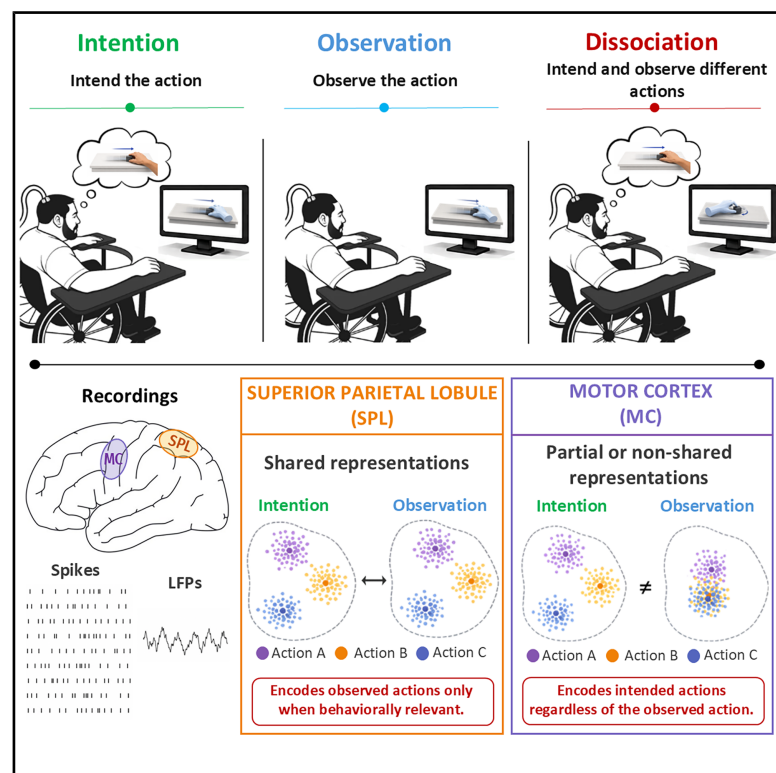


Hierarchical and context-dependent encoding of actions in human posterior parietal and motor cortex

Graphical abstract



Authors

Vasiliki Bougou, Jorge Gamez, Emily R. Rosario, Charles Liu, Kelsie Pejisa, Ausaf Bari, Richard A. Andersen

Correspondence

richard.andersen@vis.caltech.edu

In brief

Interpreting others' actions is widely thought to rely on mirroring, yet direct neurophysiological evidence in humans remains scarce. Using single-neuron recordings from motor and posterior parietal cortex in tetraplegic participants, this study reveals hierarchical, context-dependent action coding: motor cortex encodes intentions, while posterior parietal cortex links observed actions to goals.

Highlights

- Superior parietal lobule (SPL) encodes intended and observed actions via a shared code
- Motor cortex neurons encode intention, but the population retains a latent observation code
- SPL flexibly gates observed actions based on behavioral relevance
- Frontoparietal cortex shows hierarchical and context-dependent action encoding



Article

Hierarchical and context-dependent encoding of actions in human posterior parietal and motor cortex

Vasiliki Bougou,¹ Jorge Gamez,¹ Emily R. Rosario,³ Charles Liu,^{3,4} Kelsie Pejisa,¹ Ausaf Bari,⁵ and Richard A. Andersen^{1,2,6,*}

¹Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA 91125, USA

²T&C Chen Brain-Machine Interface Center, California Institute of Technology, Pasadena, CA 91125, USA

³Casa Colina Hospital and Center for Healthcare, Pomona, CA 91767, USA

⁴Department of Neurological Surgery, Keck School of Medicine of USC, Los Angeles, CA 90033, USA

⁵Department of Neurological Surgery, University of California, Los Angeles, Los Angeles, CA 90095, USA

⁶Lead contact

*Correspondence: richard.andersen@vis.caltech.edu

<https://doi.org/10.1016/j.cell.2026.07.032>

SUMMARY

Action understanding requires internal models that link visual input to motor goals. In monkeys, mirror neurons have been proposed to support this process through motor resonance during observation, yet single-unit evidence in humans remains scarce. We recorded neural activity from the motor cortex (MC) and the superior parietal lobule (SPL) in two tetraplegic participants implanted with Utah arrays while they intended or observed hand actions. MC strongly encoded intention but showed weak responses during observation, evident primarily at the population level. In contrast, SPL supported shared representations across intended and observed actions at both the single-unit and population levels. When incongruent instructed and observed actions occurred simultaneously, SPL encoded the observed actions only when they were behaviorally relevant, whereas MC remained intention dominant. Our results identify a context-dependent gating mechanism in SPL and suggest a hierarchical organization in which MC maintains intention-specific codes while SPL flexibly links observed actions with internal goals.

INTRODUCTION

The ability to interpret the actions of others, to identify what is being done, by whom, and in what context, is a fundamental feature of intelligent behavior. It relies on internal representations that can flexibly map visual input onto motor knowledge, enabling the brain to simulate, predict, and interpret observed movements. For decades, these processes have been framed through the lens of mirror mechanisms: the idea that shared neural codes support both action execution and observation.^{1–3} Yet the scope, structure, and cortical distribution of these representations remain the subject of considerable debate.⁴ Is action observation associated with motor-constrained reactivations of one's own movements or with more abstract, generalizable codes that extend beyond the observer's motor repertoire? And how are these representations organized across regions traditionally implicated in sensory-motor control versus those thought to support higher-order cognitive functions?

Mirror neurons, first identified in macaques, respond during both execution and observation, supporting the motor simulation theory that action recognition relies on covert motor reactivation.^{1,2,5,6} Although influential, this framework has been refined

by evidence that mirror responses are shaped by prior experience and task context (contextual flexibility), that individual units often encode only subsets of action features (partial tuning), and that visuomotor congruence emerges more robustly at the ensemble than at the single-cell level (population encoding).⁷ Moreover, several neurophysiological studies suggest that observation-related responses in motor regions may reflect processes involved in motor preparation or suppression,^{8–10} in line with evidence linking this system to imitation, social facilitation, and broader action-affordance phenomena.^{7,11} These findings point to more complex and heterogeneous mechanisms than strict motor mirroring can explain. Converging theoretical work suggests that cognitive flexibility emerges from internal models that are compositional, building complex structures from separable and recombinable elements, and generalizable, capturing abstract structure that extends across contexts and tasks to support adaptive behavior.^{12,13} Yet in humans, direct single-neuron evidence remains scarce, and the broader principles underlying action representations, particularly their flexibility and generalizability, are still not well defined.

These theoretical shifts, from rigid mirroring to flexible, predictive, and context-sensitive encoding, highlight the need to



reassess how action representations are structured in the human brain. Action observation provides a unique lens into this flexibility. Unlike execution, it allows testing whether neural populations encode actions disentangled from the specific motor commands or the output normally coupled to them. This is particularly relevant in higher-order regions such as the parietal cortex, where mixed selectivity is prevalent^{14–18} and encoding may reflect abstract and goal-directed representations rather than motor-specific plans.^{19–21} Assessing such generalizability requires moving beyond single-neuron selectivity toward the structure of population codes.

Recent advances in neural manifold analysis²² and representational geometry^{23,24} (the structure of neural population activity, often captured as trajectories) have shown that population activity can reveal compositional and generalizable subspaces, but these approaches have only recently started to be applied to investigate action representations.^{25,26} To date, only a limited number of studies have examined action observation responses at the single-neuron level in humans. Observation-related activity has been reported in medial frontal and temporal regions²⁷ and in the posterior parietal cortex.^{25,28} In the motor cortex (MC), intracortical recordings have shown that movement-related activity is present during observation, imagery, and attempted movement but that population structure is primarily organized by volitional state.^{29,30} However, whether representational geometry generalizes across intention and observation, and how such generalization differs between core regions of the motor system, remains unclear.

To investigate how intended and observed actions are represented in the human brain, we recorded single-unit activity (SUA) and multiunit activity (MUA) together with local field potentials (LFPs) from the motor and posterior parietal cortices in two tetraplegic participants implanted with Utah arrays. Using a factorial design that manipulated hand, action type, and movement direction, we compared neural representations across two task formats, one in which participants intended actions and another in which they passively observed them. Our analyses revealed a gradient of representational overlap, with the posterior parietal cortex encoding action identity in a format-general manner across intention and observation, whereas MC representations were predominantly intention-specific, with observation responses reflecting only a latent projection of the intention-related structure. To test whether this overlap was modulated by behavioral relevance, we introduced a dissociation task that decoupled instructed and observed actions. This manipulation demonstrated that parietal representations of observed actions emerged only when those actions were behaviorally relevant, consistent with a gating mechanism driven by task demands, whereas MC consistently reflected only the instructed movement. Together, these findings provide direct human electrophysiological evidence that action representations extend beyond motor mirroring, supporting a hierarchical and context-dependent coding framework in the frontoparietal system.

