



OPEN Temporal hierarchy of observed goal-directed actions

Shahar Aberbach-Goodman & Roy Mukamel

During social interactions, we continuously integrate current and previous information over varying timescales to infer other people's action intentions. Motor cognition theories argue for a hierarchical organization of goal-directed actions based on temporal scales. Accordingly, transient motor primitives are represented at lower levels of the hierarchy, a combination of primitives building motor sequences at subordinate levels, and more stable overarching action goals at superordinate levels. A neural topography of hierarchal timescales for information accumulation was previously shown in the visual and auditory domains. However, whether such a temporal hierarchy can also account for observed goal-directed action representations in motor pathways remains to be determined. Thus, the current study examined the neural architecture underlying the processing of observed goal-directed actions using inter-subject correlation (ISC) of fMRI activity. Observers ($n = 24$) viewed sequential hand movements presented in their intact order or piecewise scrambled at three timescales pertaining to goal-directed action evolution (Primitives: ± 1.5 s, Sub-Goals: ± 4 s, and High-Goals: ± 10 s). The results revealed differential intrinsic temporal capacities for integrating goal-directed action information across brain areas engaged in action observation. Longer timescales ($> \pm 10$ s) were found in the posterior parietal and dorsal premotor compared to the ventral premotor (± 4 s) and anterior parietal (± 1.5 s) cortex. Moreover, our results revealed a hemispheric bias with more extended timescales in the right MT+, primary somatosensory, and early visual cortices compared to their homotopic regions in the left hemisphere. Our findings corroborate a hierarchical neural mapping of observed actions based on temporal scales of goals and provide further support for a ubiquitous time-dependent neural organization of information processing across multiple modalities.

Humans perform voluntary actions to accomplish goals spanning varying complexities and temporal scales. Although complex, the goals of observed action sequences are effortlessly inferred by typically developing individuals during social interactions. Despite three decades of studies of the brain areas involved in the processing of observed actions, known as the action observation network (AON)^{1,2}, there is currently no consensus in the literature regarding the organizing principle according to which observed sequences of goal-directed actions are parsed and mapped in the brain. Behaviorally, it was shown that people utilize the information of observed kinematics to infer other's intentions³. At the neural level, several brain regions within the AON have been implicated to play a role in predicting action goals from the observed kinematics, including the superior parietal lobule (SPL), inferior frontal gyrus (IFG), middle frontal gyrus (MFG), and the inferior parietal lobule (IPL), with the latter contributing the most to goal prediction⁴. Moreover, previous studies show that the feature of the realized goal or outcome of observed actions is also encoded in fronto-parietal regions, independent of the action's kinematic features⁵⁻⁷.

The neuroimaging findings of goal encoding are in line with behavioral findings from children, showing enhanced learning from imitation of new action sequences when provided with knowledge about the overarching goals that bind the actions together⁸. However, considering that observed actions are encoded in neural activity both at the kinematic as well as at the conceptual goal level⁹, it is yet unclear how the intention-specifying information conveyed by action kinematics is accumulated and integrated to encode the overarching action goals.

Previous neuroimaging studies have argued for a hierarchical neural organization of observed actions. Findings from a series of repetition suppression fMRI studies^{7,10,11} showed sensitivity for observed low-level kinematics of hand-object interactions in the inferior occipital cortex and the SPL, while the premotor cortex and the intraparietal sulcus (IPS) were shown to be sensitive to objects and action outcomes (e.g. open/closed box), regardless of the performed movement (e.g., pushing/pulling). The levels of action representation in the ventral premotor cortex (PMv) and IPS were later dissociated in additional fMRI studies of action observation, showing low-level, concrete, action representation in the PMv while action encoding was generalized across different

Sagol School of Neuroscience and School of Psychological Sciences, Tel Aviv University, 6997801 Tel-Aviv, Israel.
 email: rmukamel@tau.ac.il

perceptual-object and kinematic features in the IPS and the lateral occipital temporal cortex (LOTC)^{12,13}. These findings indicate hierarchical levels of the neural representation of goal-directed actions within action observation pathways—a concrete representation of the movement’s perceptual-sensory and motor features in motor/premotor areas, and an action goal representation that is generalized across these physical attributes in parietal cortices. Nevertheless, there is no consensus regarding the specific relation between the levels of the hierarchy (e.g., the kinematics, the manipulated object, the visual outcome, etc.) or the level at which each aspect is represented relative to the others (i.e., objects at the same or higher level than kinematics). Moreover, the actions examined in these studies consisted of a discrete outcome. Thus, the neural substrates for processing natural continuous action sequences, comprised of multiple sub-actions directed to an overarching goal, remain an open question.

Other than arguments for an action hierarchy based on the abstraction of sensory-motor features, Uithol and colleagues argued for a hierarchy of temporal stability for the organization of goal-directed action sequences that is governed by the actions’ dynamics over *time*¹⁴. Specifically, that high levels of the hierarchy encode more stable (goal-related) representations, whereas lower levels represent more transient (motor primitives) kinematics. Accordingly, different brain regions are postulated to have different intrinsic time scales on which they can represent sequential contingencies.

Previous studies, using Inter-Subject Correlation (ISC)^{15,16} revealed a topographical mapping of processing time scales in brain areas of the auditory and visual systems^{17–19}. Specifically, the authors defined a temporal receptive window (TRW) for a given region as the length of time during which sensory information is integrated and may affect the response in that region. In their findings, TRW sizes were shown to increase from primary to high-level perceptual and cognitive areas. Accordingly, the shortest TRWs (< ~4 s) for visual or auditory stimuli were found in the corresponding early sensory regions (V1/MT+, and A1), while long TRWs (> ~30 s) were found in the frontal eye field (FEF) and temporal-parietal junction (TPJ) for visual and auditory stimuli, respectively. These studies support a neural representation of hierarchical information accumulation according to temporal scales. Interestingly, areas engaged in action observation were shown to be sensitive to the predictability of action information over time^{20,21}. Therefore, a potential principle linking the different hierarchical levels of goal-directed actions may be a temporal processing capacity. However, such a temporal organization of hierarchical observed (and executed) goal-directed actions is currently mainly supported by computational models and action recognition of human every-day activities^{22–24}. Thus, a precise neural mapping of levels in the temporal hierarchy, encoding goal-directed actions, in sensory-motor pathways is still missing.

To this end, in the current study we adapted the ISC approach to examine whether brain regions previously found to encode action goal information have different intrinsic time scales for processing observed actions. Specifically, we manipulated the sequence coherency of observed goal-directed hand actions at three timescales: basic motor primitives, Sub-Goals, and High-Goals. Building on previous work indicating that areas involved in the processing of observed actions are responsive to different action elements, we hypothesized to find a graded neural topography of intrinsic processing time scales compatible with the representation of action goals across the areas sensitive to action observation. Action observation engages areas of the motor, somatosensory, and visual systems. Previous studies using similar methods suggested predictive coding as a potential mechanistic delineation of the feed-forward and feedback connections between different action observation areas^{20,21}. These studies show that the PMv, IPS, and inferior parietal lobe (IPL) are sensitive to action predictability and therefore to the temporal coherency of action information. Nevertheless, it is yet unclear whether these frontoparietal regions have different temporal processing capacities and whether or not they are positioned at similar or different levels of the action hierarchy. We thus have no specific prediction regarding the topographical gradient direction of the TRW sizes. In addition to its sensitivity to action coherency, the parietal lobe is considered in the literature as an integrative hub for sensory-motor and visual information, and was indicated to have a prominent role in high-level action representation²⁵ and coding of intention^{26,27}. We thus hypothesized the regions of the parietal lobe to exhibit a relatively long TRW, allowing high-level goal representation that is more stable over time.

Methods

Participants

Twenty-eight subjects were recruited for the study. Three subjects were excluded from data analysis due to excessive movement in the scanner, and one due to low performance (see experimental procedure below). Accordingly, the data for the study was acquired from a final sample of 24 participants (15 females, mean age 24.8, range 18–34 years). All participants were healthy with no history of neurological or psychiatric disorders, right-handed (self-report), had normal or corrected-to-normal vision, and were naive to the purposes of the study. All methods in the study were carried out in accordance with the approved guidelines of the ethical committee at Tel-Aviv University and the Helsinki Committee of the Sheba Medical Center. All participants provided written informed consent to participate and were compensated for their time.

