after they grasped the dial in the left or right start position, respectively (Figure 1C). We decided on the index-middle finger grip instead of a more natural variant of precision grasp as this would ensure the grip would be highly similar in all conditions and limit changes to grip angle to optimize end-state comfort (e.g., putting the right-hand thumb further up when planning to turn clockwise and further down when planning to turn the dial counterclockwise). This avoids contaminating neural activation with low-level sensorimotor confounds such as digit positioning. Participants were instructed to keep the timing of all movements as similar as possible, such that the right hand reached from and returned to the home position at the same
time (see next paragraph ‘Trial Design’). To isolate the visuomotor planning response from the visual and motor execution responses, we used a slow event-related paradigm with 32-s trials, consisting of three phases: ‘Presentation’, ‘Plan’ and ‘Execute’ (see Figure 1D). We adapted this paradigm from previous fMRI studies that successfully isolated delay-period activity from transient neural responses following the onset of visual stimuli and movement execution (Beurze et al., 2007, 2009; Pertzov et al., 2011). Furthermore, using this paradigm in previous work from our group, we were able to successfully isolate and decode planning-related neural activation prior to action execution (Gallivan et al., 2011; Gallivan, McLean, et al., 2013).

**Trial design.** Before each trial, subjects were in complete darkness except for the fixation LED upon which participants were instructed to maintain their gaze. The trial began with a 6-s (3 TRs of 2 s each) Presentation phase in which the illumination LED lit up the rotatable dial. After the Presentation phase, the 12-s (6-TR) Plan phase was initiated with a voice cue (0.5-s duration) saying either ‘Grasp’ or ‘Turn’ to instruct the upcoming action to the participant. Participants could see the object during the presentation phase, thus perceiving the starting orientation of the yellow handle. However, they were instructed to only begin the action after they received the ‘Go’ cue. After the Plan phase, a 2-s (1-TR) Execute phase began with a 0.5-s beep (the ‘Go’ cue’) cueing participants to initiate and execute the instructed action. Performing the entire action, consisting of reaching and grasping (with or without dial turning), took approximately 2 s. The dial remained illuminated for 2 s after the ‘Go’ cue allowing visual feedback during action execution. After the 2-s Execution phase, the illumination LED was turned off, cueing participants to let go of the dial and return their hand back to the home position. Participants remained in this position for 12 s (6 TRs) in the dark (i.e., the intertrial interval) to allow the BOLD response to return back to baseline before the next trial would be initiated.

**Functional runs.** Within each functional run, each of the four trial types (grasp: left or right – turn: from left to right or right to left) was presented five times in a pseudo-randomized manner for a total of 20 trials. Each participant performed eight functional runs, yielding a total of 40 trials/condition (160 trials in total). Each functional run took between 10 and 11 minutes. For each participant, trial orders were counterbalanced across all functional runs so that each trial type was preceded and followed equally often by every trial type (including the same trial type) across the entire experiment. During testing, the experimenter was positioned next to the set-up to ensure that the dial was in the correct position before each trial, and manually adjust it if needed.

Participants performed a separate practice session before the actual experiment to familiarize them with the motor task and ensure proper execution. The experimental session took approximately

---

**bioRxiv preprint doi:** [https://doi.org/10.1101/2023.01.02.522486](https://doi.org/10.1101/2023.01.02.522486); this version posted January 2, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a [CC-BY 4.0 International license](https://creativecommons.org/licenses/by/4.0/).
three hours and consisted of preparation (i.e., informed consent, MRI safety, placing the participant in
the scanner and setting up the behavioral task), eight functional runs and one anatomical scan. The
anatomical scan was collected between the fourth and fifth functional runs to give participants a break
from the task.

MRI acquisition

Imaging was performed using a 3-Tesla Siemens TIM MAGNETOM Trio MRI scanner at the Robarts
Research Institute (London, ON, Canada). The T1-weighted anatomical image was collected using an ADNI
MPRAGE sequence (time to repetition (TR) = 2300 ms, time to echo (TE) = 2.98 ms, field of view = 192 mm
x 240 mm x 256 mm, matrix size = 192 x 240 x 256, flip angle = 9°, 1-mm isotropic voxels). Functional MRI
volumes sensitive to the blood oxygenation level-dependent (BOLD) signal were collected using a T2*-weighted single-shot gradient-echo echo-planar imaging (EPI) acquisition sequence (TR = 2000 ms, slice
thickness = 3 mm, in-plane resolution = 3 mm x 3 mm, TE = 30 ms , field of view = 240 mm x 240 mm,
matrix size = 80 x 80, flip angle = 90°, and acceleration factor (integrated parallel acquisition technologies,
iPAT) = 2 with generalized auto-calibrating partially parallel acquisitions (GRAPPA) reconstruction. Each
volume comprised 34 contiguous (i.e., with no gap) oblique slices acquired at an approximate 30° caudal
tilt with respect to the anterior-to-posterior commissure (ACPC) plane, providing near whole brain
coverage. We used a combination of parallel imaging coils to achieve a good signal to noise ratio and to
enable direct viewing of the rotatable dial without mirrors or occlusion. Specifically, we placed the
posterior half of the 12-channel receive-only head coil (6-channels) beneath the head and tilted it at an
angle of approximately 20°. To increase the head tilt to approximately 30°, we put additional foam padding
below the head. We then suspended a 4-channel receive-only flex coil over the forehead (Figure 1A).

fMRI anatomical data processing

All fMRI preprocessing was performed in BrainVoyager version 22 (Brain Innovation, Maastricht
Netherlands). For the present study, we defined regions of interest (ROIs) in surface space instead of
volumetric space as it has been shown that cortical alignment improves group results by reducing
individual differences in sulcal locations (Fischl et al., 1999; Frost & Goebel, 2012). As such, we performed
surface reconstruction (mesh generation) and cortex-based alignment. Given that we relied on the
recommended approach and standard settings of BrainVoyager 22, these steps are explained only in brief.

Folded mesh generation for each participant. In volumetric space we performed the following
steps. First, we corrected for intensity inhomogeneity. Then, we rotated the anatomical data in ACPC
space, due to the experimental head tilt, and normalized to Montreal Neurological Institute (MNI) space. Next, we excluded the subcortical structures (labelled as white matter) and the cerebellum (removed from anatomical) from mesh generation. We then defined the boundaries between white matter and gray matter and between gray matter and cerebrospinal fluid. Finally, we created a folded mesh (surface representation) of only the left hemisphere after removing topologically incorrect bridges (Kriegeskorte & Goebel, 2001). We decided to only investigate the left hemisphere because the motor task involved the right hand only, which previous work has shown predominantly activates the left (contralateral) hemisphere (Cavina-Pratesi et al., 2010, 2018; Gallivan et al., 2011; Gallivan, McLean, et al., 2013).