RESULTS

We recorded SUA and MUA as well as LFPs from two tetraplegic participants (JJ and RD) implanted with Utah arrays targeting the

motor and posterior parietal cortices. Participant JJ had 96-channel arrays in the hand knob area of the primary MC and superior parietal lobule (SPL). Participant RD had 64-channel arrays in the medial motor cortex (MCM), lateral motor cortex (MCL), and SPL. An additional array targeting the supramarginal gyrus (SMG) in RD was excluded from all analyses. Signal quality and unit isolation were comparable to those of the other arrays; however, activity in this region did not exhibit reliable task-related modulation across sessions or conditions.

To investigate action encoding across intention and observation, we designed a fully crossed 2 (hand: left/right) $\times$ 3 (action: lift, slide, rotate) $\times$ 2 (direction: left/right) design (Figure 1A). Direction refers to an abstract left/right factor defined for each action type (see STAR Methods). In the intention format, participants intended the cued action; in the observation format, they passively viewed the same action. This design allowed independent assessment of effector, action type, and direction encoding across formats.

Unit counts per array are reported in Tables S1, S3, and S4; implant locations are shown in Figure 1B. Based on the quality factor classification (see STAR Methods), 81 units in MCM/RD, 109 in MCL/RD, and 92 in SPL/RD were well isolated (quality 1–2), with the remainder classified as MUA (quality 3–4). In JJ, whose arrays had been implanted for an extended period, the majority of signals were classified as MUA, yielding only 5 well-isolated MC units and 4 well-isolated SPL units. The dissociation tasks yielded 134 well-isolated units in MCM/RD, 107 in MCL/RD, and 254 in SPL/RD. Because recordings were obtained from the same arrays across multiple sessions, we evaluated potential re-identification of units across days in RD; the median proportion of putative matches was 12.0% (range 0.0%–44.4%) (Figure S1), indicating limited cross-session overlap.

Eye-tracking confirmed stable fixation across conditions (Figure S2). Gaze features neither predicted MUA (R^2 not different from shuffled controls, corrected $p > 0.05$) nor reliably decoded task variables, confirming that neural responses were not explained by eye movements.

Distinct tuning profiles in parietal and motor cortices across action formats

We first examined condition-specific responses at the single-unit level across regions. Figure 2A shows example units from RD: one in SPL tuned to action “lift” and one in MCM tuned to right-hand actions. Both showed consistent tuning across formats, with stronger responses during intention than observation, though cross-format tuning was not ubiquitous; selectivity often differed between formats.

Response latencies were computed per condition (Figure 2B; see STAR Methods). During intention, SPL responses preceded MC (median: SPL 1.2 s, MC 1.9 s; $p < 0.05$); no significant timing difference was observed during observation (SPL 1.4 s, MC 1.5 s). This pattern held within each participant (Figure S3A), though in JJ, SPL also significantly preceded MC during observation. These results point to a leading role of SPL during action intention, with less consistent dynamics during observation.

We quantified the number of units responsive to at least one condition in each format and brain area, including both increases and decreases in firing rate (Figures 2C and S3B; see STAR

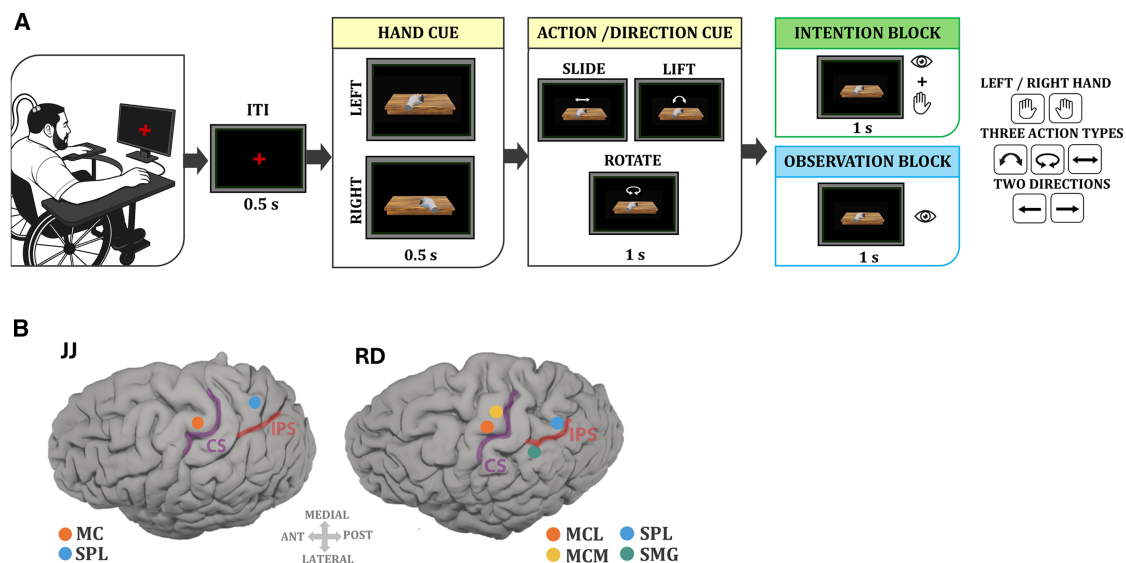


Figure 1. Experimental design and implant locations

(A) Schematic of the action intention and observation task. Participants either intended or passively observed hand actions comprising 12 conditions (2 hands $\times$ 3 actions $\times$ 2 directions).

(B) Implant locations. Participant JJ had 96-channel arrays targeting the MC and SPL. Participant RD had four 64-channel arrays targeting the medial and lateral motor cortex (MCM and MCL) and the posterior parietal cortex (SPL and SMG).

Methods. In SPL, 395 units were responsive during intention and 225 during observation, whereas in MC, 607 responded during intention and 153 during observation. Relative to the total recorded units (MC: 1341; SPL: 858), this corresponds to 45.3% and 11.4% in MC, and 46.0% and 26.2% in SPL for intention and observation, respectively. Thus, responsiveness during intention was comparable across regions, but SPL showed a higher proportion of observation-responsive units. This pattern was consistent across participants, although JJ's MC intention responses were concentrated for right-hand rotation, whereas RD's motor arrays showed broader tuning with a right-hand bias (Figure S3B). Both facilitated and suppressed responses were observed (Figure S3H), with facilitation generally predominating except for left-hand intention trials in MC.

To assess tuning to specific task features, we performed a three-way ANOVA on binned firing rates with action type, effector (hand), and direction as factors (Figures 2D and S3C). In MC during intention, effector tuning peaked at 2.25 s (145 units), followed by action tuning at 2.75 s (167 units). In SPL, action was the most prevalent feature (77 units at 2.25 s), followed by effector (42 units). Tuning profiles differed across participants (Figure S3C): JJ's MC was dominated by action tuning, consistent with the concentration of rotation responses (Figure S3B), whereas RD's motor arrays showed stronger effector tuning. During observation, SPL preserved action tuning (33 units at 2.75 s), while MC exhibited minimal observation-related tuning (7 action-tuned units). Interaction effects were rare, indicating predominantly single-feature modulation. Overall, MC was dominated by effector or action tuning during intention, whereas SPL consistently encoded action type across formats.

We next quantified format-specific responsiveness by counting units active for at least one condition during intention only,

observation only, or both (Figures 2E and S3D). Percentages are reported relative to the total number of units that were responsive in at least one format within each region (i.e., the union of intention- and observation-responsive units; MC: 676, SPL: 458). In MC, most units were responsive exclusively during intention (77.4%), compared with 10.2% during observation only and 12.4% responsive in both formats. In contrast, SPL showed a higher proportion of units responsive across both formats (35.4%), with 50.9% intention only and 13.7% observation only. The distribution of response categories differed significantly between regions ($\chi^2(2) = 97.96$, $p < 0.05$). Similar patterns were observed within each participant (Figure S3D). Condition-averaged activity for each response category is visualized as heatmaps aggregated by area (Figure S3G).