Stimuli

The experimental stimuli consisted of four videos depicting the hand movements of a right-handed actor preparing a cake. The videos were edited using ADOBE Premiere Pro CC 15.0 and DaVinci Resolve 16. First, we constructed the *intact* video. This movie contained the actor’s movements in a chronological sequential order. Then, the intact movie was subdivided into clips containing basic motor primitives (e.g., grasping; mean duration 1.5 s ± 0.06 s). These clips were randomly reordered to construct the video for the shortest timescale scramble—the *Primitives* condition. For the two other timescale scramble conditions, several clips consisting of primitives such as *reaching*, *grasping*, *transporting the cup*, etc. were joined in their original chronological order to construct meaningful action goals at the *Sub-Goal* level (e.g. *pouring the egg*; mean duration 4.1 s ± 0.2 s), and *High-Goal*

level (e.g. 'adding eggs to cake batter'; mean duration $10 \text{ s} \pm 0.56$) consisting of several sub-goals such as 'pouring the egg', 'blending with batter', etc.). See Fig. 1 for illustration of study design and Table S1 for a complete list of primitives and action goals. Each of the four videos contained all primitive clips (with a different ordering depending on the condition's timescale) and had a total duration of 309 s. Representative 30 s snippets from each of the four experimental video conditions are provided in the supplementary materials S1–S4.

Experimental procedure

Participants were presented with the four goal-directed action videos during four separate runs within a single fMRI session and were instructed to attend the videos. Each video was presented only once during a single scanning run. Video order was randomized and counterbalanced across participants. To assess participants' level of attention, following each video presentation they were requested to verbally answer a different yes/no question (e.g., 'Did the actor drop something?', 'Did the actor rotate the cake?', 'Did the actor use scissors?', 'Did the actor clean the surface?'). The data from one subject who failed to answer the question correctly on more than 2 videos was discarded from further analysis. The anatomical scan (MP2RAGE) was performed following the first two experimental runs. In each scan, a fixation mark was presented at the beginning and end of the video presentation. In order to prevent any predictability differences between the conditions with regard to the main goal of the actor in the videos, participants were informed at the beginning of the session that they will observe videos of a person preparing a cake. Videos were presented using Psychtoolbox-3 (Brainard 1997, www.psychtoolbox.org) on MATLAB 2016b (The MathWorks, Inc., Natick, MA, USA). Stimuli were presented on a 32" monitor and viewed by the participants through a mirror placed on the MRI head coil.

Image acquisition

Functional imaging was acquired on a Siemens Magnetom Prisma 3T Scanner (Siemens Healthcare) with a 64-channel head coil at the Tel-Aviv University Strauss Center for Computational Neuroimaging. An interleaved multiband gradient-echo, echo-planar pulse sequence was used in all functional scans. Whole-brain coverage was provided by acquiring 66 slices for each volume (slice thickness 2 mm; voxel size 2 mm isotropic; time to repetition (TR) = 1000 ms; time to echo (TE) = 30 ms; flip angle = 82° ; field of view = 192 mm; acceleration factor = 2). For anatomical reference, a whole-brain high-resolution T1-weighted scan (slice thickness 1 mm; voxel size 1 mm isotropic; TR = 2530 ms; TE = 2.99 ms; flip angle = 7° ; field of view = 224 mm) was acquired for each participant.

Video conditions

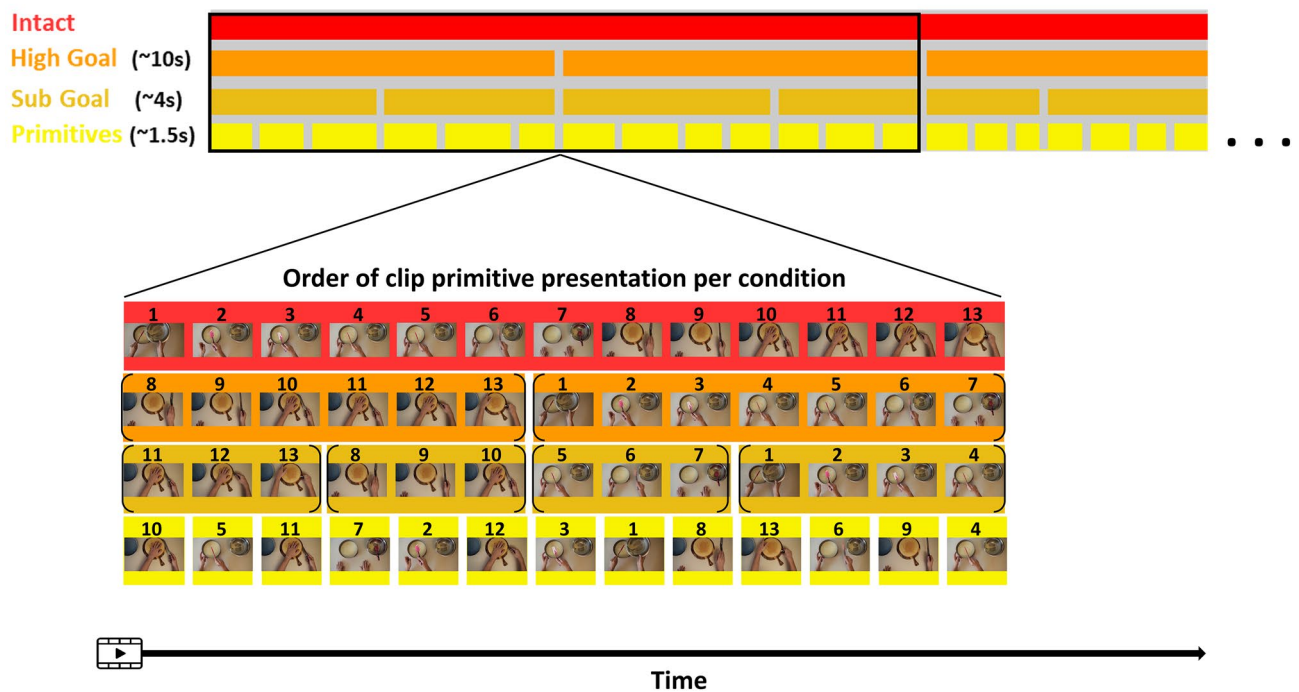


Figure 1. Task design. Participants observed four movies of goal-directed actions, performed by an actor preparing a cake. One video consisted of the actions in their intact order and the three other videos consisted of the actions scrambled at High-Goal (coherent information within each $\pm 10 \text{ s}$ time bins of the movie), Sub-Goal (coherent information within each $\pm 4 \text{ s}$ time bins of the movie), and Primitives (coherent information within each $\pm 1.5 \text{ s}$ portion of the movie) time scales. In the illustration, each image represents a primitive movie clip with its number representing the clip temporal order. All four movies had a duration of 5 min and consisted of the same 200 primitive clips (concatenated in different ordering depending on the scramble condition).

Image preprocessing

fMRI data preprocessing was conducted using the FMRIB's Software Library's (FSL v5.0.7) fMRI Expert Analysis Tool (FEAT v6.00)²⁸. The data from each experimental run underwent the following preprocessing procedures: brain extraction, slice-time correction, high-pass filtering at 100 s (0.01 Hz), motion-correction to the middle time-point of each run, smoothing with a 5-mm FWHM kernel, and correction for autocorrelation using pre-whitening (as implemented in FSL). Participants with an absolute displacement value exceeding 2 mm were excluded from analysis. All images were registered to the high-resolution anatomical data using boundary-based reconstruction and normalized to the Montreal Neurological Institute (MNI) template using nonlinear registration.

Data analysis

Inter-subject correlation analysis

Inter-Subject Correlation (ISC) is a measure of reliability of the stimulus-locked neural response, in which the shared variance across subjects is used as a data-driven estimate of the evoked response. Accordingly, the time course of a voxel, or brain region, in a target subject is modeled with the time course of the corresponding location in a different source subject¹⁵. Brain activity in a single voxel or region of interest (ROI) can include three main components: a stimulus-evoked component that is common across subjects, a subject-specific stimulus-evoked component, and a stimulus-unrelated or noise component. By modeling one subject with another subject's time course, the subject-specific and unrelated noise components are effectively filtered out—thus isolating the stimulus evoked responses that are shared across subjects¹⁶. Under the ISC approach, the level of engagement of a certain brain region during a task is determined according to the level of neural synchronization within that region across multiple subjects. The main advantage is that it allows the use of continuous dynamic naturalistic stimuli to discover the functional characteristics of different brain regions without requiring a pre-defined hemodynamic response model. The underlying logic of the ISC method is that, if a brain region is responsive to a certain stimulus and has the capacity to integrate over the continuous stimulus information, then randomizing the coherent order of the stimulus and destroying the high-level information reduces the processing efficacy in that region, and therefore decreases the degree of similarity (synchronization) across participants²⁰.