**Standardized folded mesh generation for each participant.** Briefly, folded meshes created from anatomical files often result in different numbers of vertices between participants. To facilitate cortex-based alignment between participants, folded meshes were first transformed into high-resolution standardized folded meshes. Each folded mesh was first morphed into a spherical representation by smoothing (thus removing differences between sulci and gyri) and correcting for distortion. The spherical representation of each participant mesh was then mapped to a high-resolution standard sphere to create a high-resolution standardized spherical representation of the participant mesh. The vertex position information of the original participant folded (not spherical) mesh was then used to generate a standardized folded mesh for each participant.

**Cortex-based alignment.** Cortex-based alignment was performed following the approach of Frost & Goebel (2012) and Goebel et al. (2006): We aligned all individual standard meshes to a dynamically generated group average target mesh. Before aligning to the dynamic group average, we performed pre-alignment (i.e., rigid sphere alignment). The actual steps of cortex-based alignment generate a dynamic group average (a surface mesh based on all individual meshes) and sphere-to-sphere mapping files for each participant that enable transporting the functional data from each individual to the dynamic group average. Inverse sphere-to-sphere mapping files were also generated, which allows transporting of data (such as regions of interest) from the dynamic group average back to individual meshes.

**fMRI functional data processing**

**General preprocessing.** All functional runs were screened for motion and magnet artifacts by examining the movement time courses and motion plots created with the motion correction algorithms. Based on this screening, data from two participants was discarded.

Functional runs were co-registered with the anatomical data using boundary-based registration (Greve & Fischl, 2009). We subsequently performed slice-scan time correction, motion correction and MNI...
normalization. Next, we performed linear-trend removal and temporal high-pass filtering (using a cut-off of three sine and cosine cycles on the fast Fourier transform of the time courses).

Volumetric to surface-based time courses. Volumetric time courses were first aligned with the respective participant’s anatomical scan using boundary-based registration to ensure optimal alignment (Greve & Fischl, 2009). Volumetric time courses were then transformed to the respective participant’s standardized folded mesh using depth integration along the vertex normal (sampling from -1 to 3 mm relative to the gray-white matter boundary). Once functional runs were transformed into surface space, they were spatially smoothed. Finally, functional runs in individual standard mesh space, could then be mapped onto the dynamic group average mesh using sphere-to-sphere mapping files that were generated during cortex-based alignment. This approach enabled us to model the averaged brain activation across all participants onto the group mesh and define ROIs based on hotspots of strongest activations in surface space.

Defining regions of interest

To localize our regions of interest on the group mesh, we applied a general linear model (GLM) on our data in surface space. Predictors of interest were created from boxcar functions convolved with the two-gamma haemodynamic response function (HRF). We aligned a boxcar function with the onset of each phase of each trial within each run of each participant, with a height dependent upon the duration of each phase. We used 3 TRs for the presentation phase, 6 TRs for the Plan phase and 1 TR for the Execution phase. The six motion parameters (three translations and three rotations) were added as predictors of no interest. Each incorrect trial was also assigned a unique predictor of no interest. All regression coefficients (beta weights) were defined relative to the baseline activity during the intertrial interval. In addition, time courses were converted to percent signal change before applying the random effects GLM (RFX-GLM).

To specify ROIs for our analyses, we searched for brain areas, on the group level, involved in the experimental task. We contrasted brain activation for all three phases with baseline. This contrast was performed across all conditions to ensure that ROI selection based on local activation patterns was not biased by differences between specific conditions (e.g., grasp versus turn). Specifically, the contrast was:

\{\text{Presentation } [\text{Grasp left} + \text{Grasp right} + \text{Turn left to right} + \text{Turn right to left}] + \text{Plan } [\text{Grasp left} + \text{Grasp right} + \text{Turn left to right} + \text{Turn right to left}] + \text{Execute } [\text{Grasp left} + \text{Grasp right} + \text{Turn left to right} + \text{Turn right to left}]\} > \text{baseline}.\]

This contrast enabled us to identify regions that showed visual and/or motor activation associated with the task.
In BrainVoyager, area selection on a surface generates a hexagon surrounding the selected vertex. We decided on an area selection size of 100 (arbitrary units) as this value provided a good balance between inclusion of a sufficient number of vertices/voxels around each hotspot for multivariate analyses and avoiding overlap between ROIs (especially in the parietal lobe).

Regions of interest
Selected regions of interest (Figure 2) were defined in the left hemisphere. Most regions of interest were defined using the contrast above (all conditions > baseline), with two exceptions, as detailed below.

First, the involvement of the cortical grasping network (Gerbella et al., 2017) was investigated by including sensorimotor and visuomotor regions (Ariani et al., 2018; Cavina-Pratesi et al., 2018; Fabbri et al., 2016; Gallivan et al., 2011; Gallivan, McLean, et al., 2013).

- Primary motor cortex (M1): Hotspot of strongest activation near the ‘hand knob’ landmark (Gallivan et al., 2013; Yousry et al., 1997).
- Dorsal premotor cortex (PMd): Hotspot of strongest activation near the junction of the precentral sulcus and the superior frontal sulcus (Picard & Strick, 2001; Pilacinski et al., 2018).
- Ventral premotor cortex (PMv): Hotspot of strongest activation inferior and posterior to the junction of the inferior frontal sulcus (Tomassini et al., 2007).
- Primary somatosensory cortex (S1): Hotspot of strongest activation anterior to the anterior intraparietal sulcus, encompassing the postcentral gyrus and sulcus (Gallivan et al., 2011; Gallivan, McLean, et al., 2013)
- Anterior intraparietal sulcus (aIPS): Hotspot of strongest activation directly at the junction of the intraparietal sulcus and the postcentral sulcus (Culham et al., 2003).
- Caudal intraparietal sulcus (cIPS): Hotspot of strongest activation on the lateral side of the brain, anterior and superior to the junction between the intraparietal sulcus and the posterior occipital sulcus (Beurze et al., 2009; Greffkes & Fink, 2005). Comparisons with the Julich atlas (Richter et al., 2019) suggested overlap with hIP7 and/or hP01.
- Anterior superior parietal occipital sulcus (aSPOC): Hotspot of strongest activation on the medial side of the brain, anterior and superior to the parietal-occipital sulcus (Cavina-Pratesi et al., 2010), thought to correspond to area V6A (Pitzalis et al., 2015).
- Posterior superior parietal occipital sulcus (pSPOC): Hotspot of strongest activation on the medial side of the brain, posterior and inferior to the parietal-occipital sulcus (Cavina-Pratesi et al., 2010), thought to correspond to area V6 (Pitzalis et al., 2015).
Second, the following medial and frontal regions were selected due to their involvement in motor planning and decision making (Ariani et al., 2015; Badre & Nee, 2018; Cavina-Pratesi et al., 2018).

- Dorsolateral prefrontal cortex (DLPFC): Hotspot of strongest activation near the middle frontal gyrus (Mylius et al., 2013).
- Supplementary motor area (SMA): hotspot of strongest activation adjacent to the medial end of the cingulate sulcus and posterior to the plane of the anterior commissure (Picard & Strick, 2001).
- Pre-supplementary motor area (pre-SMA): Hotspot of strongest activation superior to the cingulate sulcus, anterior to the plane of the anterior commissure and anterior and inferior to the hotspot of strongest activation selected for SMA (Picard & Strick, 2001).