To characterize tuning structure and generalization, we applied a linear model framework adapted from Chivukula et al. (2025),²⁵ classifying units as showing consistent tuning across formats (shared), different tuning (idiosyncratic), tuning in only one format (single-format), or no reliable selectivity (unselective) (see STAR Methods). Figure 2F shows the distribution of model classifications across regions, and Figure S3E presents the results separately for each array. Percentages are reported relative to all units that exhibited significant tuning in at least one format (per region: MC: 709, SPL: 375; per participant: MC/JJ: 285, MCM/RD: 235, MCL/RD: 189, SPL/JJ: 137, SPL/RD: 238). MC and SPL differed significantly in their distribution of tuning categories ($\chi^2(5) = 77.02$, $p < 0.05$). In MC, the majority of units were classified as single format (52.6%), followed by shared action (21.3%) and shared hand (10.9%), with smaller proportions of idiosyncratic (8.3%) and invariant units (5.8%). Notably, JJ's MC showed a relatively high proportion of shared-action models (41%), nearly comparable to single-

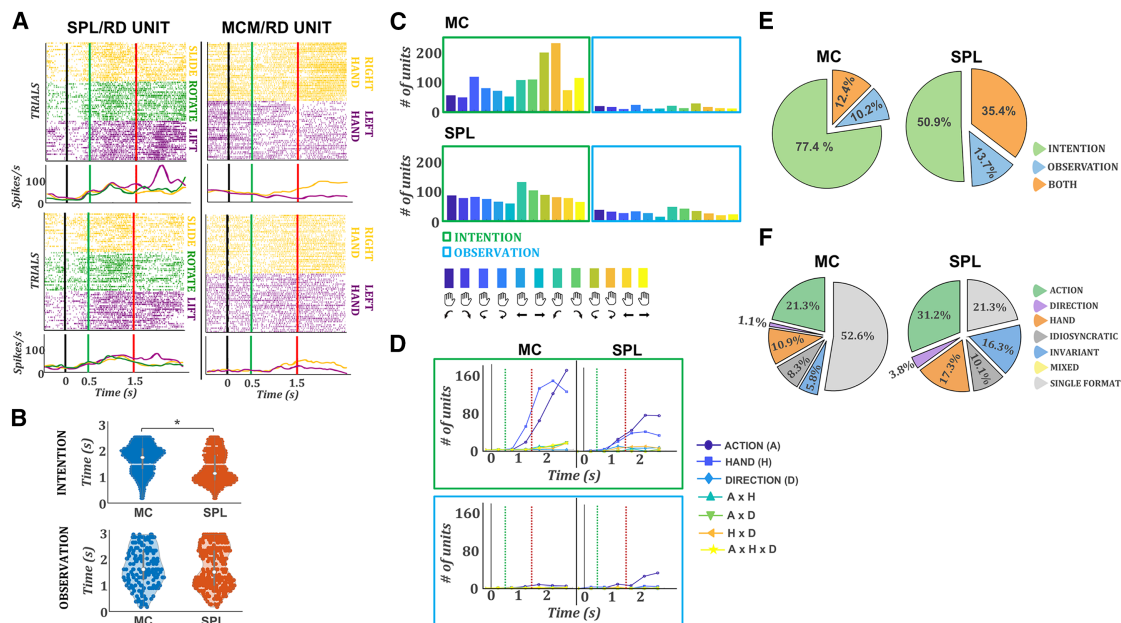


Figure 2. Single-unit selectivity across intention and observation

(A) Example neurons from RD in SPL (left) and MCM (right). Top: intention; bottom: observation. SPL trials are color coded by action (slide, rotate, and lift); MCM trials are color coded by hand (right, left). Raster plots and corresponding peristimulus time histograms (PSTHs) are shown below. Vertical lines indicate the hand cue (black), action cue (green), and go cue (red). The SPL neuron is action-selective across formats, whereas the MCM neuron is selective for the right hand. (B–F) Results pooled across arrays within each region (MC: MC/JJ, MCM/RD, MCL/RD; SPL: SPL/JJ, SPL/RD); per-array results are shown in Figure S3. (B) Violin plots of onset latencies for intention (top) and observation (bottom). Asterisks indicate statistically significant differences ($p < 0.05$). (C) Number of responsive units per condition during intention (green outline) and observation (blue outline). (D) Three-way ANOVA tuning profiles over time (factors: action, hand, and direction); intention (green outline), observation (blue outline). Lines represent main effects and all two- and three-way interactions. Vertical lines indicate the hand cue (black), action cue (green), and go cue (red). (E) Format selectivity: percentage of units responsive in intention only (green), observation only (blue), or both (orange). (F) Linear model classification based on cross-validated R^2 . Units are assigned to shared (action, hand, and direction), invariant, mixed, idiosyncratic, or single-format categories.

format units (45.6%), whereas RD's motor arrays (MCM and MCL) were clearly dominated by single-format units, which exceeded 50% in both (Figure S3E). SPL exhibited consistent shared structure across intention and observation. Shared action was the most frequent model (31.2%), followed by shared hand (17.3%), invariant (16.3%), and idiosyncratic (10.1%). A comparable distribution of models was observed in each participant individually, with a nearly identical ranking of model types (Figure S3E). These findings further support the view that SPL contains a population of units that encode action features in a consistent and generalizable format across observation and intention.

To ensure that collapsing leftward and rightward variants of each action did not inflate shared-action fits, we repeated the model comparison using a 6-level action factor (action $\times$ direction) (Figure S3F). The overall pattern remained unchanged, indicating that shared-action tuning, particularly in SPL, reflected genuine action identity encoding rather than directional overlap.

We tested whether signal type influenced results by comparing SUA and MUA in RD (Figures S4A, S4C, and S4E). Distributions of significant responses, format-specific proportions, and linear model classifications did not differ between signal types (all $p > 0.05$), though MCM MUA showed a numerically higher proportion of shared-hand tuning than SUA, consistent with the greater

shared structure in JJ's MUA-dominated MC. SUA and MUA were therefore pooled for subsequent analyses. Results were also qualitatively unchanged after removing putatively matched cross-session units (Figures S4B, S4D, and S4F).

Overall, SPL exhibited stable shared structure across intention and observation in both participants, with a substantial proportion of units responsive in both formats and shared action emerging as the most frequent model. In contrast, MC responses were predominantly intention-specific, and shared structure was comparatively limited.

Cross-format representational similarity is robust in SPL but conditional or absent in MC

Single-unit analyses revealed shared structure in SPL but largely intention-dominant, single-format responses in MC. We next examined population-level structure using representational similarity analysis (RSA) (see STAR Methods).

Cross-format RSA matrices for each array and task variable are shown in Figure 3A (within-format matrices in Figures S5A and S5B; time-window robustness in Figure S5C). In JJ's MC, cross-format similarity was strongest for rotation ($R^2 = 0.3$); lift and slide were near zero. In RD's MCM, correlations were strong for the right hand (0.6) and moderate for lift and rotation (0.4); left-hand and slide were weaker (≤ 0.3). RD's MCL showed no meaningful

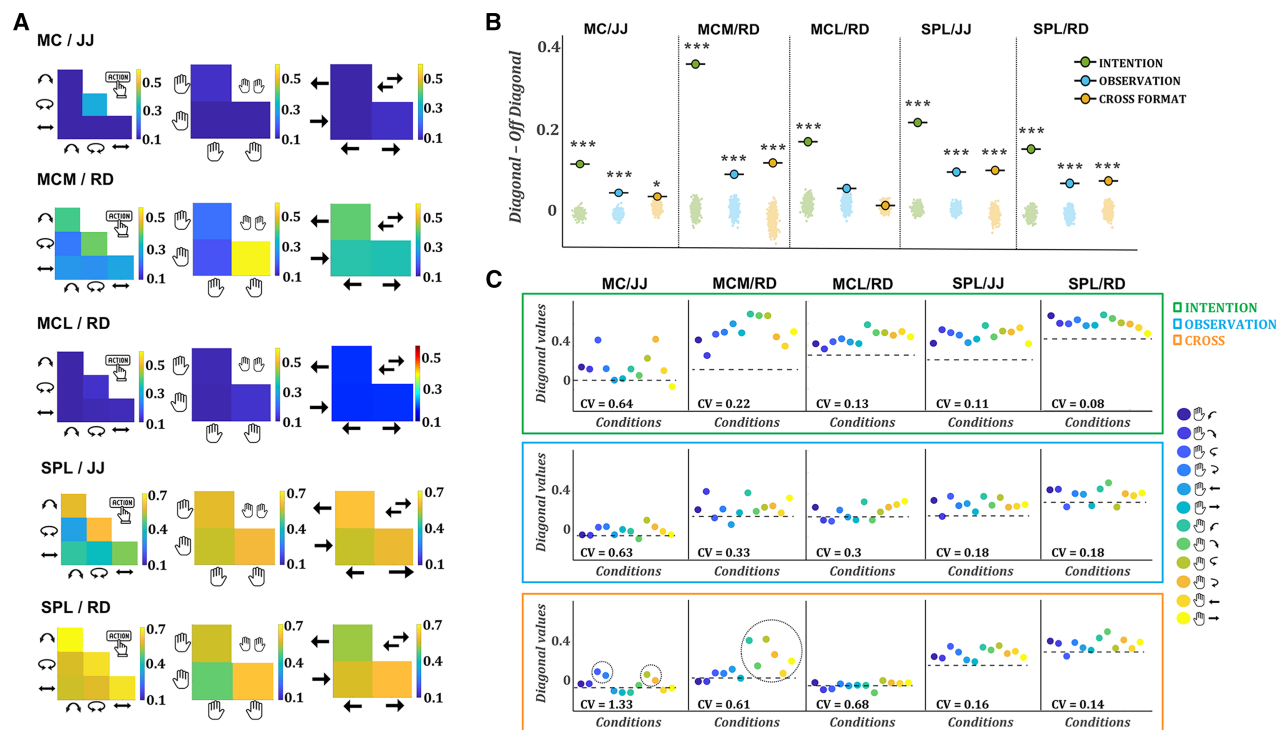


Figure 3. RSA within and between formats

(A) Cross-format RSA heatmaps for each array, computed separately for action (left column), hand (middle), and direction (right). Matrices show correlations between condition-specific population vectors from intention and observation.