Reliable voxels

To identify voxels with significant shared responses to the presentation of observed actions, we followed the leave-one-out approach for ISC analysis^{16,19}. In this procedure, for each of the experimental videos, the time course of a given voxel in one subject was correlated with the average time course of the same voxel across all remaining $n-1$ subjects (excluding that subject). The time course included all fMRI blood oxygenation level dependent (BOLD) volumes starting five seconds following movie onset until five seconds following movie offset (to account for the inherent delay of the hemodynamic response function)^{17,29}. The analysis produced a whole-brain map of Pearson's linear correlation coefficients for each subject (in MNI space). Finally, a whole-brain group level ISC map was computed by averaging all individual maps, resulting in an average r value for each voxel. Only voxels with positive ISC values were considered in this analysis. The statistical significance of the group-level ISC values was computed using a phase randomization bootstrapping procedure¹⁹. Accordingly, we applied the Fourier transform to the empirical time series of each voxel (in each subject) and randomized the phase of the Fourier components. Then, we performed the inverse Fourier transformation, thus preserving the power spectrum of the signal but disrupting the temporal alignment¹⁶. We then computed a group ISC map in a similar manner to the way it was computed for the original data. This was repeated 1000 times, generating a null distribution of the correlation values. For each experimental video, the original ISC values were compared with their null distributions to compute the ISC p values. Finally, the p values across all experimental videos were corrected for multiple comparisons using the Benjamini–Hochberg false-discovery rate procedure (FDR)³⁰, with $q = 0.01$. For the main comparison of ISC values across conditions, we used a mask that included all voxels that had significant ISC values in at least one of the experimental conditions.

ROI-based ISC and Paired-sample comparisons

To examine the temporal capacity of different cortical areas to accumulate information from observed, goal-directed actions, we parceled the brain into ROIs based on the Human Brainnetome Atlas³¹. Only ROIs that contained at least 30 voxels that reliably responded to at least one of the experimental videos (see voxel-based ISC analysis described above), were taken for further analysis. For each ROI included in this analysis, we computed the average time course across all significant voxels of each subject, separately for each of the four video presentations (i.e., Intact, High-Goal, Sub-Goal and Primitives). The ROI-based BOLD time courses were taken to compute the ISC in each condition using the same leave-one out approach described above for the analysis at the single voxel level.

For a given ROI, ISC should be highest for stimuli presented at a scrambling level that corresponds with the region's TRW. Presentation of scrambled stimuli at shorter time-scales than the optimal TRW will disrupt processing, and yield lower ISC values¹⁶. Presentation of scrambled stimuli at longer time-scales is not expected to disrupt processing, therefore ISC levels for longer time-scales should be similar (or at least not higher) than those of the optimal time-scale. The intact movie contains coherent information within all temporal time windows (i.e., short, intermediate and long time-windows). Therefore, the ISC value of the intact condition should be either maximal (in case the true TRW of the region is the longest time-scale examined), or not higher than the maximal ISC value obtained for a scramble condition (in case the true TRW of the region corresponds with a shorter time-scale). In order to identify the relative TRW type of each ROI at the group level, the ROI's ISC in the intact condition was compared to the ISC in each of the three scramble conditions using a one-tailed paired

t-tests (ISC_Intact > ISC_scramble). Since correlation values do not follow a normal distribution, we performed a Fisher-z transformation of the ISC values prior to conducting the paired-sample t-tests. The TRW of each ROI was then determined according to the shortest time scale of scrambled videos, for which the ISC was not significantly smaller than the ISC in the intact (unscrambled) condition. For example, if a certain region does not show lower ISC levels in any of the scrambled conditions compared to the Intact, this region would be considered to have the shortest TRW—corresponding to the representation of the level of motor Primitives alone ($\leq \sim 1.5$ s). Similarly, a region exhibiting a reduction in ISC compared to the intact video, only for the Primitives condition but not for the higher time scales of Sub-Goal and High-Goal scrambles, would be considered to have a higher TRW—representing the observed information at the Sub-Goal and primitives level ($\leq \sim 4$ s). According to this rationale, a region with a certain TRW will be considered sensitive to the information presented within the time scale of that TRW and all time scales beneath it, i.e., a region with a long TRW representing high-goals, also represents the lower-level sub-goal and primitives, that unfold within shorter time windows. See Fig. 2 for a schematic depiction of different ISC patterns and their corresponding TRW levels.

Results

Voxel-based ISC analysis (whole brain)

In order to examine the sensitivity of different regions to the temporal scales of observed goal-directed actions, we first identified the voxels that reliably responded to observed actions during *at least one* of the videos (see “Methods”). The voxel based ISC analysis of the BOLD response revealed voxels that are significantly correlated across participants ($p < 0.01$, FDR corrected) in extensive portions of the brain, including cortical and sub-cortical regions. ISC values were highest in early visual cortices, MT+, lateral superior occipital gyrus, superior parietal lobule and the somatosensory cortex, see Fig. 3.

ROI based ISC analysis

In order to identify the temporal capacity of different brain regions to accumulate information during observation of goal-directed actions, we extended our analysis from the single voxel level to the level of ROI. To this end, in each ROI for each subject, we first averaged the time series of the reliable voxels identified from the voxel based ISC analysis described above. Next, we computed the ISC value in each condition using the average ROI time course. Only ROIs with more than 30 reliable voxels were analyzed (see Table S2 for the complete list of ROIs and their ISC values in the four video conditions). In the current analysis, areas in which neural activity is primarily driven by transient dynamics of the stimuli (i.e., short TRW), would not exhibit a difference in ISC across the scrambled video conditions compared to the Intact video presentation. In contrast, in brain regions where activity depends on information accumulated over several seconds, the ISC should decrease with increasing scrambling levels (see “Methods” for the full schematic explanation).

ISC values across different brain regions and timescales of observed goal-directed actions

Long TRW

The longest TRWs, of more than ± 10 s, were found bilaterally in the Precuneus (A7), in the left inferior frontal junction (IFJ) and left dorsal premotor (PMD) cortex (A6) of the middle frontal gyrus (MFG) and in the left IPS (Hip3). In the right hemisphere, the longest TRWs were also found in the SPL and PFM of the posterior inferior parietal lobule. In these areas, ISC values were significantly lower in all scrambled conditions compared to the intact condition. In addition, the right lateral superior occipital gyrus (lsOccG), and early visual areas in the medio ventral occipital cortex (MVOcC) and medial superior occipital gyrus (msOccG) also exhibited the

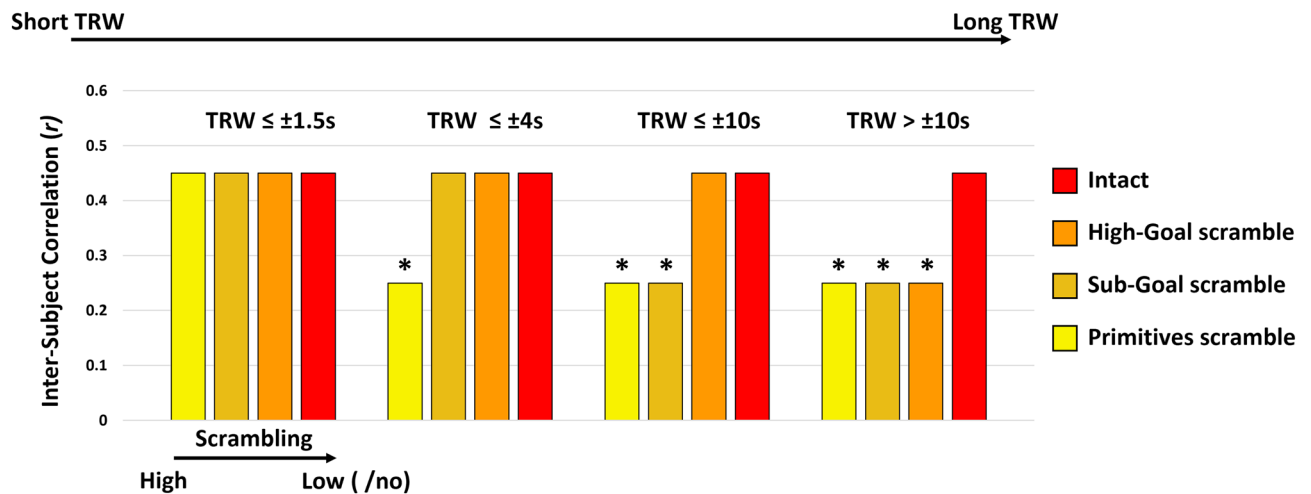


Figure 2. Schematic of four ISC patterns across conditions and their corresponding TRW types. The TRW of an ROI is determined according to the shortest time scale of scrambled videos (i.e., High-Goal/Sub-Goal/Primitives) for which the ISC was not significantly lower compared to the ISC in the intact (unscrambled) condition. *Significant reduction in ISC in scrambled compared to the intact condition.