Third, we included the superior parietal lobule (SPL) due to the extensive activation evoked by our task and in hand actions more generally (Ariani et al., 2018a; Cavina-Pratesi et al., 2018). We defined the SPL as the area on the lateral/superior side of the brain that is bordered anteriorly by the postcentral sulcus, inferiorly by the intraparietal sulcus and posteriorly by the parietal-occipital sulcus (Scheperjans, et al., 2008a; Scheperjans et al., 2008b). Due to the large swathe of activation evoked by our contrast, we defined four ROIs within the SPL based on their relative position to each other (to ensure minimal overlap) and the anatomical landmarks bordering the SPL. We decided on four ROIs as well as their names based on Scheperjans et al. (2008a).

- 7PC: Hotspot of strongest activation located on the posterior wall of the postcentral sulcus and superior to the intraparietal sulcus. Given we defined post-AIPS as well, 7PC was also defined as superior to post-AIPS.
- 5L: Hotspot of strongest activation located just posterior to the postcentral sulcus and superior to area 7PC.
- 7P: Hotspot of strongest activation superior to the intraparietal sulcus and anterior to the parietal-occipital sulcus.
- 7A: Hotspot of strongest activation in the postcentral gyrus, superior to 7PC, posterior to 5L and anterior to 7P.

Finally, two visual regions were selected from a probabilistic functional atlas that utilized cortex-based alignment (Rosenke et al., 2021), to which we aligned our participant surface meshes.

- Dorsal primary visual cortex (V1d): Given that primary visual cortex (V1) responds to visual orientation (Kamitani & Tong, 2005a) we wanted to investigate its response here. Because our target objects (and participants hands) fell within the lower visual field (below the fixation point),
they would stimulate the dorsal divisions of early visual areas (Wandell et al., 2007). As such, we investigated only the dorsal division of V1, V1d, using the Rosenke atlas.

- Extrastriate body area (EBA): We were also interested in examining the response of the extrastriate body area, which has been implicated not only in the visual perception of bodies (Greenfield et al., 1996) but also in computing goals during action planning (Astafiev et al., 2004; Zimmermann et al., 2016). Because the EBA is not easy to distinguish from nearby regions of the lateral occipitotemporal cortex, we utilized the EBA from the Rosenke atlas.

After ROIs were defined in group surface space, they were transformed to individual surface space, using the inverse transformation files generated during cortex-based alignment and then to individual volumetric MNI space, using depth expansion (inverse of depth integration; see ‘fMRI functional data processing: Volumetric to surface-based time courses’) along the vertex normals (-1 to 3 mm). This approach allowed us to define ROIs on the group surface but extract functional data from the individual volumetric level.

**Analysis of functional data**

Functional data were extracted to perform deconvolution general linear models (deconvolution GLMs; Hinrichs et al., 2000), representational similarity analysis (Kriegeskorte, 2008) and temporal multivoxel pattern analysis (tMVPA; Vizioli et al., 2018). All analyses described below, excluding initial processing of fMRI data and the univariate analyses (which where both done in BrainVoyager), were done with in-house MATLAB scripts.

**Univariate analysis.** We used a random-effects GLM with deconvolution to extract the time courses of activation in each ROI during the experimental task. For each experimental condition, we used 15 matchstick predictors, the first of which was aligned with time point zero (Figure 1D; “light on”). For each functional voxel in each ROI, baseline z-normalized estimates of the deconvoluted BOLD response, representing the mean-centered signal for each voxel and condition relative to the standard deviation of signal fluctuations, were extracted for all conditions. Voxel activation was then averaged across voxels within each ROI. This provided us with an estimate of the averaged time course of brain activation for each condition for each ROI and for each participant.

**Temporal multivoxel pattern analysis.** For tMVPA we relied on the approach developed in Ramon et al. (2015) and Vizioli et al. (2018). Briefly, tMVPA has been developed to investigate the temporal development of neural representations by relying on multivariate analyses with a trial wise approach. The tMVPA methods are schematized in Figure 3. For each voxel of each ROI of each participant, we performed
a deconvolution GLM for every trial separately, which was then transformed into BOLD percent signal change by dividing the raw BOLD time course by its mean. Note that for tMVPA we performed deconvolution GLMs for each trial (of the same condition) separately whereas for the univariate analysis and RSA (see below) we ran deconvolution GLMs for each condition (i.e., across same-condition trials). After computing the single-trial activation patterns, we computed single-trial representational dissimilarity matrices (stRDMs; dissimilarity = 1 - r) using Pearson correlation for each condition of each ROI of each participant. As shown in Figure 3A, a stRDM is generated by computing dissimilarity between the activation patterns of voxels at each timepoint of a given trial (‘trial m’) with the activation pattern of the same voxels for each of the timepoints for another trial (‘trial n’). This process is iteratively repeated until all unique within-condition trial pairings are run. stRDMs were then averaged across the main diagonal to yield a diagonally symmetric matrix. We performed this averaging as we were only interested in between-condition differences and not in differences between same-condition trials.

To further clarify calculation of the dissimilarity metric: in Figure 3B and Figure 3C, the green highlighted square indicates the dissimilarity between all the voxels (of a given ROI) at timepoint 2 of Turn Left to Right trial m with the activation of the same voxels at timepoint 2 of Turn Left to Right trial n. The magenta highlighted squares indicate the averaged dissimilarity between timepoints 2 and 10 of trial m and n. As explained before, we averaged the dissimilarity between timepoint 2 of trial m and timepoint 10 of trial n (magenta highlighted square below diagonal) with the dissimilarity between timepoint 10 of trial m and timepoint 2 of trial n (magenta highlighted square above diagonal). Finally, the blue highlighted square indicates the dissimilarity between timepoints 10 of trial m and n. In sum, values on the main diagonal show dissimilarity between within-condition trials at the same timepoint. Values that are off the main diagonal show the dissimilarity between within-condition trials at different (i.e., earlier or later) timepoints.