(B) Permutation tests comparing diagonal/off-diagonal differences to shuffled null distributions. Bold markers indicate real values; transparent dots indicate permutations. Colors denote format (green: intention; blue: observation; orange: cross). Asterisks mark significant deviations from null ($p < 0.05$; $***p < 0.001$).

(C) Diagonal values of the 12-condition RSA matrices for each array and format (top: intention; middle: observation; bottom: cross-format). Each dot represents one of the 12 conditions; circled dots indicate conditions with selectively elevated cross-format similarity; dashed lines indicate the average off-diagonal value.

correlations ($R^2 < 0.13$). These patterns reflect selective generalization in MC/JJ and MCM/RD, aligned with their dominant intention-driven features (rotation and right-hand tuning) (Figure S3B), despite minimal observation tuning at the unit level. In contrast, SPL showed robust cross-format similarity in both participants, most prominently for action type, which exhibited the clearest diagonal structure. Diagonal correlations were high for all actions (JJ: 0.50–0.60; RD: 0.66–0.71), consistent with a stable shared representation across formats.

To evaluate RSA structure and quantify whether it exceeded noise levels, we compared the difference between mean diagonal and off-diagonal RSA values against a null distribution generated via 1,000 label permutations (Figure 3B). SPL showed significant structure across all formats in both participants (JJ: int = 0.22, obs = 0.11, cross = 0.11; RD: int = 0.15, obs = 0.07, cross = 0.07; all $p < 0.05$), as did MC/JJ and MCM/RD (MC: 0.12, 0.04, 0.04; MCM: 0.35, 0.08, 0.12; all $p < 0.05$). MCL/RD reached significance only during intention (0.16, $p < 0.05$); observation and cross-format effects were nonsignificant.

We also quantified variability across conditions using the coefficient of variation (CV) (Figure 3C). During intention, diagonal values exceeded the off-diagonal mean across all arrays; this was strongest and most uniform in SPL and RD's motor arrays (CVs: SPL/JJ = 0.11, SPL/RD = 0.08, MCL/RD = 0.13, MCM/

RD = 0.22), while MC/JJ showed weaker, more variable structure (CV = 0.64), driven largely by rotation. During observation, the same pattern was preserved but weaker. In the cross-format analysis, MC/JJ and MCM/RD showed selective structure with high condition variability (CV = 1.33 and 0.61), MCL/RD showed none, and SPL exhibited consistently elevated diagonal values with low variability (CV = 0.16 and 0.14).

Together, these results demonstrate that SPL consistently encoded task structure across formats at the population level, most prominently for action type. In contrast, MC showed cross-format similarity primarily for features that are strongly represented during intention, resulting in more condition-specific cross-format structure.

SPL supports robust cross-format decoding, while MC shows asymmetric or absent generalization

We assessed how reliably task features could be decoded from population activity over time using time-resolved decoding analyses for action type, effector, and movement direction, separately for intention and observation (Figure 4A; individual arrays in Figures S6A, S6B, S7A, and S7B). During intention, both MC and SPL showed robust decoding of action type (peak: MC 93.7%, SPL 98.6%) and effector (peak: MC 97.2%, SPL 95.2%); direction decoding remained weak. During observation,

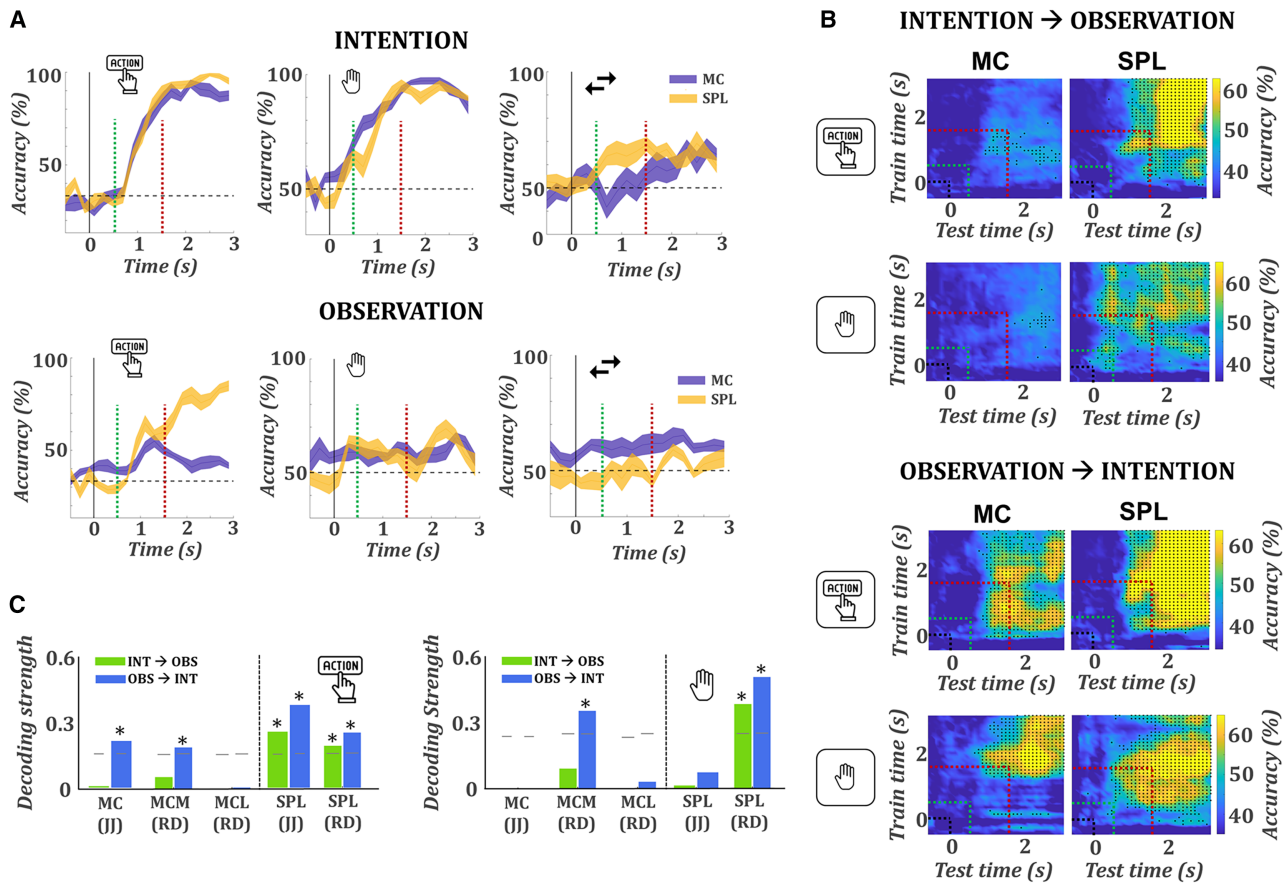


Figure 4. Decoding task variables within and across formats

(A) Time-resolved decoding accuracy for action (left), hand (middle), and direction (right). Top: intention; bottom: observation. Results are shown for pooled MC arrays (purple) and pooled SPL arrays (orange). Vertical lines indicate the hand cue (black), action cue (green), and go cue (red). Shaded regions denote SEM across cross-validations; the horizontal dashed line indicates chance.

(B) Cross-format decoding matrices. Classifiers were trained in one format and tested in the other. Top: intention → observation; bottom: observation → intention. Heatmaps show decoding accuracy across training (y axis) and testing (x axis) time bins for action (top row) and hand (bottom row). Dashed lines mark task events (black: hand cue, green: action cue, red: go cue); dots indicate significant decoding (permutation test, $p < 0.05$).

(C) Cross-format decoding strength per array for action (left) and hand (right). Bars show intention → observation (green) and observation → intention (blue). Horizontal dashed lines indicate chance levels computed from permutation tests, and asterisks denote significant differences from chance ($p < 0.05$).