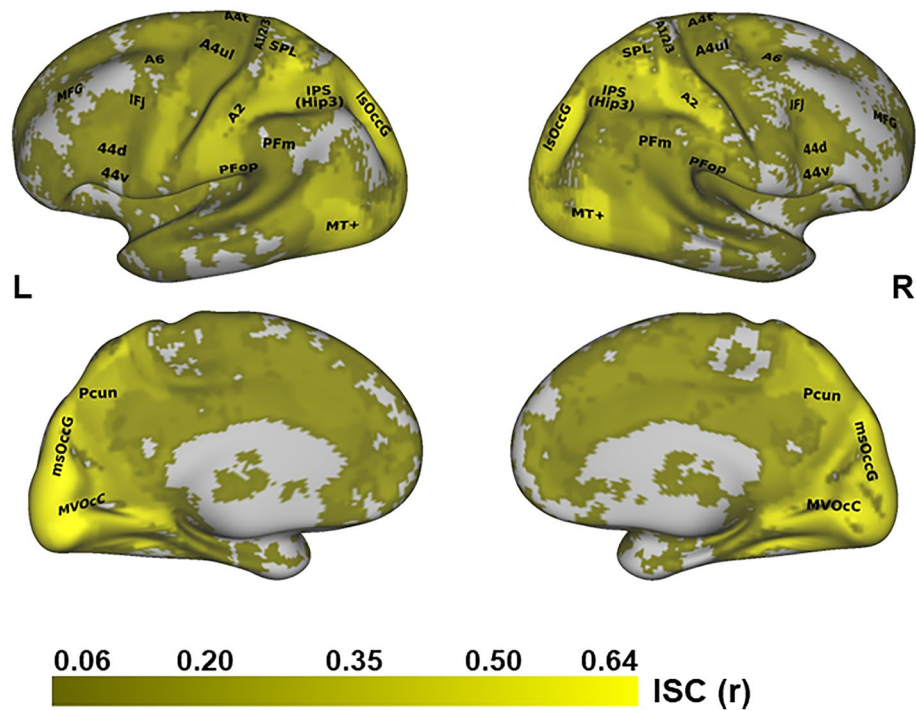


Figure 3. Voxels reliably responsive across subjects during action observation. Lateral (top) and medial (bottom) maps of ISC values in significantly responsive voxels during at least one of the experimental video presentations (tested versus a shuffle distribution of ISC values). *L* left hemisphere, *R* right hemisphere. Regions names on map according to the Brainnetome Atlas nomenclature.

longest TRW, while in the left hemisphere these regions exhibited shorter processing time scales. See Fig. 4 for a visualization of the TRWs across regions. Figure 5 shows ISC values in selected regions, and Table S2 provides the complete list of ISC values across regions and experimental conditions.

Intermediate high-goal TRW

Compared to the intact condition, ISC values were significantly lower in the Primitives and Sub-Goal scramble conditions, but not in the High-Goal scramble condition, in the bilateral middle frontal gyrus (MFG), left SPL and right primary somatosensory cortex (A1/2/3) and MT+ area. Thus, these regions exhibit a relatively long TRW of more than ± 4 s but less than 10 s (see Figs. 4 and 5).

Intermediate sub-goal TRW

The lower TRW of the Sub-Goal processing time scale of more than ± 1.5 s but less than ± 4 s includes the bilateral somatosensory A2, the right PMd cortex (A6), and the left ventral premotor (A44d), primary motor (A4ul) and MT+ areas. Thus, these regions show a decrease in ISC only in the Primitive scramble (± 1.5 s) compared to the intact video presentation. The Sub-Goal processing time scale was additionally found in the left IsOccG and early visual cortices (MVOcc, msOccG) while their homotopic regions in the right hemisphere exhibited longer TRWs, as mentioned above (see Figs. 4 and 5).

Short primitives TRW

The shortest TRW, of the Primitives time scale, was attributed to regions in which ISC values did *not* decrease in any of the scrambled conditions compared to the intact video presentation (see Fig. 2 for the TRW schematic). Bilaterally, these regions include the anterior inferior parietal lobule (PFOP), ventral A44, and inferior occipital gyrus (iOccG). Short TRWs were also found in the right IFJ, dorsal A44 (A44d) and primary motor cortex (A4t). The right IPS (Hip3) also exhibited the lowest TRW time scale while the left IPS exhibited a long processing time scale. The opposite mapping was found for the PFM, such that this region exhibited a short TRW in the left hemisphere and a long TRW in the right hemisphere (see Figs. 4 and 5).

In addition to the aforementioned ROIs, several ROIs with over 30 significant voxels across the experimental conditions showed patterns of ISC levels between conditions that were inconsistent with any of the TRW schemes described in Fig. 2. Thus, these regions could not be assigned with one of the TRWs tested in the current study (see Table S2).

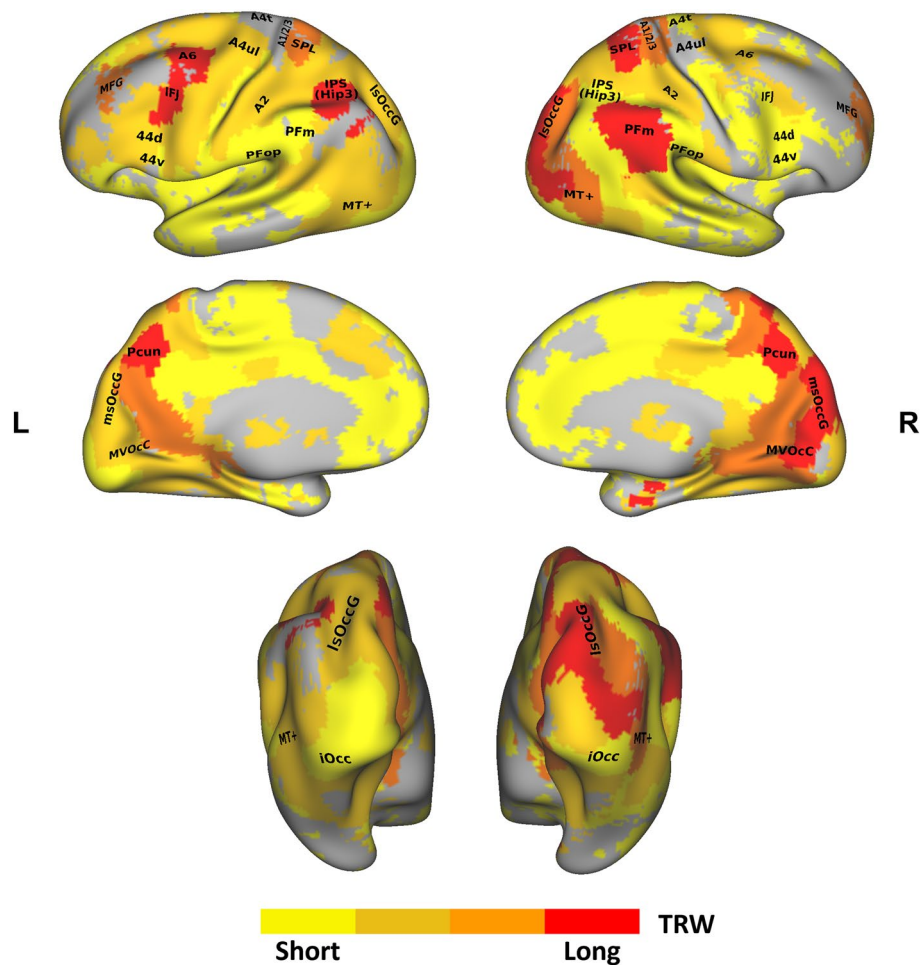


Figure 4. Hierarchical topography of TRWs across ROIs. Each ROI in the maps is colored according to its TRW type, as revealed by the paired sample comparison results, and determined according to the shortest time scale of scrambled videos in which the ISC was not lower compared to the intact (unscrambled) condition. *L* left, *R* right. Regions names on map according to the Brainnetome Atlas nomenclature (see “Methods”).