stRDMs were calculated for the four conditions separately, Fisher z-transformed and then averaged (10 % trimmed mean; Vizioli et al., 2018) within conditions (e.g., average RDM for grasp left). Finally, the averaged RDMs were then averaged across orientations (e.g., averaging of RDMs between left and right start orientations to produce average Grasp and Turn RDMs; Figure 3D and 3E respectively). This was done as we were primarily interested in investigating differences in neural representations between singular (grasping) and sequential actions (grasping then turning) and we did not expect that these differences would depend upon the start orientation. As such, this approach resulted in two RDMs for each participant.
In line with Vizioli et al. (2018), we performed statistical analysis on the Fisher z-transformed data; however, we used the non-transformed data for visualization purposes to render the values visually more interpretable. To test for statistically significant differences between the grasp and turn RDMs, we subtracted for each participant the turn RDM from the grasp RDM resulting in a subtraction matrix (Turn versus Grasp: Figure 3F) and investigated where the subtraction differed significantly from zero. As such, a given value in the subtraction matrix that is positive indicates that grasping trials are more dissimilar than turning trials at that given timepoint. Conversely, a value that is negative indicates that turning trials are more dissimilar than grasping trials at that given timepoint. Note that we decided to statistically test whether each cell differed from zero instead of the increasingly larger sliding window analysis used in Ramon et al. (2015) and Vizioli et al. (2018). Our rationale was that the earlier studies relied on a visual task (face recognition) whereas our study relied on a motor task. Arguably, during motor preparation/execution, neural representations might evolve differently than purely visual responses. As such, we argued that statistically testing each cell separately might reveal more information on the temporal evolution of motor execution as, for instance, transitions between activation patterns between planning and execution might be brief or abrupt. Note that because we had averaged values across the main diagonal when calculating the dissimilarity matrices (Figure 3 and 5), these matrices were symmetrical reflections across the diagonal. However, we performed statistical analyses only on the values on and above the main diagonal of the Turn versus Grasp subtraction matrices (Figure 3F). To test for statistical significance, we performed (1 - alpha) bootstrap confidence-interval analysis by sampling participants with replacement 500 times. Within each ROI, we accounted for multiple comparisons using Bonferroni correction. Importantly, we decided to exclude the first two timepoints (TRs) and the last one from the statistical analyses because no differences were expected before the BOLD response to emerge at the start, because later timepoints reflect the post-stimulus undershoot phase of the BOLD response, and to reduce the number of comparisons requiring Bonferroni correction for multiple comparisons.

Hierarchical clustering. The tMVPA matrices (stRDMs) for Grasp and Turn revealed that different ROIs showed different temporal unfolding. For example, some regions showed similar activation patterns throughout the entire trial; whereas, other regions showed similar patterns only during motor execution (See Figure 5). To make it easier to group these different ROI timing patterns for display and discussion, we used hierarchical clustering. To do so, we averaged within-participants the grasp and turn RDMs to generate one RDM per participant. Next, we calculated for each participant the Spearman correlations between all ROIs using the averaged RDMs. The ROI correlation matrix for each participant was then
transformed into a dissimilarity measure (1-\(r\)). Finally, hierarchical clustering was then performed using the ROI dissimilarity matrix of all participants.

*Representational similarity analysis.* We used RSA to examine the degree to which the pattern of activation across voxels within each ROI at each time point represented the start orientation, the end (goal) orientation, and the task, as illustrated in Figure 4A-C. For RSA, we relied on the methods described in Kriegeskorte (2008). By examining the degree to which different conditions evoke a similar pattern of brain activation within a ROI, the nature of neural coding (or representational geometry; Kriegeskorte & Kievit, 2013) can be assessed. First, we utilized the deconvolution analysis without averaging across voxels within ROIs. For each voxel, the average activation was normalized to a mean of zero by subtracting the mean across conditions from the value of each condition in that voxel (Haxby et al., 2001). This was done for each run separately. This resulted in average temporal development of brain activation for each condition for each voxel of each ROI and for each run of each participant. Then, we computed representational dissimilarity matrices (RDMs) for each timepoint for each ROI of each participant. By computing a Pearson correlation of voxel activation patterns within each ROI (dissimilarity = 1 - \(r\)), the RDMs can quantify the dissimilarity in activation patterns of one condition with different conditions. To test within-subject reliability of the RDM calculations, we used cross-validation by splitting data into all potential combinations. Split data provides data-based estimates of dissimilarity even for the same condition (e.g., Grasp left in one data split vs. Grasp left in another data split); whereas, in unsplit data, the dissimilarity is necessarily zero. Data for all cells in the RDM is necessary for RSA on factorial designs to ensure that contrasts are balanced across orthogonal factors. Pearson correlations (\(r\)) were computed between the activation patterns across voxels in each condition during odd and even runs and then Fisher transformed to have a similarity metric with a Gaussian distribution.

To test whether each region contained information about condition differences, we measured correlations between the RDM in each region (Figure 4D-E) and three separate models that capture orthogonal components of the experimental task (Figure 4A-C). This was done for each timepoint separately (15 matchstick predictors; time-resolved decoding). We used three models that capture specific task attributes: (A) Start orientation, irrespective of task (i.e., grasp or turning) — a timepoint with an RDM that correlates with this model would indicate an encoding of the initial orientation of the handle (i.e., left, or right), (B) end orientation, irrespective of the performed action (i.e., grasp or turning) — a timepoint with an RDM that correlates with this model would indicate an encoding of the final orientation of the handle (i.e., left, or right) after task execution, (C) motor task, irrespective of the initial/final handle’s orientation (i.e., left, or right) — a timepoint with an RDM that correlates with this model would
indicate an encoding of the task goal task (i.e., grasping or turning). The metric was calculated by computing the Spearman correlation ($\rho$) between the Fisher-transformed split RDMs and each model for each ROI for each participant. We used one-way Student’s $t$ tests to assess whether model correlations were significantly greater than zero and performed corrections for multiple comparisons using the false discovery rate. Corrections for multiple corrections were done for each ROI separately (i.e., 15 timepoints for three models compared to zero for 45 comparisons in total). For each ROI we also calculated the upper and lower bound of the noise ceiling. Briefly, the noise ceiling is the expected RDM correlation achieved by the (unknown) true model, given the noise of the data and provides an estimate of the maximum correlations that could be expected for a given model in a given ROI (Kriegeskorte, 2008). RDMs were first rank transformed, following the original z-transformation. The upper bound of the noise ceiling, considered an overestimate of the maximum correlation, was calculated as iteratively correlating one participant’s RDM with the average RDM of all participants (thus including the given participant) and then averaging across all participants. The lower bound of the noise ceiling, considered an underestimate of the maximum correlation, was calculated as iteratively correlating between one participant’s RDM with the average RDM of all other participants (thus excluding the given participant), and then averaging across participants.

**Results**

Qualitative examination of the deconvolution time courses (Figure 2) and the tMVPA results indicated that different regions had different temporal profiles of activity. As shown in Figure 5, hierarchical clustering of the RDMs averaged across Grasp and Turn facilitated conceptual grouping of the ROIs for further investigation.