SPL maintained robust action decoding (peak: 84.9%), whereas MC showed only modest above-chance decoding (peak: 54.9%). Effector decoding was modest in both regions (MC: 65.4%, SPL: 69.3%), and direction remained at chance. The same pattern was observed at the individual array level. Across sessions (Figures S7C and S7D), decoding performance was stable: action and hand decoding remained consistently above chance during intention in all regions, with SPL showing the strongest action decoding during observation, whereas direction decoding remained uniformly poor.

To assess generalization across formats, we trained classifiers on one format and tested them on the other across all time-point pairs, running analyses bidirectionally (intention → observation and observation → intention) (Figures 4B, S6C, and S7D; direction results in Figure S8A; for thresholding via permutation testing, see STAR Methods).

SPL showed robust bidirectional cross-decoding, strongest for action. Action decoding peaked at 81.9% (intention → observation) and 93.0% (observation → intention), with similarly high hand decoding (77.8% and 84.0%). In contrast, MC exhibited weaker, asymmetric generalization. Action decoding was lower for intention → observation than for observation → intention (48.6% vs. 63.9%), and hand decoding showed the same pattern (62.5% vs. 85.4%).

At the array level (Figures S6C and S6D), SPL showed bidirectional generalization in both participants. In MC, asymmetry was evident in MC/JJ and MCM/RD but not MCL/RD. In MC/JJ, cross-decoding was primarily observed for action in the observation → intention direction (62.5%). In MCM/RD, generalization was strongest for hand and again favored the observation → intention direction (88.2% vs. 72.2%), whereas MCL showed no significant effects.

To summarize decoding across the full train $\times$ test time space, we quantified cross-format decoding strength within the task window (0–2.5 s) as the product of mean significant accuracy and the fraction of significant time-bin pairs (Figure 4C). Significance was assessed against permutation-derived null distributions. SPL showed the strongest cross-format strength for action in both participants and directions (RD: 0.20/0.20; JJ: 0.27/0.39, $p < 0.05$) and robust bidirectional decoding for hand in RD (0.20/0.26, $p < 0.05$). In contrast, MC exhibited weaker and asymmetric generalization. In MC/JJ, action decoding was largely unidirectional (0.01/0.23, $p < 0.05$), and in MCM/RD, both hand and action showed asymmetry (hand: 0.05/0.20, $p < 0.05$; action: 0.06/0.20, $p < 0.05$). No measurable decoding strength was observed in MCL.

Cross-format decoding in RD was qualitatively unchanged after separating SUA and MUA signals and excluding putative cross-session matches (Figures S8B and S8C), confirming that the observed asymmetries were not driven by signal type or unit re-identification.

To test whether cross-format generalization was driven by shared visual input, participant JJ completed three additional sessions in which intention trials were performed while viewing only a static hand image rather than the movement video (MC: 270 units; SPL: 230 units) (Figure S9; see STAR Methods). Action decoding during observation remained robust only in SPL, and MC again showed asymmetric cross-format decoding restricted to the observation $\rightarrow$ intention direction, with SPL showing significant but weaker bidirectional generalization. These results suggest that cross-format effects cannot be fully explained by shared visual input.

The unidirectional decoding observed in MC/JJ and MCM/RD is particularly notable given the weak or absent within-format decoding during observation in these regions. Despite limited separability within observation, classifiers trained on observation generalized above chance to intention. Thus, observation-related population activity retained condition-specific structure aligned with intention representations, even when within-format decoding was weak. In SPL, the robust and bidirectional decoding indicates the presence of a stable, format-invariant code, particularly for action type, consistent with its generalized encoding across tasks and formats.

SPL and MC exhibit distinct representational geometries across action formats

A central question in understanding action encoding is whether population activity occupies a similar geometric structure across formats. We examined the geometric organization of neural trajectories by projecting trial-averaged responses into a shared principal-component analysis (PCA) space (Figures 5A, S10A, and S10B; variance explained in Table S2). Explained variance exceeded 60% in all cases, with observation variance consistently lower than intention in MC but comparable across formats in SPL.

In MC/JJ, trajectories appeared similar across formats only for the rotation action, consistent with its selective generalization in RSA (Figure 3A). In MCM/RD, responses were more similar for right-hand conditions, in line with effector-specific cross-format structure (Figure 3A). MCL/RD showed clearly segregated tra-

jectories across formats for all conditions, matching the absence of generalization observed throughout prior analyses (Figures S6C and S6D). In contrast, SPL exhibited qualitatively similar trajectories for all actions across formats, consistent with its robust and generalized encoding.

To quantify cross-format similarity, we applied Procrustes analysis between observation and intention trajectories for each task variable and condition (Figures 5B, 5C, and S10C), which aligns trajectories via translation, rotation, and scaling and returns a residual dissimilarity distance. SPL showed low alignment distances across all conditions ($d < 0.1$), indicating consistent geometric overlap. In MC/JJ, low distance was found only for rotation ($d = 0.25$), with poor alignment for lift ($d = 0.89$) and slide ($d = 0.64$). In MCM/RD, right-hand trajectories aligned more closely ($d = 0.13$) than left-hand ones ($d = 0.30$). No condition in MCL/RD yielded meaningful alignment. When pooling distances across arrays, alignment was significantly stronger in SPL than in MC across all task dimensions. For action, the mean distance was 0.45 in MC and 0.07 in SPL (Wilcoxon rank-sum, $p < 0.05$). For hand, the mean distance was 0.34 in MC and 0.07 in SPL ($p < 0.05$). For direction, the mean distance was 0.40 in MC and 0.08 in SPL ($p < 0.05$). These results confirm that SPL supports a shared representational geometry across formats, while MC/JJ and MCM/RD exhibit selective overlap, and MCL/RD none. Although Procrustes alignment does not imply that observation and intention occupy the same neural space, it tests whether the relative configuration of conditions is preserved after removing global differences in translation, rotation, and scale. Successful alignment therefore indicates a shared representational structure that can be linearly mapped across formats.

To further assess population structure across formats, we applied uniform manifold approximation and projection (UMAP) to embed single-trial activity from both intention and observation into a shared low-dimensional space (Figures 5D and S11). Three representative embeddings (MCM/RD, MCL/RD, and SPL/JJ) revealed distinct regimes (Figure 5D): in MCM/RD, right-hand trials formed overlapping clusters across formats while left-hand trials remained separated; in MCL/RD, intention and observation occupied distinct non-overlapping regions; in SPL/JJ, strong cross-format overlap was evident for all action types, with well-separated but aligned clusters, particularly for rotation. These embeddings reinforced the dissociation across regions: SPL in both participants supported robust, shared population codes across intention and observation; MCM/RD and MC/JJ exhibited conditional overlap; and MCL/RD remained format specific.

High-gamma LFP responses reveal format-consistent population encoding at the single-channel level

To evaluate neural population structure beyond spiking activity, we analyzed high-gamma (60–120 Hz) LFP power as a complementary population-level metric (Figure 6). Across all arrays and participants, all channels showed significant modulation for at least one condition in both intention and observation. Strikingly, all MC arrays exhibited robust observation-related responses, in contrast to the sparse or absent observation tuning in the SUA/MUA data. Figure 6A shows example power spectra from JJ,