Discussion

Current theories of action organization postulate hierarchical structuring that is dependent on the temporal dynamics of action goals¹⁴. Yet thus far a neural mapping of such an action hierarchy was lacking. Therefore, in the current study, we examined whether brain areas in sensory-motor pathways involved in the processing of observed actions, integrate the goal-directed action information over graded timescales. As expected, primary motor areas exhibit low processing temporal scales while regions of the intraparietal sulcus and inferior parietal cortex exhibit the most extended timescales. Interestingly, the primary somatosensory cortex and early visual areas showed relatively long timescales as well. Moreover, the long TRWs in sensory regions were found to be biased to the right hemisphere. Our results conform with a hierarchical topographical mapping of action goals according to increasing processing temporal scales and contribute to prominent arguments of predictive coding mechanisms for inferring other people’s goals.

Areas engaged during action observation exhibit hierarchical temporal scales for processing goal-directed action sequences

In the current study, we examined the neural topography of TRWs among areas responsive to observed everyday actions. The areas found responsive encompass regions of the AON as defined by traditional univariate contrast analysis of BOLD responses^{1,32}. In addition to the traditional areas of the AON, the current ISC analysis also revealed voxels in the Precuneus and MFG, and in several early visual cortices. The ISC analysis is indeed complementary to response amplitude methods and can therefore uncover additional phenomena of sensitivity to dynamic stimuli in the brain that are undetectable using univariate measures of response amplitude^{16,17}. The voxel-based analysis in the current study revealed the highest ISC values during observation of goal-directed actions in inferior and superior occipital regions, SPL, and somatosensory cortex. Our voxel-based findings are compatible with the study of Thomas and colleagues, who used ISC analysis and showed voxels that are responsive to observed actions in somatosensory, and visual cortices³⁰. Beyond the overall high voxel-based ISC levels, our

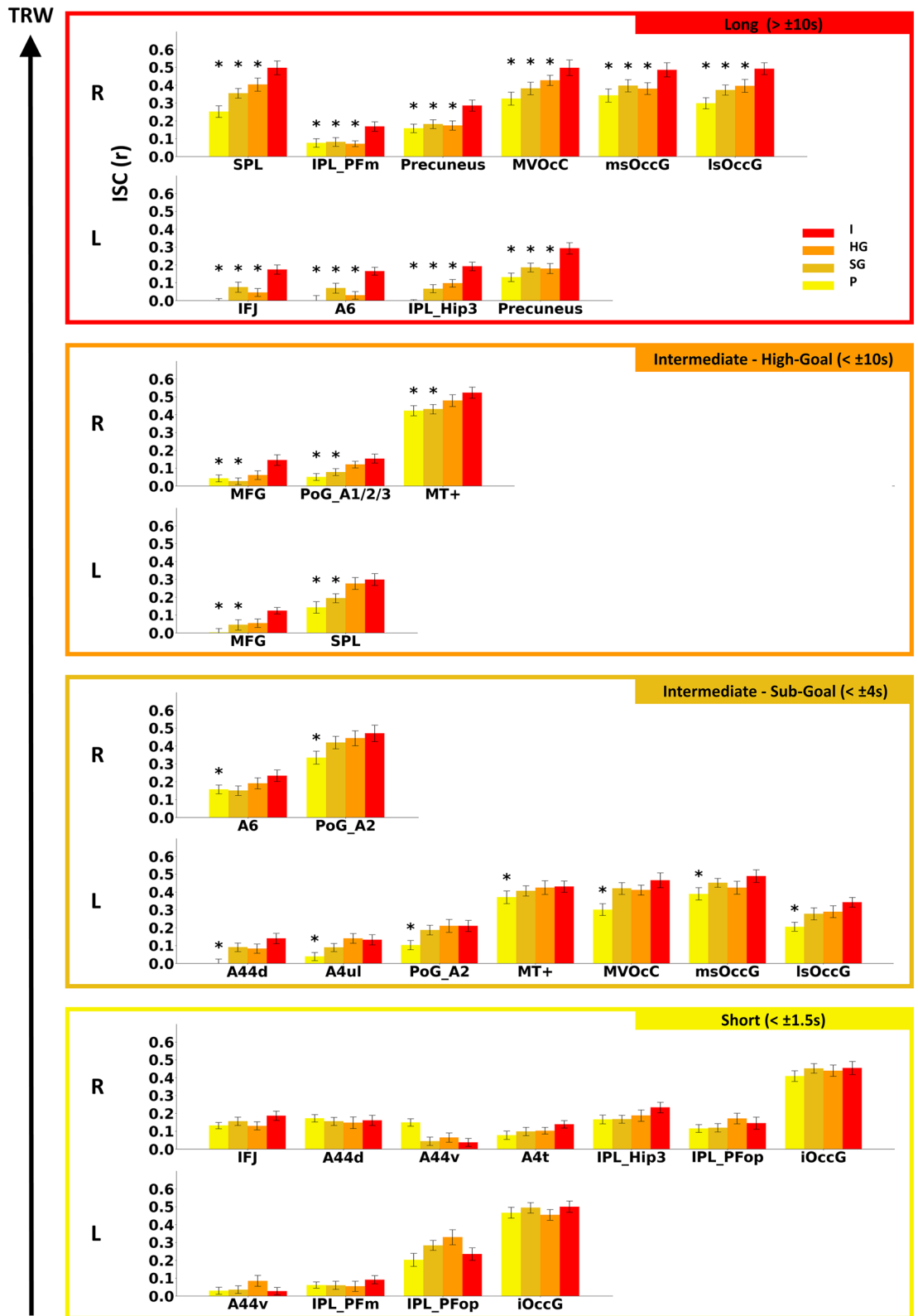


Figure 5. TRW hierarchy according to ISC values in several of the ROIS. ISC values in each movie condition for the right and left hemispheres (upper/lower row in each panel respectively). Regions with the longest to shortest TRWs are presented from top to bottom panels. Note the differences in absolute ISC levels across regions. Importantly, the TRW is determined by the relative pattern of correlation across conditions within the same region. *P* primitives, *SG* sub-goal, *HG* high-goal, *I* intact.

ROI-based ISC analysis revealed hierarchical topographical mapping of goal-directed action sequences across regions in sensory-motor pathways, based on their temporal sensitivity. In both hemispheres, the results show relatively long processing timescales of high-level goal representation in the Precuneus, SPL and MFG, whereas relatively short TRWs, of sub-goal and primitive levels, in the primary motor and ventral premotor cortices.

Our result of long TRW found in the Precuneus is in line with previous studies showing long TRW in this region based on ISC during the processing of visual dynamic stimulus or auditory narratives^{17,19}. Moreover, previous reports indicate strong sensitivity of the Precuneus to the end-goal compared with the means of observed actions³³ and to its contribution to understanding of other people's intentions and 'mind-reading' during first-person perspective taking³⁴. Taken together, the findings are compatible with the temporal hierarchy hypothesis¹⁴, suggesting that intention and goal representation is more stable over time and as such, should require relatively long temporal capacity for information representation.

Although not a core AON region, the MFG was previously reported to encode movement kinematics during goal-directed action observation⁴. While the MFG is frequently associated with cognitive control and action selection³⁵ it was also suggested to play a role in maintaining goal-related representations for top-down modulation of sensory activity during working memory³⁶. Our results of a long TRW in the MFG are in line with this view, indicating the region's capacity to integrate action information over a long time window.

In the left hemisphere, the longest TRWs ($> \pm 10$ s) were additionally found in the IPS and PMd cortex while in the right hemisphere, these areas exhibit short TRWs (± 1.5 s $<$ TRW $<$ ± 4 s, or even TRW $<$ ± 1.5 s). Interestingly, the opposite trend is evident for sensory regions. In the right hemisphere, longest TRWs were also found in the lateral superior and medio ventral occipital regions and in the PFm, together with a relatively long TRW (± 4 s $<$ TRW $<$ ± 10 s) in the primary somatosensory cortex, whereas their homotopic regions exhibit relatively short sub-goal processing time scales (± 1.5 s $<$ TRW $<$ ± 4 s or TRW $<$ ± 1.5 s) in the left hemisphere.

Within and across hemispheres, the current results indicate that different brain regions process observed actions on (at least) four different temporal scales tested in the current study. Compatible with the framework of action organization suggesting hierarchical structuring that is governed by the temporal dynamics of action goals¹⁴, the current results indicate a neural mapping of action organization in frontoparietal and visual regions according to the regions' temporal scales for processing observed actions. Note that several regions did not conform with one of the four TRW categories tested in the current study (see the list of regions in the 'Unspecified' TRW section of Table S2). Thus, we cannot determine their position within the temporal hierarchy for processing of goal-directed action sequences.