Notably, ROIs differed in the time ranges over which they showed reliable activation patterns across trials of the same type. Most strikingly, V1d and cIPS showed strong temporal similarity in activity patterns throughout the Plan and Execute phases of the trial, as indicated by the large block of high similarity (dark red) beginning early in planning and continuing through late execution. That is, in these regions, voxel activation patterns were not only similar during the same phase (along the diagonal cells) but were also similar across timepoints throughout planning and execution (off-diagonal cells in the planning and execution phases). This indicates consistency in the neural representation throughout the trials. Similar but weaker similarity patterns were also observed in other regions, particularly aSPOC, pSPOC, 7P, EBA, 7A, 7PC, and 5L.
In contrast, other regions -- aIPS, M1, S1, pre-SMA, SMA, PMd -- showed trial-consistent activation patterns predominantly for the peak execution period (dark red blocks in the execution phase; i.e., in lower right corners of Grasp and Turn matrices). In some cases, there was also some consistency of patterns between the peak execution phase and earlier timepoints during Plan and Execute of which M1 is a clear example in Figure 5 (off-diagonal cells between planning and execution phase, as indicated by the reddish “wings” above and left of the peak similarity). PMv and DLPFC showed only weak consistency of patterns across trials, highest in the peak of execution.

Interesting patterns were also observed in the tMVPA difference in trial-by-trial consistency between turn and grasp (final column of Figure 5). Notably many of the regions showed higher consistency during Turn actions than Grasp actions.

Based on these observations from Figure 5, we focussed subsequent analysis and interpretation on a subset of ROIs in Figure 6.

Visual regions of V1d and cIPS

Qualitative assessment of the univariate analysis. The first column in Figure 6 shows the deconvolution time courses for V1d and cIPS. V1d and cIPS show a steep increase in average voxel activation following object presentation (timepoint 0). This initial increase is followed by a second smaller increase halfway through the planning phase and, last, by a third and final increase following action execution. Although V1d and cIPS are primarily visual regions, the univariate analysis shows that activation does not remain constant throughout the visual presentation of the stimuli, but is modulated by the dynamics of the task.

Qualitative assessment of the condition RDMs. In Figure 6 the second and third columns show the condition RDMs for grasp and turn, respectively. For both regions, activation pattern similarity within conditions strongly increases following object presentation. Due to the sluggish nature of the hemodynamic response and our findings for the univariate analysis, it is plausible that this early increase of similarity in the activation patterns is primarily driven by a visual response. Notably, the cIPS within-condition similarity in activation patterns increases towards the peak execution (approx. timepoint 11-13).

Temporal multivoxel pattern analysis. The fourth column in Figure 6 depicts the subtraction Turn minus grasp on which we performed statistical analyses as previously described. For V1d, tMVPA is unable to discriminate the grasp and turning condition for matched timepoints (along the diagonal) during motor planning. That is, subtracting the grasp dissimilarity from the turn dissimilarity (second and third column...
in Figure 6) does not result in a difference that differs statistically from zero. Interestingly, tMVPA is able
to discriminate the activation patterns of grasping and turning during late execution (i.e., the “plus sign"
significant squares in the grasp versus turn RDM for V1d). In addition, tMVPA was also able to discriminate
activity patterns related to grasping versus turning when dissimilarity calculations were done between
timepoints of the late planning phase and timepoints of the late execution phase (i.e., the three adjacent
significant squares in the grasp versus turn RDM for V1d). These findings suggest that the activation
patterns in V1d become tuned towards either grasping or turning during late motor planning. Similar but
stronger effects were found for cIPS. tMVPA was able to discriminate cIPS activation patterns during most
of the execution phase as well as when dissimilarity calculations were done between timepoints of the
execution phase and early and late timepoints of the planning phase (i.e., the vertical and horizontal
“wings” of significant squares above and left of the execution phase). This suggests that in cIPS,
representations of the task emerge early in the planning period.

Representational similarity analysis. For the RSA, we tested how well multivariate activation
patterns fit three models for the representation of start orientation, end orientation, and task (Turn versus
Grasp) over time (Figure 6; last column). Notably, although all three types of information appear to be
represented in both V1d and cIPS during action execution, only starting orientation appeared to be
represented during the planning phase (Figure 6; last column; V1d and cIPS; green trace). Although the
decoding of start orientation during planning and end orientation during execution may reflect the
processing of simple visual orientation information (Kamitani & Tong, 2005), coding of both start and end
orientation overlapped during execution (Figure 6; last column; V1d and cIPS; green and red trace),
suggest a much more complex representation. Notably, V1d, represented the task during late planning,
before any action had been initiated, perhaps due to anticipation of the visual consequences of the
upcoming action.

In sum, while our findings are consistent with a role for V1d and cIPS as predominantly visual
regions, the two multivariate analyses (tMVPA and RSA) suggested that these regions discriminate
between grasping and turning actions not only during action execution but also during motor planning.

Regions of the SPL: 7A, 7PC, 5L and 7P

Qualitative assessment of the univariate analysis. As shown in the first column in Figure 6, our
results suggest multiple ROIs within the SPL have a visuomotor involvement in hand-object interactions.
All four regions show three distinct peaks, two of comparable amplitude following the onset of visual
stimulation and the planning instruction, with a larger peak during action execution. This change in
average activation is more distinctly visible for regions 7A and 5L as their activation peaks higher during
the execution phase compared to the regions 7PC and 7P.

Qualitative assessment of the condition RDMs. The condition RDMs of Figure 6 show that the ROIs
within the SPL show activity patterns that are similar to each other. First, similarity between trials within
each condition increases following object presentation, in particular, during early planning. Second,
similarity in activation patterns is then maintained until a second increase during the execution phase.
Notably, this pattern (in the condition RDMs of the SPL regions) is similar to that of V1d and cIPS, albeit
much weaker. Most importantly, the increase in within-condition trial similarity following object
presentation is much weaker for the SPL regions than for V1d and cIPS. This suggests that these regions
encode visual information to a lesser extent than V1d and cIPS and might be more involved in visuomotor
transformation of visual cues into executable motor commands.

Temporal multivoxel pattern analysis. In line with the qualitative assessments of the univariate
analysis and the condition RDMs, tMVPA suggests a role for the regions of the SPL throughout the trial,
but particularly during motor execution (Figure 6; column 4). For 7A, tMVPA was able to discriminate
activation patterns for grasping and turning during the later execution phase. Interestingly, we found
similar results for 7PC as for cIPS (as explained in “Visual regions of V1d and cIPS”). tMVPA was able to
discriminate grasping and turning throughout both earlier and later stages of the execution phase as well
as when dissimilarity calculations were done between timepoints of the planning phase and timepoints
of the execution phase. These findings indicate that 7PC might anticipate turning actions already during
the planning phase. The results for 5L are somewhat similar to the findings for 7A and 7PC. That is, tMVPA
is able to differentiate between grasping and turning during the later execution phase but also when
dissimilarity was calculated on timepoints of the early/late planning phase and timepoints of the
execution phase. For 7P, tMVPA could significantly discriminate activation patterns associated with
grasping and turning only during the execution phase. Finally, it can be seen that tMVPA can discriminate
between grasping and turning by relying on the activation patterns of each of the four regions. However,
this ability to discriminate the task is the weakest in 7PC, stronger in 7A and 5L and the strongest in 7PC.