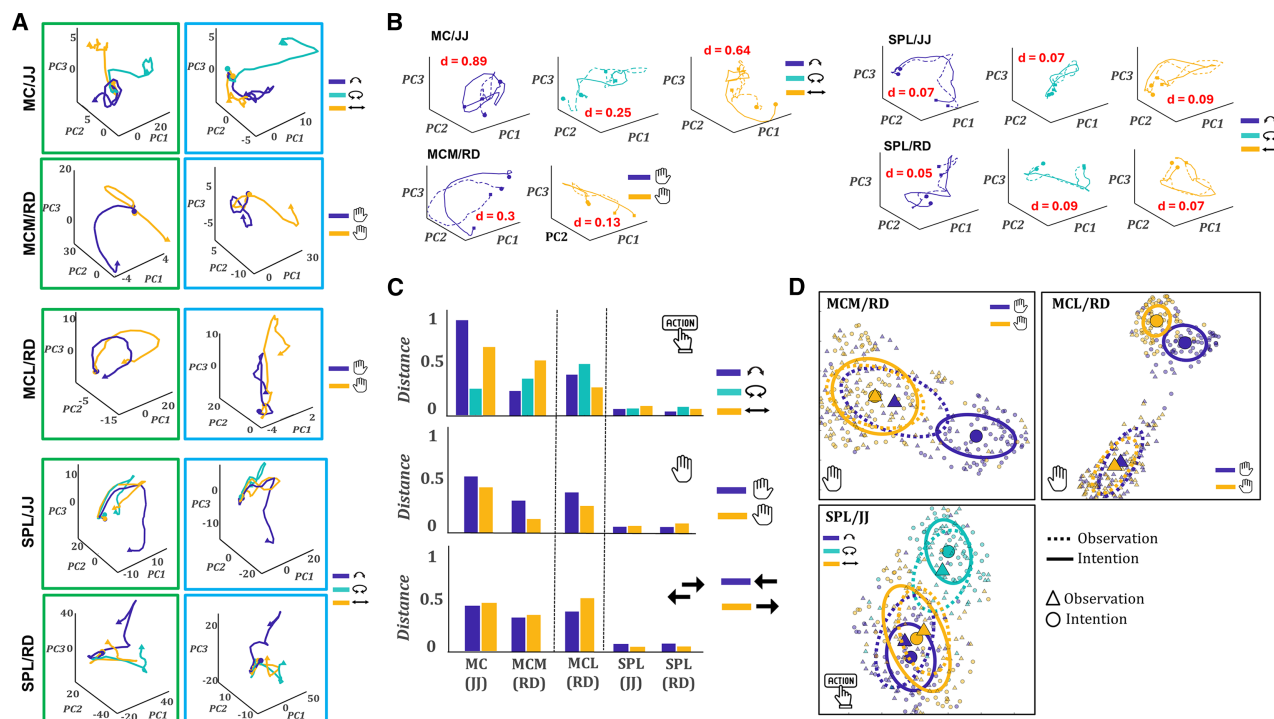


Figure 5. Geometric organization of population activity across formats

(A) Example PCA trajectories for intention (left column, green outline) and observation (right column, blue outline), plotted in the first three principal components for action (MC/JJ, SPL/JJ, and SPL/RD) and hand (MCL/RD and MCM/RD). (B) Procrustes-aligned trajectories from the same examples as in (A), with intention shown as solid lines and observation as dashed lines. Alignment distance (d) quantifies dissimilarity after transformation. (C) Summary of alignment distances across all arrays and task variables (action, hand, and direction). (D) UMAP embeddings illustrating three representative cases of cross-format geometry. Triangles represent observation trials, and circles represent intention trials. Colors denote task conditions, and ellipses enclose clusters (solid: intention, dashed: observation).

illustrating increased high-gamma activity in MC during intention specifically for rotation actions, with a similar but more subtle effect during observation. In SPL, strong responses were evident for all actions in both formats. Figure 6B quantifies the distribution of significantly modulated channels across conditions, highlighting that format-specific tuning was preserved: rotation tuning in JJ MC and effector tuning in MCM/MCL of RD appeared in both intention and observation. SPL again exhibited widespread modulation across all conditions in both participants. Figure 6C shows the evolution of tuning over time. In MC, tuning profiles were nearly identical between intention and observation: action-selectivity in JJ's MC and hand-selectivity in MCM and MCL of RD. SPL showed consistent tuning for action during observation in both participants; during intention, tuning was action-dominant in JJ and hand-dominant in RD. These findings indicate that high-gamma activity captures structured population-level encoding even when single-unit selectivity appears sparse. This pattern does not imply a distinct representational format in the LFPs but instead reflects differences in sampling. Spike recordings capture a limited subset of neurons, whereas high-gamma signals pool activity over a broader local population. Consistent with this, spike-based population analyses already revealed latent representational geometry despite weak single-unit tuning, which high-gamma activity expressed

more robustly. In SPL, the convergence of SUA, LFP, and population metrics supports its role as a stable hub for shared-action representations across intention and observation.

SPL flexibly represents observed actions only when they are behaviorally relevant

To determine whether instructed and observed actions are encoded in parallel or selectively modulated by task demands, we designed two dissociation tasks that explicitly separated intention from observation (Figure 7). On each trial, participant RD received instruction cues specifying the intended hand and action, followed by a video showing either a congruent or incongruent action. Although both the instruction and video were present on every trial, task demands determined which source of information was behaviorally relevant. This design allowed us to dissociate intention-related and observation-related activity and test how behavioral relevance shapes neural representations across regions. Session-level unit counts are reported in Tables S3 and S4.

When only the instruction was relevant (no-probe intention block)

In the no-probe intention block, RD intended the instructed action while watching an action video that was behaviorally

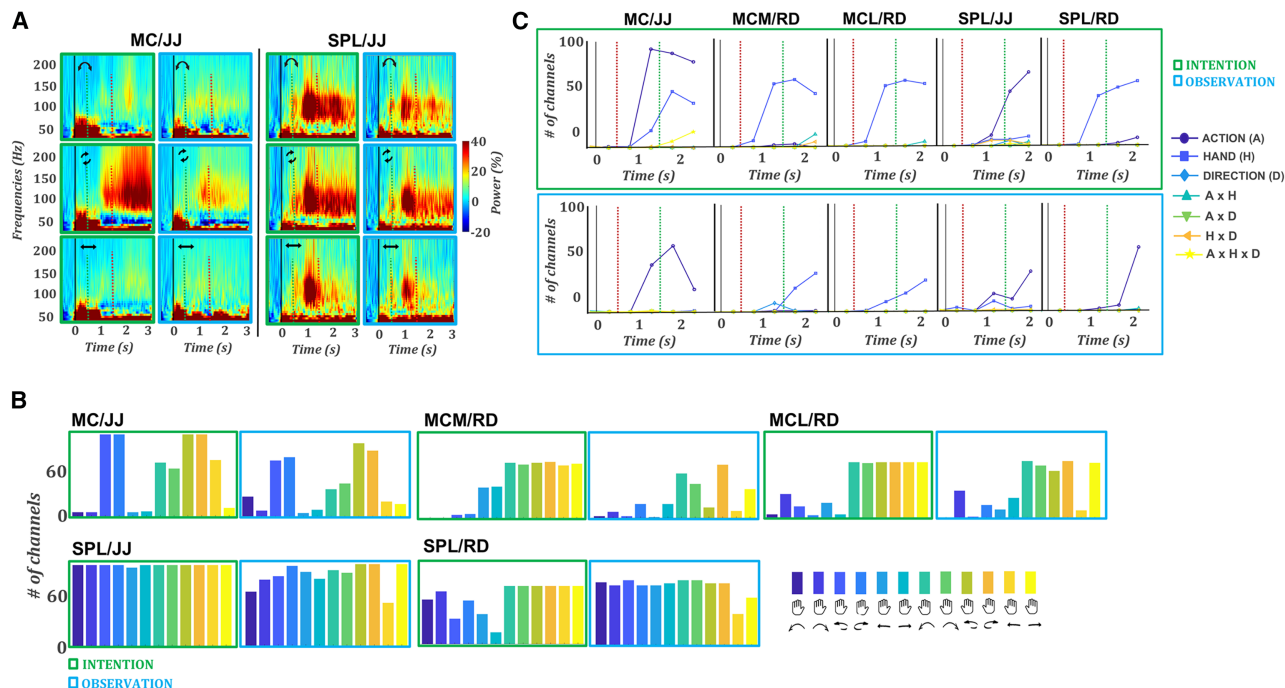


Figure 6. High-frequency LFP activity across formats and cortical regions

(A) Example spectrograms showing event-related changes in LFP power for action type averaged across all channels per array (MC/JJ and SPL/JJ). Colors indicate percentage power change relative to baseline. Within each array, the left column (green outline) shows intention, and the right column (blue outline) shows observation. Vertical lines indicate the hand cue (black), action cue (green), and go cue (red).
(B) Number of channels with significant power increases. Green and blue outlines correspond to intention and observation, respectively.
(C) Three-way ANOVA tuning profiles over time (factors: action, hand, and direction). Top: intention (green outline); bottom: observation (blue outline). Lines indicate the number of channels significantly tuned to each main effect and interaction. Vertical lines indicate the hand cue (black), action cue (green), and go cue (red).

irrelevant, as it could mismatch the intended action. [Figure 7A](#) (left panel) illustrates the task design and decoding time courses for both the instructed and video actions, with green marking behaviorally relevant epochs and red marking irrelevant epochs. Decoding followed this relevance structure: the instructed action reached near-ceiling levels across regions (MCM: 99.4%; MCL: 97.5%; SPL: 91.9%), whereas decoding of the concurrent video action remained at chance (<37% in all areas). Consistent with this, tuning analyses ([Figure S12](#), green outline) showed robust encoding of the instructed movement, with strong hand tuning in MC (MCM: 163 units at 1.75 s; MCL: 90 units at 1.25 s) and action tuning in SPL (60 units at 2.25 s). No significant tuning to the video action was observed in any region.