Shorter time scales of sequence information processing in the left ventral premotor cortex than dorsal premotor and IPS

Compatible with previous findings, we show that the neural response in parietal and premotor regions is modulated by the temporal coherency of observed actions. Moreover, our results reveal that the different frontoparietal regions exhibit differential levels of temporal sensitivity. Thus, the findings delineate the varying capacities of regions to integrate action information. While the IPL, IPS and ventral premotor (BA44) were previously shown to be sensitive to the temporal structure of observed actions^{20,21}, here, we further show that these regions exhibit different temporal capacities to process action goals. Specifically, the left ventral premotor (BA44d) exhibit a Sub-Goal timescale (± 1.5 s $<$ TRW $<$ ± 4 s) of action processing, mapped to an intermediate level in the action hierarchy, while the right IPL, the left IPS and PMd have the capacity to integrate information over longer time scales, thus mapped to the top of the action hierarchy (TRW $>$ ± 10 s). The difference we find between the ventral and dorsal regions of the premotor cortex is in line with an anatomical and functional distinction of the two regions, suggesting that they are involved in different aspects of motor behavior in primates. According to this account, the PMv process direct sensorimotor information while the PMd retrieves and integrates, indirect, multiple sets of motor information from sensory signals³⁷. A similar segregation between PMv and PMd was also found in an fMRI study using repetition suppression for observed manual object-manipulation in humans³⁸. The study showed stronger suppression for hand kinematics in PMv while responses in the PMd were more sensitive to the desired end state of the observed action. The longer TRW we find in the PMd is line with these accounts of higher-level integration of actions in the PMd compared to PMv.

In addition to TRW difference between premotor regions, different TRWs were also found between PMv and parietal regions. Specifically, longer TRWs were found in the left IPS, and right IPL (PFm) compared to the left PMv. While previous results from action execution studies dissociate the level of representation in PMv and parietal regions, pointing to low-level sensory motor representation in the former and a more generalized action goal representation in the latter^{39,40}, action observation studies show a less conclusive dissociation. Several action observation studies place the two regions at the same level, showing their involvement in a generalized goal representation independent of low-level sensory motor features^{5-7,11}, while others show a different level of representation with the PMv being associated with effector and kinematic dependent action representation, and IPL with more generalized representation^{12,13,27,41}. The current results shed new light on these previous findings demonstrating differences in the temporal capacity to integrate observed action information such that PMv process information on a relatively short TRW, compatible with the representation of motor primitives, while the posterior IPL process information on a longer time scale representing the high-level goal information during observation. In light of these reports, our results support the notion that relative to low-level kinematics, observed overarching goals are more stable in time and are thus encoded in areas with a longer temporal capacity to accumulate and integrate action information. In other words, different brain areas have different intrinsic processing characteristics which may determine the neural encoding of observed actions. Accordingly, areas with longer temporal processing capacities constitute the higher levels of the action hierarchy in which low level

information is integrated to encode overarching goals, whereas areas with shorter timescales constitute the lower levels and encode only the transient, motor act kinematics¹⁴.

In the current study, we examined action observation. As similarities were previously found between neural activity during action observation, execution and planning^{1,42}, our results regarding temporal timescales might generalize from observation to action planning and execution. Indeed, the long TRWs we find in the IPS and PFm are compatible with patient studies demonstrating a correlation between lesions in these areas and deficits during the planning and execution of goal-directed actions, known as *Ideational apraxia*^{43–45}. A reduced temporal processing capacity due to such lesions might interfere with the stability of goal representation and be manifested with action planning deficits. It should be noted that in the current study, the PFm and IPS show long temporal processing capacities while other regions of the IPL, such as the more anterior PFop exhibit shorter temporal scales. The parietal lobe is found to be engaged in varying motor as well as other cognitive tasks⁴⁶ and its different sub-regions were shown to play different functional roles both in the current study and in previous findings²⁵. Our results, showing a relatively short TRW in the PFop while a longer TRW is found in the PFm are in line with a connectivity-based segregation between the anterior PFop and posterior PFm areas. In a recent thorough investigation of the functional attributes of the parietal cortex using effective connectivity and diffusion tractography, it was shown that compared with the anterior parietal, posterior parietal regions have stronger effective connectivity with visual cortical areas and the hippocampal system, combining visuomotor, visual object, and reward input, while the anterior regions have greater effective connectivity with somatosensory and premotor cortical areas⁴⁷. This is also in line with a recent high-resolution 7T fMRI study showing increased functional connectivity between deep layers of anterior IPL (area PFt) and PMv (BA44) during the observation of intact vs. scrambled action sequences²¹.

Both primary and high-level areas exhibit long TRWs

Previous studies using ISC analysis of fMRI data parametrically manipulated the temporal coherency of visual¹⁷, and auditory¹⁹ input. Their findings revealed that both the visual and auditory systems exhibit hierarchically increasing TRWs that are topographically mapped from primary to high-level areas of the relevant sensory system. Our findings confirm that for action observation as well, brain areas exhibit a hierarchy of TRWs for processing goal-directed action information. As expected, frontoparietal and lateral occipital areas, traditionally considered to encode high-level conceptual information in both action observation^{6,7,10,12,13,48} and execution^{39,49–52}, exhibit relatively long TRWs. Interestingly, our results further show that several regions traditionally considered as primary areas exhibit relatively long TRWs as well. Specifically, we find that the primary somatosensory cortex and early visual cortex (EVC) in the right hemisphere exhibit relatively long TRWs of ± 10 s and even longer. Since action observation engages motor, somatosensory and visual pathways^{1,2,42,53–55} the interaction between the different systems may have an effect on the functional characteristics of neural responses within the different areas.

Previous studies have demonstrated the involvement of primary somatosensory cortex (S1) during action observation^{1,56,57}. It was suggested that S1, namely BA2, is recruited by the vicarious aspects of the observed action. This may indicate an instantaneous activation, locked to local somatosensory features at the sight of other people's actions⁵⁶. Interestingly however, it was demonstrated that prior knowledge regarding the goal of an observed action modulates activity within the somatosensory cortex. This was evident even when the object-goal of the action was not visible⁵⁷. This can be taken to suggest that the somatosensory representation of action goals is driven not only by the instantaneously perceived physical features of the movement, but also by the global goal information which is continuous and stable over time. Our findings of long TRWs in right S1 are in line with this notion, indicating the region's capacity to integrate over goal related action information in a relatively long time window rather than only respond to momentary sensory features of the stimuli.

It was previously suggested that predictive feedback to the occipital cortex from the motor cortex during action execution⁵⁸, and from premotor and parietal areas during observation^{20,21} modulate the neural representation of the visual input such that visual responses represent and send feedforward action prediction errors. Accordingly, observing a scrambled presentation of motor acts hinders goal predictability and therefore elevates prediction errors. This, in turn, might induce divergent neural activity across subjects and, therefore, a reduction in ISC. In addition to feedback connections from the motor cortex, it was shown that primary visual cortex receives strong feedback connections from MT+, LOC and the PFm of the parietal lobe and that high-level information is encoded in EVC during evidence accumulation, prediction and spatial attention^{47,59,60}. The involvement of the EVC in such complex perceptual operations was suggested to reflect top-down modulation of activity in primary visual cortex by feedback projections⁶⁰. Moreover, it was suggested that such feedback projections from higher hierarchical levels modulate the temporal dynamics of lower levels^{61,62}. Thus, the existence of such top-down modulation by feedback connections may play a role in the reduced ISC levels found in the current study for scrambled actions, and the corresponding long TRWs in primary sensory regions.

Hemispheric biases

In addition to the differences in processing scales across frontoparietal and occipital regions, our results further revealed differences in TRWs between hemispheres. Specifically, longer TRWs were found in primary somatosensory, MT+, and early visual areas (msOccG and MVOcC) of the right hemisphere. We did not initially expect to find such right lateralization for longer TRWs. Yet, a right hemispheric dominance for processing static frames depicting coherent vs. non-coherent visual events, was previously shown in areas including the postcentral gyrus as well as middle and posterior superior temporal cortex⁶³. Hasson and colleagues directly examined the topographical mapping of TRWs in the visual system using dynamic stimuli, however their results were averaged across hemispheres¹⁷. Cerliani and colleagues directly examined sensitivity to temporal structures

in action observation, however, their analysis was focused on specific regions only in the left hemisphere (Pft, BA44, lateral occipital cortex)²¹.