Representational similarity analysis. The RSA results (Figure 6 column 5) further support that the
regions of the SPL are rather visuomotor than purely visual: starting orientation cannot be decoded from
the SPL (Figure 6; last column; 7P, 7A, 7PC and 5L; green trace) during the planning phase as for V1d and
cIPS. In support of our tMVPA results, grasping versus turning can be decoded from all four regions in the
SPL (Figure 6; last column; 7P, 7A, 7PC and 5L; blue trace). RSA is primarily able to do this when relying on
the execution phase significant decoding is also observed during the late planning phase for the regions
Interestingly, in 7PC only turning versus grasping can be decoded, suggesting that this region might be mainly involved in rotating the wrist, irrespective of the direction. Indeed, from the other regions (7A, 5L and 7P) both start and end orientation can be decoded during the action execution phase suggesting that these regions encode specific wrist orientations throughout the hand-object interaction.

In sum, our multivariate analyses provide evidence for the involvement of the SPL in decoding sequential actions. That is, sequential (grasping then turning) actions can be differentiated from singular grasping actions based on the activation patterns from the regions 7A, 7PC, 5L and 7P, in some cases even before action execution.

EBA and aIPS

Qualitative assessment of the univariate analysis. As shown in Figure 6, aIPS shows a modest visual response, with a distinct “bump” in the planning phase and following execution, indicating involvement in both motor planning and execution. EBA shows an initial visual response following object presentation after which activation wanes until action execution. However, following action execution, activation increases again suggesting an involvement of EBA in action execution, perhaps related to the visualization of the arm and hand with respect to the object.

Qualitative assessment of the condition RDMs. In line with the univariate analysis, the condition RDMs for aIPS show that similarity in activation patterns remains very low until motor action and then sharply increases following action execution. As such, the condition RDMs indicate that aIPS is primarily involved in action execution. In the condition RDMs for EBA, a clear visual response can be seen as within-condition similarity increases early during trials. Interestingly, in line with the univariate analysis, within-condition similarity increases again following task execution further supporting the notion of EBA’s involvement in motor execution and/or visual feedback of the hand.

Temporal multivoxel pattern analysis. tMVPA can discriminate activation patterns in aIPS related to grasping and turning during the early execution phase, again supporting the notion that aIPS is involved in the execution of hand-object interactions. Interestingly, tMVPA was able to discriminate activation patterns related to grasping and turning during the planning phase as well as when dissimilarity was calculated between timepoints of the planning phase and timepoints of the execution phase. EBA showed sporadically significant differences between the two tasks during the early visual response and planning.

Representational similarity analysis. From aIPS, RSA could discriminate the activation patterns associated with grasping and turning during the planning phase as well as during the execution phase (Figure 6; last column; aIPS; blue trace). For EBA, the RSA revealed an early visual response as it was able
to decode starting orientation from the activation patterns following object presentation (Figure 6; last column; EBA; green trace). Interestingly, in with the previous analyses, RSA was able to decode grasp versus turn in both the planning phase as well as the execution phase (Figure 6; last column; EBA; blue trace).

In sum, our findings for aIPS are in line with previous studies that showed its involvement in motor planning and execution (e.g., Singhal et al., 2013). Interestingly, our multivariate analyses support the suggestion that EBA, perhaps along with hand-selective divisions of the lateral occipitotemporal cortex, is not a purely visual region but is also involved in analyzing visual feedback of the body and hand during motor execution (Astafiev et al., 2004; van den Heiligenberg et al., 2018; Zimmermann et al., 2016).

Other regions of interest

We focused the presentation of the results above in the most robust and interesting findings about how hand-object interactions unfold over time. Other regions showed effects that were weaker or less surprising. The condition RDMs and tMVPA results for the previously discussed ROIs and other ROIs can be found in Figure 6. Full results figures for all ROIs can be found on GitHub (https://github.com/GuyRens/OvenDials).

Briefly, we found no robust effects for the parietal regions of pSPOC and aSPOC which have been considered to be typically involved in hand-object interactions. A similar absence was found for the DLPFC and for the secondary motor regions PMv, SMA and pre-SMA. However, tMVPA was able to discriminate activation patterns related to grasping and turning during the execution phase from both PMd and S1 which is in line with previous work suggesting their involvement in the execution of hand-object interactions. Finally, for PMd and S1, tMVPA was able to discriminate grasping and turning activation patterns during the early execution phase.

Discussion

Our results (summarized by the outlined circles in Figure 2) provide new insights into how neural representations over time for a simple prehension task vs. a more complex manipulation task. While prior work has shown that different types of grasp sequences can be decoded (Gallivan et al., 2016), the current study examines the temporal coding of this information across planning and execution, as well as the nature of the information being represented in individual brain regions over time. In addition, our study examines the activity of regions within the SPL, not previously explored in the prior work. Our additions
corroborate earlier findings (Gallivan et al., 2016, 2019) that action sequences are represented in visual areas (such as V1 and EBA), but also implicate subdivisions of the SPL (areas 7A, 7PC, 7P and 5L).

Using RSA, we found that while some regions, particularly V1, coded start orientation during planning, multiple regions (V1, cIPS, SPL and EBA) coded a combination of start orientation, end orientation and task during execution. This finding suggests that the activation patterns in these areas reflect more than simple visual-perceptual information such, as the visual orientation of the dial (Kamitani & Tong, 2005); instead, these patterns reflect the combination of the initial goal, the motor specific act, and the final outcome. Moreover, the similarity of activation patterns across planning and execution, as identified through our tMVPA, suggests that an action and its outcome are anticipated well before the movement begins.

Temporal Unfolding of Complex vs. Simple Actions

We investigated the temporal unfolding of complex vs. simple actions using multiple measures of neural responses: univariate time courses (Figure 1), RSA for sequential volumes (Figures 4 and 6) and, tMVPA (Figures 3, 5 and 6). These different measures provide a fuller picture of how actions unfold over time than considering each approach in isolation. For example, univariate signals reveal that both V1 and cIPS show robust visual signals in all three phases of the trial (presentation, plan and execute), as do SPL regions and EBA to a lesser degree. tMVPA shows consistency in the patterns of activation in these areas throughout the trial. However, RSA shows that different areas are representing different kinds of information at different points in the trial. For example, V1 represents start orientation early in the trial and task later, while 7P represents both pieces of information only just before execution begins. Moreover, tMVPA shows that trial similarity remains high in all these regions through both Grasp and Turn trials, but trial-by-trial similarity during execution and between execution and planning is higher for Turn than Grasp trials.