When only the video was relevant (no-probe observation block)

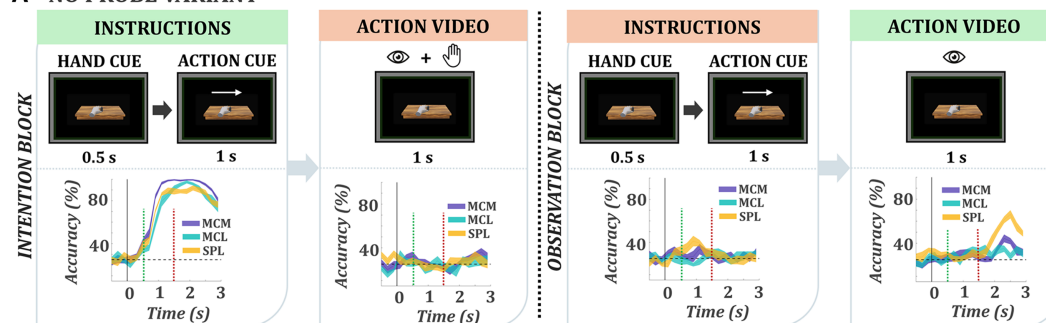
In the no-probe observation block, RD received the action instruction but passively viewed the video without intending the instructed movement. In this context, the instruction cue was behaviorally irrelevant, whereas the video action was the only relevant source of information ([Figure 7A](#), right panel). Consistent with this, decoding of the instructed action remained near chance across all arrays (<42%). In contrast, decoding of the video action increased to 66.3% at 2.5 s in SPL but remained weak in MC (<45%). Tuning analyses showed the same pattern

([Figure S12](#), blue outline): SPL selectively encoded the video action (23 tuned units at 2.25 s), predominantly for action identity rather than hand or interaction terms, while tuning to the instructed action returned to baseline across regions. Thus, in the no-probe task, SPL selectively encoded the behaviorally relevant source of information, whereas MC consistently represented the instructed movement and did not encode the action video.

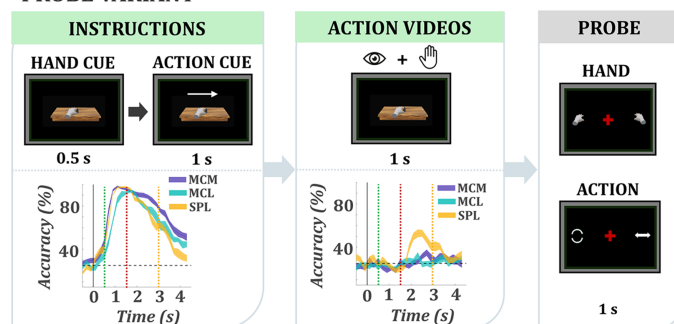
When both sources were relevant (probe variant)

In the probe variant, RD intended the instructed movement while viewing the action video, as in the no-probe intention block. However, after movement completion, he was required to report a feature of the video via a saccade ([Figure 7B](#)), rendering both the instruction and the video behaviorally relevant within the same trial. Decoding of the instructed action remained high across all regions (MCM: 97.9%; MCL: 94.2%; SPL: 97.5%) ([Figure 7C](#)). Crucially, SPL now also showed reliable decoding of the video action (52.9% at 2.3 s), whereas decoding in MC remained at chance. Tuning analyses showed the same structure ([Figure S12](#), gray outline). Tuning to the instructed action remained robust in MC (MCM: 122 hand-tuned units at 2.25 s; MCL: 74 at 1.25 s) and SPL (96 action-tuned units at 1.25 s). SPL also showed video-related tuning (29 units at 2.25 s), whereas motor arrays exhibited minimal video-related tuning

A NO PROBE VARIANT



B PROBE VARIANT



C

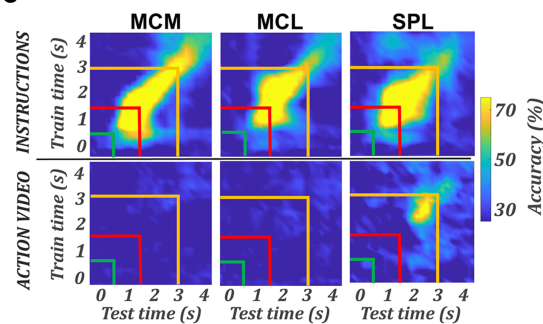


Figure 7. Dissociation tasks: Neural encoding for instructed and observed actions

Dissociation task: green and red denote the task-relevant and task-irrelevant sources of information within each block.

(A) No-probe variant. Trial structure for the intention and observation blocks. In the intention block (left), RD executed the instructed action while ignoring the video. In the observation block (right), RD passively viewed the video while ignoring the instruction. Decoding plots show time-resolved accuracy for each block. Colored lines indicate decoding performance for MCM (purple), MCL (teal), and SPL (orange); shaded regions denote $\pm$ SEM across cross-validation folds. The dashed horizontal line marks the chance level (25%; 2 actions $\times$ 2 hands). Vertical lines indicate task events: hand cue (black), action cue (green), and video/movement onset (red).

(B) Probe variant. Trial structure, including the post-trial probe requiring a saccadic report of either the video hand or video action. Decoding plots show instructed action decoding (left) and video action decoding (right). The dashed horizontal line marks the chance level (25%; 2 actions $\times$ 2 hands). Vertical lines indicate the hand cue (black), action cue (green), video onset (red), and probe onset (orange).

(C) Heatmaps of LDA decoding accuracy in the probe variant across all train $\times$ test time windows for a given brain area (MCM, MCL, and SPL). Top row: decoders trained on the instructed action; bottom row: decoders trained on the video action. Color indicates mean decoding accuracy across folds. The diagonal represents within-time decoding, whereas off-diagonal values reflect temporal generalization.

(<5 units). Thus, behavioral relevance reshaped representations in SPL but not in MC: SPL concurrently encoded both intended and observed actions, whereas MC maintained intention-specific representations.

Control analyses: Congruency, temporal stability, and full-condition structure

Repeating all analyses on incongruent trials only yielded nearly identical decoding performance (Figure S13A), and neither tuning nor decoding revealed significant selectivity for congruency itself (Figure S13B), confirming that population activity reflected genuine task-related representations. The qualitative regional dissociation was preserved when decoding was repeated separately for SUA and MUA (Figure S14), ruling out signal composition as a driver of the gating effect.

We next examined the temporal stability of instructed and video action representations within the probe task using cross-time decoding (Figure 7C). Instructed actions showed stable decoding across all regions, with SPL exhibiting a broader diagonal during the Go epoch, indicating more sustained representations.

In contrast, video action decoding was transient and restricted to SPL during video presentation, while MC showed no reliable decoding.

Finally, classifiers trained to decode all 16 trial types (Figure S15) revealed clustering by instructed action in the no-probe intention block, weak structure in the observation block, and full trial-type identity representation in SPL, but not MC, during the probe task. Notably, this parallel encoding emerged despite minimal differences from the no-probe intention task.

Together, these results reveal a context-dependent gating mechanism in SPL: while MC consistently encoded the instructed movement, SPL represented observed actions only when behaviorally relevant, despite identical visual input across task variants.

DISCUSSION

We conducted a systematic comparison of intended and observed actions using single-unit recordings in the human MC and SPL. In MC, intended actions were robustly encoded,

but passive observation failed to elicit mirror-like responses at the single-unit level. Still, population analyses revealed a weak but structured geometry during observation, partially aligned with the execution format and centered on the most strongly tuned features. In SPL, observed actions were reliably encoded, with most selective units maintaining shared tuning for action type across formats. These representations generalized across the population and were sensitive to behavioral relevance: when observed actions were not task-relevant, their encoding was suppressed. Our findings suggest that action observation engages distinct encoding schemes across cortical regions, reflecting a flexible, context-dependent system rather than a fixed mirror mechanism.

In MC, the absence of mirror-like tuning at the single-unit level during passive observation contrasts with prior reports in nonhuman primates, where neurons in M1 and premotor areas exhibit consistent observation-driven responses.^{31–33} However, population-level analyses revealed a more nuanced structure: neural activity projected into a low-dimensional space revealed partially aligned trajectories between intention and observation, successful cross-decoding from observation to intention, and localized overlap in UMAP space, specifically for intention-tuned features (Figures 3, 4, and 5). These results suggest that while single units do not overtly mirror observed actions, MC maintains a latent representational geometry during observation that reflects core aspects of the intention structure.