Taken together, the current results point to a potential advantage of the right hemisphere in the temporal capacity to integrate over observed action sequence information. Hemispheric asymmetry is known for spatial processing of visual input with a right hemispheric specialization for global vs. local processing due to greater sensitivity to low vs. high spatial frequencies in the right compared to the left hemisphere⁶⁴. This is also supported by evidence from patients showing that left hemisphere lesions are associated with deficits in processing local visual features, while patients with right hemisphere lesions are more impaired in global visual processing⁶⁵. A right hemispheric advantage to slow frequencies and hence to global processing in the temporal, as in the spatial aspect of visual inputs, might account for the pattern of results found in the current study. Nevertheless, the question of hemispheric differences in processing scales of observed action information remains for future studies.

Note that due to the approach we used to construct the scrambled videos from the intact video, the four videos differ with respect to low-level visual transients. Thus, the amount of visual transients is greater in the Primitive scramble and is progressively less in the other conditions. It is possible that visual transients induce an evoked response that is common across subjects, thus increasing ISC. In this case, our results would be biased towards higher ISC levels in the scrambled conditions and consequently shift the identification of regions towards having low TRWs. Intriguingly, in early visual cortices the current results show a significant *decrease*, rather than increase, in ISC for the shortest scrambling timescale (Primitives \pm 1.5 s) which is compatible with a long TRW and sensitivity to observed action coherency rather than synchronization that is locked to low-level visual transients. In addition, a modulation of ISC due to visual transients would be similar for the two hemispheres. However, the current results show longer TRWs in the right compared to the left hemisphere. Therefore, low-level visual features are less likely to be linked to the increased ISC during the Intact video presentation in the right hemisphere and not in the left. Taken together, the increased number of visual transients in the scrambled conditions seems less likely to explain the pattern of results we obtained in visual cortex. Previous studies, have equated the number of visual transients across conditions by using video recordings from two camera angles and changing the viewpoint angle after each primitive action^{20,21}. In the current study we did not implement this approach since we were interested in keeping the intact video, and coherent segments in the scrambled conditions, as natural and similar to action stimuli that we normally perceive in real life.

The current study established the existence of a TRW hierarchy for processing goal-directed actions. While here we show four different time-scales for processing baking-related, goal-directed actions, future investigation are advisable in order to provide a complete characterization of the hierarchy manifestation and its developmental trajectory. Specifically, to understand how the TRWs develop and are shaped by experience, what is the upper limit of processing time-window across regions and whether the hierarchy found here can also be generalized to processing of other types of actions such as intransitive non-goal directed actions or to other contextual domains.

Conclusion

The current study systematically examined four temporal scales of action goals and demonstrates that, instead of a uniform network, areas in sensory-motor pathways exhibit graded neuro-physiological characteristics of processing capacities during the integration of observed goal-directed action sequences. Our findings delineate a range of timescales across action observation areas showing that the right posterior IPL, and the left IPS and PMd, exhibit the longest TRW, compared to relatively short TRWs in the PMv and primary motor cortices. This is in line with conceptual goal representation previously found in parietal regions, while motor primitives are more frequently associated in the literature with the PMv. Moreover, the results point to a right-hemisphere bias for long TRWs, mainly in the medial occipital cortex. Taken together, the current results shed light on the neural architecture of action organization, support the theory of temporal hierarchical representation of actions (Uithol et al., 2012) and contribute to arguments of predictive coding mechanisms for neural communication between action observation areas. The findings may further provide insight and help delineate dysfunctions in motor disorders such as Ideational Apraxia, in which action goal representation is compromised and can inspire the development of computational models for human action understanding and recognition.

Data availability

According to the conditions of our ethics approval, the data of this study cannot be uploaded to open repository, however, anonymized data can be provided upon request.

Code availability

Code for data analysis is available at https://github.com/ShaharAb/Temporal_Hierarchy_study.

Received: 5 July 2023; Accepted: 7 November 2023

Published online: 11 November 2023

References

1. Caspers, S., Zilles, K., Laird, A. R. & Eickhoff, S. B. ALE meta-analysis of action observation and imitation in the human brain. *Neuroimage* **50**, 1148–1167 (2010).
2. Molenberghs, P., Cunnington, R. & Mattingley, J. B. Brain regions with mirror properties: A meta-analysis of 125 human fMRI studies. *Neurosci. Biobehav. Rev.* **36**, 341–349 (2012).
3. Cavallo, A., Koul, A., Ansuini, C., Capozzi, F. & Becchio, C. Decoding intentions from movement kinematics. *Sci. Rep.* **6**, 1–8 (2016).
4. Koul, A. et al. Action observation areas represent intentions from subtle kinematic features. *Cereb. Cortex* **28**, 2647–2654 (2018).