These results build upon earlier work examining univariate and multivariate time courses during grasping (Ariani et al., 2018b; Gallivan et al., 2016a). They also provide a valuable complement for EEG studies that reveal temporal coding with even finer temporal resolution but poorer spatial resolution. An EEG study by (Guo et al., 2019) showed that grasp orientation (defined by instruction rather than object attributes) can be classified during both a visual preview and action execution, with similar representations between the two phases; moreover, their most informative electrodes were over left caudal parietal cortex, though source localization indicated diverse potential sources (that may include cIPS, SPL, and EBA).
One of the more surprising results of the current study was the strong involvement of SPL in showing differences between Turn and Grasp actions. The role of the SPL in hand actions has been relatively less studied than areas in the IPS (especially cIPS and aIPS) and SPOC (V6/V6A), for which clear homologies between humans and other primates have been proposed (Culham & Kanwisher, 2001; Grefkes & Fink, 2005). One partial exception is area 7PC, an SPL region dorsal to aIPS, that has been proposed as the human homologue of the macaque medial intraparietal sulcus (mIPS) and/or parietal reach region (PRR), albeit with limited consensus about locus (Gallivan & Culham, 2015) and putative homologies (Culham et al., 2006). These SPL areas likely form part of the dorsomedial (or dorsal-dorsal) visual stream (Rizzolatti & Matelli, 2003) and have been postulated as an integration zone for visual, motor, and somatosensory information for goal-directed movements (Passarelli et al., 2021). Human and macaque SPL have similar structural organization with the anterior sector -- 7PC (macaque homologue: VIP; Caminiti et al., 2015) and 5L (macaque homologue: PE; Gamberini et al., 2020) -- being predominantly somatosensory and the more caudal parts -- 7A (macaque homologue: PEC; Gamberini et al., 2020) and 7P (macaque homologue: V6A; Gamberini et al., 2020) -- being both somatosensory and visual (Gamberini et al., 2020). The anterior regions of the SPL (5L and 7PC) fall on the more somatosensory side of the visual-to-somatosensory gradient from the posterior to anterior SPL. As such, they may be involved in anticipating the sensory (perhaps largely somatosensory) and motor requirements and consequences of the action.

**Orientation Processing**

Our experiment explicitly involved tasks that required processing of object orientation for grasping and turning. While processing of object location for reaching and object shape and size for grasping have been well characterized behaviorally and neurally (Cavina-Pratesi et al., 2010), the influence of object orientation on reach-to-grasp actions has only recently been investigated in the monkey and human. These studies have suggested that object orientation and/or wrist/grip orientation rely on V6A/aSPOC (Battaglini et al., 2002; Monaco et al., 2011) and the adjacent cIPS (Rice et al., 2007; Shikata et al., 2001; Valyear et al., 2006) perhaps along with aIPS (Breveglieri et al., 2022). One fMRI study that had participants grasp a dial much like ours found that “bistable” object orientations (which equally afforded two possible grip postures) evoked greater activation in cIPS than stable orientations (which afforded only one grip posture) (Wood et al., 2017). Moreover, this study found that two neuropsychological patients with unilateral damage to cIPS showed deficits in grasping bistable orientations with the contralateral hand. We find that cIPS codes a combination of start orientation, end
orientation, and task during action execution. Moreover, cIPS activation is more consistent throughout a trial in a complex turning task, where object orientation must be changed through a grip rotation, than a simpler grasp-only task, corroborating the proposed role of cIPS in processing orientation for action. To our surprise, we did not observe any orientation coding in aSPOC or pSPOC.

Akin to cIPS, early visual cortex has been increasingly implicated in motor planning and execution (Gallivan et al., 2014, 2019; Gutteling et al., 2015; Monaco et al., 2020). We show that the visual cortex (V1) codes the start orientation of the object throughout the trial with increases in the task representation toward the end of planning and during execution. This suggests that V1 activation patterns are not solely driven in a bottom-up fashion by current visual cues (Kamitani & Tong, 2005) but come to incorporate the complexity of the task and visual information as the action unfolds.

Although studies of EBA have focused on its role in body perception (Kontaris et al., 2009), EBA has also been implicated in processing body posture during planning and execution of goal-directed actions (Gallivan et al., 2016; Zimmermann et al., 2018). We found some evidence for this notion: EBA has strong representations of task and orientations during execution but only limited ones during planning. As such, EBA may be mainly driven by the perceptual consequences of the action.

**Methodological considerations**

Our “grasping” actions were atypical in that they required gripping the dial between the knuckles of the index and middle fingers, in a way that was similar between grasping, turning to the left, and turning to the right. This choice of posture was deliberate, to ensure that the representations of the upcoming task were not confounded with differences in initial grip posture. Nevertheless, this strategy differs considerably from the approach that participants would likely have undertaken without such instruction. That is, based on the “end-state comfort effect” (Rosenbaum et al., 1990), given free choice, participants likely would have used the right hand to grasp the object between the thumb and index-finger knuckle, and would have adjusted the grip such that the thumb was lowest when planning to turn left, aligned with the index-finger knuckle when grasping without turning, and highest when planning to turn right. Indeed, had they done so, we would expect even stronger multivariate differences between conditions. Nevertheless, we found multivariate differences even in the case when the low-level grasp posture remained constant (Ariani et al., 2015; Ramon et al., 2015).

**Conclusion**
We investigated the neural underpinnings of how hand-object interactions unfold on a moment-to-moment basis. Importantly, we found that the neural underpinnings of grasping and turning actions could best be differentiated in occipital cortex (particularly cIPS and V1) and SPL. These results corroborate the importance of cIPS in processing object and grip orientation and suggest that SPL may play a larger role in action sequences than previously realized. Our results also show how tMVPA can be used to understand how motor actions are represented across different phases of the trial.

Bibliography


<table>
<thead>
<tr>
<th>Region of Interest</th>
<th>Number of Functional Voxels</th>
</tr>
</thead>
<tbody>
<tr>
<td>M1</td>
<td>98 ± 7</td>
</tr>
<tr>
<td>S1</td>
<td>123 ± 11</td>
</tr>
<tr>
<td>PMv</td>
<td>107 ± 8</td>
</tr>
<tr>
<td>PMd</td>
<td>217 ± 9</td>
</tr>
<tr>
<td>aSPOC</td>
<td>178 ± 14</td>
</tr>
<tr>
<td>pSPOC</td>
<td>138 ± 7</td>
</tr>
<tr>
<td>aIPS</td>
<td>268 ± 17</td>
</tr>
<tr>
<td>cIPS</td>
<td>172 ± 10</td>
</tr>
<tr>
<td>Pre-SMA</td>
<td>108 ± 7</td>
</tr>
<tr>
<td>SMA</td>
<td>132 ± 10</td>
</tr>
<tr>
<td>DLPFC</td>
<td>129 ± 9</td>
</tr>
<tr>
<td>V1d</td>
<td>152 ± 13</td>
</tr>
<tr>
<td>Area-5L</td>
<td>167 ± 10</td>
</tr>
<tr>
<td>Area-7A</td>
<td>194 ± 12</td>
</tr>
<tr>
<td>Area-7P</td>
<td>150 ± 9</td>
</tr>
<tr>
<td>Area-7PC</td>
<td>198 ± 13</td>
</tr>
<tr>
<td>EBA</td>
<td>1163 ± 23</td>
</tr>
</tbody>
</table>

Table 1. The average number of functional (1-mm iso-)voxels per Region of Interest. Values shown are mean ± SEM.
Figure legends

**Figure 1.** Experimental set-up and design. **1A.** Picture of participant set-up in the fMRI scanner shown from side view. **1B.** Close up of participant using the instructed grip to grasp the experimental stimulus, i.e., rotatable dial with a handle. **1C.** Experimental conditions in a two (start orientation: left or right) by two (action: grasp or turn) design. **2D.** Timing of each event-related trial. Trials began with the stimulus being illuminated (TR 0-3) indicating the start of the preparation phase. At TR 3 participants received task instructions indicating the start of the planning phase (TR 3-9). At TR 9 participants received the “go” cue indicating onset of action execution (TR 9-10). After 2 s (1 TR) the light was turned off and the intertrial interval was initiated (TR 10-16).