5. Johnson-Frey, S. H. *et al.* Actions or hand-object interactions? Human inferior frontal cortex and action observation. *Neuron* **39**, 1053–1058 (2003).
6. Iacoboni, M., Molnar-Szakacs, I., Gallese, V., Buccino, G. & Mazziotta, J. C. Grasping the intentions of others with one's own mirror neuron system. *PLoS Biol.* **3**, 0529–0535 (2005).
7. Hamilton, A. F. de C. & Grafton, S. T. Action outcomes are represented in human inferior frontoparietal cortex. *Cereb. Cortex* **18**, 1160–1168 (2008).
8. Loucks, J., Mutschler, C. & Meltzoff, A. N. Children's representation and imitation of events: How goal organization influences 3-year-old children's memory for action sequences. *Cogn. Sci.* **41**, 1904–1933 (2017).
9. Nicholson, T., Roser, M. & Bach, P. Understanding the goals of everyday instrumental actions is primarily linked to object, not motor-kinematic, information: Evidence from fMRI. *PLoS ONE* **12**, e0169700 (2017).
10. Hamilton, A. F. D. C. & Grafton, S. T. Goal representation in human anterior intraparietal sulcus. *J. Neurosci.* **26**, 1133–1137 (2006).
11. Grafton, S. T. & Hamilton, A. F. D. C. Evidence for a distributed hierarchy of action representation in the brain. *Hum. Movem. Sci.* **26**, 590–616 (2007).
12. Wurm, M. F. & Lingnau, A. Decoding actions at different levels of abstraction. *J. Neurosci.* **35**, 7727–7735 (2015).
13. Wurm, M. F., Ariani, G., Greenlee, M. W. & Lingnau, A. Decoding concrete and abstract action representations during explicit and implicit conceptual processing. *Cereb. Cortex* **26**, 3390–3401 (2016).
14. Uithol, S., Van Rooij, I., Bekkering, H. & Haselager, P. Hierarchies in action and motor control. *J. Cogn. Neurosci.* **24**, 1077–1086 (2012).
15. Hasson, U., Nir, Y., Levy, I., Fuhrmann, G. & Malach, R. Intersubject synchronization of cortical activity during natural vision. *Science* **303**, 1634–1640 (2004).
16. Nastase, S. A., Gazzola, V., Hasson, U. & Keysers, C. Measuring shared responses across subjects using intersubject correlation. *Soc. Cogn. Affect. Neurosci.* **14**, 669–687 (2019).
17. Hasson, U., Yang, E., Vallines, I., Heeger, D. J. & Rubin, N. A hierarchy of temporal receptive windows in human cortex. *J. Neurosci.* **28**, 2539–2550 (2008).
18. Hasson, U., Chen, J. & Honey, C. J. Hierarchical process memory: Memory as an integral component of information processing. *Trends Cogn. Sci.* **19**, 304–313 (2015).
19. Lerner, Y., Honey, C. J., Silbert, L. J. & Hasson, U. Topographic mapping of a hierarchy of temporal receptive windows using a narrated story. *J. Neurosci.* **31**, 2906–2915 (2011).
20. Thomas, R. M., De Sanctis, T., Gazzola, V. & Keysers, C. Where and how our brain represents the temporal structure of observed action. *NeuroImage* **183**, 677–697 (2018).
21. Cerliani, L. *et al.* Predictive coding during action observation – A depth-resolved intersubject functional correlation study at 7T. *Cortex* **148**, 121–138 (2022).
22. Yamashita, Y. & Tani, J. Emergence of functional hierarchy in a multiple timescale neural network model: A humanoid robot experiment. *PLoS Comput. Biol.* **4**, e1000220 (2008).
23. Pastra, K. & Aloimonos, Y. The minimalist grammar of action. *Philos. Trans. R. Soc. B Biol. Sci.* **367**, 103–117 (2012).
24. Kuehne, H., Arslan, A. & Serre, T. The language of actions: Recovering the syntax and semantics of goal-directed human activities. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition* 780–787. <https://doi.org/10.1109/CVPR.2014.105> (2014).
25. Binkofski, F. C., Klann, J. & Caspers, S. On the neuroanatomy and functional role of the inferior parietal lobule and intraparietal sulcus. In *Neurobiology of language* 35–47 (Elsevier, 2016). <https://doi.org/10.1016/B978-0-12-407794-2.00004-3>.
26. Hamilton, A. F. C. Goals, intentions and mental states: Challenges for theories of autism. *J. Child Psychol. Psychiatry Allied Discip.* **50**, 881–892 (2009).
27. Jastorff, J., Begliomini, C., Fabbri-Destro, M., Rizzolatti, G. & Orban, G. A. Coding observed motor acts: Different organizational principles in the parietal and premotor cortex of humans. *J. Neurophysiol.* **104**, 128–140 (2010).
28. Smith, S. M. *et al.* Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage* **23**, S208–S219 (2004).
29. Bandettini, P. A., Jesmanowicz, A., Wong, E. C. & Hyde, J. S. Processing strategies for time-course data sets in functional MRI of the human brain. *Magn. Reson. Med.* **30**, 161–173 (1993).
30. Benjamini, Y. & Hochberg, Y. Benjamini Y, Hochberg Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Stat. Soc. B* 289–300 (1995).
31. Fan, L. *et al.* The human brainnetome atlas: A new brain atlas based on connectional architecture. *Cereb. Cortex* **26**, 3508–3526 (2016).
32. Hardwick, R. M., Caspers, S., Eickhoff, S. B. & Swinnen, S. P. Neural correlates of action: Comparing meta-analyses of imagery, observation, and execution. *Neurosci. Biobehav. Rev.* **94**, 31–44 (2018).
33. Hesse, M. D., Sparing, R. & Fink, G. R. End or means—the “what” and “how” of observed intentional actions. *J. Cogn. Neurosci.* **21**, 776–790 (2009).
34. Cavanna, A. E. & Trimble, M. R. The precuneus: A review of its functional anatomy and behavioural correlates. *Brain* **129**, 564–583 (2006).
35. Matsumoto, K. & Tanaka, K. Conflict and cognitive control. *Science* **303**, 969–970 (2004).
36. Sreenivasan, K. K., Curtis, C. E. & D'Esposito, M. Revisiting the role of persistent neural activity during working memory. *Trends Cogn. Sci.* **18**, 82–89 (2014).
37. Hoshi, E. & Tanji, J. Distinctions between dorsal and ventral premotor areas: Anatomical connectivity and functional properties. *Curr. Opin. Neurobiol.* **17**, 234–242 (2007).
38. Majdandzic, J., Bekkering, H., Van Schie, H. T. & Toni, I. Movement-specific repetition suppression in ventral and dorsal premotor cortex during action observation. *Cereb. Cortex* **19**, 2736–2745 (2009).
39. Turella, L., Rumiati, R. & Lingnau, A. Hierarchical action encoding within the human brain. *Cereb. Cortex* **30**, 2924–2938 (2020).
40. Errante, A., Ziccarelli, S., Mingolla, G. P. & Fogassi, L. Decoding grip type and action goal during the observation of reaching-grasping actions: A multivariate fMRI study. *NeuroImage* **243**, 118511 (2021).
41. Hamilton, A. F. de C. & Grafton, S. T. The motor hierarchy: from kinematics to goals and intentions. *Sensorimotor Found. High. Cogn.* **22**, 381–408 (2007).
42. Molenberghs, P., Hayward, L., Mattingley, J. B. & Cunnington, R. Activation patterns during action observation are modulated by context in mirror system areas. *NeuroImage* **59**, 608–615 (2012).
43. Haaland, K. Y., Harrington, D. L. & Knight, R. T. Neural representations of skilled movement. *Brain* **123**, 2306–2313 (2000).
44. Koski, L., Iacoboni, M. & Mazziotta, J. C. Deconstructing apraxia: understanding disorders of intentional movement after stroke. *Curr. Opin. Neurol.* **15**, 71–77 (2002).
45. Goldenberg, G. Apraxia and the parietal lobes. *Neuropsychologia* **47**, 1449–1459 (2009).
46. Parvizi, J. & Wagner, A. D. Memory, numbers, and action decision in human posterior parietal cortex. *Neuron* **97**, 7–10 (2018).
47. Rolls, E. T., Deco, G., Huang, C.-C. & Feng, J. The human posterior parietal cortex: effective connectome, and its relation to function. *Cereb. Cortex* **10**, 1–29 (2022).
48. Aflalo, T. *et al.* A shared neural substrate for action verbs and observed actions in human posterior parietal cortex. *Sci. Adv.* **6**, eabb3984 (2020).

49. Gallivan, J. P., Adam McLean, D., Smith, F. W. & Culham, J. C. Decoding effector-dependent and effector-independent movement intentions from human parieto-frontal brain activity. *J. Neurosci.* **31**, 17149–17168 (2011).
50. Krasovskiy, A., Gilron, R., Yeshurun, Y. & Mukamel, R. Differentiating intended sensory outcome from underlying motor actions in the human brain. *J. Neurosci.* **34**, 15446–15454 (2014).
51. Gertz, H., Lingnau, A. & Fiehler, K. Decoding movement goals from the fronto-parietal reach network. *Front. Hum. Neurosci.* **11**, 84 (2017).
52. Aberbach-Goodman, S., Buaron, B., Mudrik, L. & Mukamel, R. Same action, different meaning: Neural substrates of action semantic meaning. *Cereb. Cortex* 1–11. <https://doi.org/10.1093/cercor/bhab483> (2022).
53. Gazzola, V. & Keysers, C. 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**, 1239–1255 (2009).
54. Mukamel, R., Ekstrom, A. D., Kaplan, J., Jacoboni, M. & Fried, I. Single-neuron responses in humans during execution and observation of actions. *Curr. Biol.* **20**, 750–756 (2010).
55. Valchev, N., Gazzola, V., Avenanti, A. & Keysers, C. Primary somatosensory contribution to action observation brain activity—combining fMRI and cTBS. *Soc. Cogn. Affect. Neurosci.* **11**, 1205–1217 (2016).
56. Keysers, C., Kaas, J. H. & Gazzola, V. Somatosensation in social perception. *Nat. Rev. Neurosci.* **11**, 417–428 (2010).
57. Turella, L., Tubaldi, F., Erb, M., Grodd, W. & Castiello, U. Object presence modulates activity within the somatosensory component of the action observation network. *Cerebral Cortex* **22**, 668–679 (2012).
58. Buaron, B., Reznik, D., Gilron, R. & Mukamel, R. Voluntary actions modulate perception and neural representation of action-consequences in a hand-dependent manner. *Cerebral Cortex* **30**(12), 6097–6107 (2020).
59. Muckli, L. & Petro, L. S. Network interactions: Non-geniculate input to V1. *Curr. Opin. Neurobiol.* **23**, 195–201 (2013).
60. Roelfsema, P. R. & de Lange, F. P. Early visual cortex as a multiscale cognitive blackboard. *Annu. Rev. Vis. Sci.* **2**, 131–151 (2016).
61. Jiang, L. P. & Rao, R. P. N. *Dynamic Predictive Coding: A New Model of Hierarchical Sequence Learning and Prediction in the Cortex.* <https://doi.org/10.1101/2022.06.23.497415> (2022).
62. Qin, C. *et al.* Predictive coding during action observation revealed by human electrocorticographic activity. <https://doi.org/10.1101/2022.12.13.519734> (2022).
63. Han, S., Jiang, Y. & Mao, L. Right hemisphere dominance in perceiving coherence of visual events. *Neurosci. Lett.* **398**, 18–21 (2006).
64. Fink, G. R. *et al.* Neural mechanisms involved in the processing of global and local aspects of hierarchically organized visual stimuli. *Brain* **120**, 1779–1791 (1997).
65. Robertson, L. C. & Lamb, M. R. Neuropsychological contributions to theories of part/whole organization. *Cogn. Psychol.* **23**, 299–330 (1991).

Acknowledgements

The authors thank lab members for constructive comments and fruitful suggestions.

Author contributions

S.A.G: conceptualization, investigation, data curation, formal analysis, software, methodology, validation, visualization, writing original draft, writing—review and editing. R.M: supervision, conceptualization, methodology, validation, writing original draft, writing—review and editing, resources, funding acquisition.

Funding

This research was supported by the Israel Science Foundation (grant No. 2392/19 to R.M.). The funding sources had no role in study design, collection, analysis, interpretation, or decision to submit this paper.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-023-46917-z>.

Correspondence and requests for materials should be addressed to R.M.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023