**Figure 2.** Brain areas selected for multivariate analysis based on a univariate contrast. Cortical areas that exhibited larger responses during the experimental trials [(presentation + plan + execute) > baseline] are shown in yellow to red activation. Results calculated across all participants (RFX GLM) are displayed on the dynamic group average surface across participants. The selected ROIs were selected in BrainVoyager software, which uses a hexagonal selection area, and then transformed to individual volumetric MNI space. Each ROI is linked to the group average of corresponding % signal change of BOLD activity (Y-axis) over time (X-axis; each X-ticks represents one TR of 2 s, based on a deconvolution analysis) for each of the four conditions. Vertical dashed lines on the graphs indicate start of the planning phase (TR 3) and the execution phase (TR 9). Sulcal landmarks are denoted by white lines (stylized according to the corresponding legend). ROI acronyms are spelled out in the methods. tMVPA results are indicated using different types of circles (e.g., dashed vs. solid). RSA results are shown by color-coding the tMVPA circles using a Venn diagram. Note that the RSA and tMVPA results are discussed in the results section as well as in Figure 5 and Figure 6.

**Figure 3.** Overview of the steps to perform temporal multivoxel pattern analysis (see methods for full explanation). The figure represents the steps taken for each participant separately. Within each condition single trial representational distance matrices (RDMs) are calculated for each timepoint of each trial pairing (**3A** shows example for grasping trials and **3B** for turning trials). **3C:** Grey cubes represent voxels of the same ROI during different timepoints for two sample trials to exemplify how cells for each RDM are calculated. After calculating single trial RDMs, a condition average is calculated (middle column; **3D** and **3E** for grasping and turning respectively) which are then subtracted from each other (last column; **3F:** red and blue matrix). Not shown on picture (for results see Figure 4 and 5): the subtraction matrices are then used for bootstrapping to determine whether the group average differs significantly from zero and to investigate whether one condition is more similar than the other for each given timepoint.

**Figure 4.** Models used for representational similarity analysis (RSA). The first row shows the models for start orientation (**4A**), end orientation (**4B**) and turn versus grasp (**4C**). The second row shows example data for one single timepoint for two regions of interest (**4D**: V1d); **4E**: Area-7A) of one participant.

**Figure 5. Left:** Hierarchical clustering of all the ROIs based on the averaged representational dissimilarity matrices (RDMs) of the Grasp and Turn condition. We refer the reader to the legend of Figure 3 for the full explanation of the heatmaps. Note that T-G represents ‘Turn minus Grasp’ and represents the final tMVPA result showing statistically significant similarities (small red boxes = Turn > Grasp; small green boxes = Grasp > Turn).
Figure 6. Selection of ROIs that showed prominent differences for the tMVPA analysis. Some data are reproduced from subsets of Figures 2 (Column 1) and 5 (Columns 2-4, at higher resolution) to facilitate inspection and comparisons across analysis methods for key regions. Column 1 shows the group average deconvolution time courses for the four conditions. The X-axis represents time (in 2 s TRs) and the Y-axis represents % signal change in the BOLD signal. Vertical dashed lines represent onset of the planning phase (TR 3) (after the initial presentation phase) and the execution phase (TR 9) following the planning phase. Columns 2 and 3 show the group average condition representational dissimilarity matrices for the grasp (column 2) and turn condition (column 3). Each cell/square of the matrix represents the dissimilarity between trials of the same condition between two timepoints. Dashed horizontal and vertical lines represent the same timepoints as those in Column 1, i.e., the first dashed horizontal/vertical line represents the start of the planning phase (TR 3) and the second dashed vertical/horizontal line represents the start of the execution phase (TR 9). Column 4. tMVPA results. Dashed lines have the same representation as explained for columns 1-3. Red squares indicate dissimilarity metrics between timepoints where turning trials are significantly more similar (less dissimilar) than grasping trials. Green squares represent the opposite (grasping more similar than turning). Column 5. Representational similarity analysis using three models, as explained in the color legend, for each timepoint of the experimental trials. Solid lines represent the correlation between the data and the model over time. The shaded grey area represents the noise ceiling with the lower and upper edges representing the lower and upper bounds. Dashed vertical lines have the same representation as explained for the previous columns. Colored asterisks indicate timepoints with statistically significant correlations for the respective model.
This document contains a diagram showing a trial plan and timing for a task. The task involves pressing buttons labeled 'Grasp' and 'Turn' with varying intervals. The diagram illustrates how the task is divided into different stages such as Start left, Start right, Grasp, Grasp and turn, Light on, "Grasp" or "Turn", and "Go". The timing is marked with intervals of 2 TRs (Time-Resolved).
Central sulcus
No turn-selectivity
tMVPA
RS during plan and/or execute
Post-central sulcus
Precentral sulcus
Intraparietal sulcus
Superior and inferior frontal sulci
Transverse occipital sulcus
Parietal-occipital sulcus
Cingulate sulcus
DLPFC
0
1.6
0.8
0 3 9
PMd
M1
S1
5L
RSA during plan and/or execute
7A
cIPS
PMv
aIPS
7PC
7P
V1d
pSPOC
aSPOC
cSPOC
pre-SMA
SMA
EBA
Imported from Rosenke et al., Cereb Cortex 2021
0
1.6
0.8
0 3 9
Grasp left
Grasp right
Turn left to right
Turn right to left
8.00
8.00
t(15)
q(FDR) < 0.05
2.41

TMVPA
- No turn-selectivity
- Turn-selective in plan or execute
- Turn-selective in plan and execute
Timepoint 2

Timepoint 10

Grasp left: trial m

Grasp right: trial n

Average across all trial pairs for Grasp Left and Grasp Right

Average across all trial pairs for Turn Left to Right and Turn Right to Left

Single subject Grasp RDM (1-r)

Single subject Turn RDM (1-r)

Subtract condition RDMs

Grasp trials are more similar than turn trials

Turn trials are more similar than grasp trials

Turn trials are more similar than grasp trials

Highly similar

Highly dissimilar

Highly similar

Highly dissimilar

Highly similar

Highly dissimilar

Highly similar

Highly dissimilar

Highly similar

Highly dissimilar
Grasp

Start Orientation

Turn

End Orientation

Turn Versus Grasp

Single subject V1d
Time point 4

Single subject Area-7A
Time point 12