Jiang et al. (2020)³⁴ reported that executed and observed movements in monkey M1 and PMd occupy a shared subspace, with closely aligned population trajectories. In our recordings from human MC, population overlap was more limited, confined to specific subregions and features strongly represented during intention. We also did not observe strong single-unit responses during passive observation, in contrast to reports in monkeys,³¹ suggesting that human MC may contribute less directly to action observation. One potential difference is that, in monkeys, observed and executed actions share closely matched kinematics, whereas our participants could not reproduce the observed movements, potentially reducing intention-observation alignment. However, the absence of overt motor output alone does not preclude shared population geometry: Dekleva et al. (2024)³⁵ demonstrated that motor imagery retains execution-like population dynamics in human MC despite differences in behavioral output. At the same time, Rastogi et al. (2020)³⁰ showed that activity in human MC was strongly shaped by volitional state, with observation responses substantially weaker than those during attempted movement. Our findings lie between these regimes. Population analyses revealed a partial and feature-specific overlap between intention and observation, suggesting that passive observation engages a restricted component of the intention-dominant space in MC rather than a fully shared intention-observation subspace.

Our population-level findings are further supported by high-gamma LFP signals, which revealed robust and condition-congruent responses in both areas during observation, fulfilling the classical criteria for mirroring (Figure 6) and closely paralleling human fMRI results that consistently report motor cortical activation during passive action viewing.³⁶ In MC, LFP re-

sponses contrasted with the sparse observation tuning at the single-unit level. Shared structure became more evident with broader spatial sampling: MUA in RD showed greater intention-observation overlap than well-isolated SUA, and shared representations were likewise more prominent in JJ's MC, which was dominated by MUA. These findings suggest that observation-related structure in MC is distributed across local circuits and is only partially captured by sparsely sampled SUA. High-gamma LFP and fMRI further emphasize this distributed activity, as BOLD signals correlate more strongly with LFP signals than with MUA and primarily reflect synaptic input and intracortical processing rather than isolated spiking output.³⁷ Together, these results support a graded organization in which intention-observation overlap is embedded in MC despite limited mirror-like tuning at the single-unit level.

In contrast to MC, SPL exhibited robust encoding of observed actions, with many selective units showing shared tuning across execution and observation (Figure 2E). This cross-format structure was strongest for action type, consistent with a higher-order representation of action identity. Similar organization was evident at the population level: neural trajectories aligned across formats, cross-decoding generalized bidirectionally for action type, and UMAP embeddings showed overlapping population structure (Figures 3, 4, and 5). These results suggest that SPL encodes a conceptual representation of action, generalizable across sensory formats and resilient to contextual mismatch (e.g., when participants do not precisely reproduce the observed movement), which is not tied to specific motor output. Our findings converge with a growing body of work suggesting that PPC supports abstract action encoding. fMRI studies have reported overlapping representations of executed and observed hand actions in PPC, comparable to those observed in premotor regions.³⁸ Aflalo et al. (2020)²⁸ demonstrated shared population codes in human PPC for action verbs and visual action stimuli. Lanzilotto et al. (2020)³⁹ showed that AIP neurons do not encode actions invariantly, but population activity allows reliable decoding across viewpoints and formats. Similarly, Chivukula et al. (2025)²⁵ showed that somatosensory representations in PPC generalize across both experienced and observed touch. Across these studies, as in ours, PPC appears to encode not the physical parameters of an event but the concept of the action itself. Our results extend this literature by showing that such conceptual generalization is supported by both single-unit tuning and aligned population geometry.

The dissociation paradigms allowed us to determine whether these abstract action representations in SPL are fixed or dynamically modulated. In the no-probe variant, SPL encoded only the instructed action, with no detectable representation of the concurrent video stream (Figure 7A). This dissociation confirms that the motor responses observed in our main task were not visually driven, ruling out the possibility that SPL activity merely reflects responses to visual input. In the probe variant (identical in sensory input but requiring participants to recall the video action on every trial), SPL encoded both the instructed and video actions (Figure 7B), whereas MC encoded only the instructed action in both task variants. We show that observed features are not represented by default; their encoding emerges only when they are task-relevant. When motor and visual goals conflict,

internally generated actions suppress irrelevant sensory input. The dissociation between motor output and visual input demonstrates that observed actions are not mirrored automatically but are flexibly gated by task demands.

Such task-dependent gating aligns with prior findings in nonhuman primates and human fMRI studies showing that sensory responses in PPC are modulated by cognitive context and behavioral relevance.⁴⁰ Classic attentional effects in parietal cortex often involve task-dependent shifts in single-unit selectivity, while unattended features remain decodable at the population level, reflecting changes in discriminability rather than representational structure.⁴¹ In contrast, we observed near-binary restructuring: decoding of action identity approached chance when behaviorally irrelevant and was robust when relevant. This extends prior work by demonstrating task-dependent gating at both the single-unit and population levels in human SPL. The role of SPL in this process is consistent with its proposed function within the frontoparietal multiple-demand (MD) network, a domain-general system implicated in cognitive control, goal-directed behavior, and adaptive task management.^{42,43} Rather than serving as a static relay of sensorimotor signals, the parietal cortex is increasingly viewed as a dynamic integrator whose representational structure is shaped by task goals. Models of attentional control further support this view, proposing that the dorsal frontoparietal network (including the intraparietal sulcus and superior parietal regions) allocates top-down attention and formulates predictions about incoming stimuli, while the ventral attention network centered on the temporoparietal junction and ventral frontal cortex detects salient or unexpected events and redirects attention accordingly.⁴⁴ These dynamics are closely related to the principle of mixed selectivity, whereby neurons encode combinations of task-relevant features across domains, a mechanism proposed to enable flexible and high-capacity representation in cognitive systems.^{14–16}

Our findings suggest that action perception does not rely on automatic mirroring but on a hierarchical organization in which MC encodes intention-dominant signals and posterior parietal cortex flexibly integrates observed input with internal goals. This geometry-based dissociation points to context-sensitive transformations rather than reflexive resonance as the basis for linking perception to action. Future studies combining broader cortical sampling with intact motor systems will determine whether the geometric principles identified here reflect local properties or a general cortical strategy. In either case, our results highlight a framework in which human action representations emerge as flexible, context-dependent geometries embedded within the broader architecture of cognitive control.

Limitations of the study

This study is based on intracortical recordings from two individuals with tetraplegia and limited spatial sampling of motor and posterior parietal cortex. Utah arrays cover a small cortical surface area, restricting measurements to localized neuronal populations and precluding assessment of broader network interactions across frontoparietal regions. Although findings were consistent across participants and analytical approaches, the limited number of individuals constrains generalization to the wider population. In addition, execution-related activity reflects

attempted rather than overt movement, which may alter sensorimotor dynamics relative to intact motor systems. The main task also included overlapping visual input between intention and observation conditions, raising the possibility that some cross-format effects may partially reflect shared sensory drive. Although additional control experiments without movement observation during intention reproduced the main cross-decoding asymmetries, these analyses were based on fewer sessions and units and were conducted in only one participant. Conclusions are also limited to the discrete set of actions and controlled laboratory conditions tested here.

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources should be directed to and will be fulfilled by the lead contact, Richard Andersen (richard.andersen@vis.caltech.edu).

Materials availability

This study did not generate any new, unique reagents.

Data and code availability

The datasets and analysis code required to reproduce the main figures reported in this paper are publicly available on CaltechDATA at <https://doi.org/10.22002/6bhj2-vy860>.

ACKNOWLEDGMENTS

We would like to express our gratitude to participants JJ and RD for making this research possible. This work was supported by the National Eye Institute (grant UG1EY032039), the Tianqiao and Chrissy Chen Brain-Machine Interface Center at Caltech, the Swartz Foundation, and the Boswell Foundation. V.B. was supported by Fonds Wetenschappelijk Onderzoek (FWO) grant 1264926N.

AUTHOR CONTRIBUTIONS

V.B. and R.A.A. conceived and designed the experiment. V.B. performed the recordings and data analysis and wrote the manuscript. J.G. and R.A.A. supervised and guided the study. E.R.R., A.B., and C.L. provided clinical support. V.B., J.G., R.A.A., and K.P. acquired funding. K.P. and R.A.A. were responsible for project administration. All authors reviewed and edited the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

DECLARATION OF GENERATIVE AI AND AI-ASSISTED TECHNOLOGIES IN THE WRITING PROCESS

During the preparation of this work, the authors used ChatGPT (OpenAI) to assist with language editing and text refinement. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
 - Participants
- METHOD DETAILS
 - Electrophysiological Recordings

- Experimental setup
- Action Intention and Observation Task
- Control Experiment Without Observed Movement During Intention
- Dissociation Task Variants
- Dissociation: No-Probe Task
- Dissociation: Probe Task
- **QUANTIFICATION AND STATISTICAL ANALYSIS**
 - Spatial Distribution of Gaze
 - Single Unit Activity (SUA) and Multiunit Activity (MUA)
 - Local Field Potential (LFP)
 - Cross-session unit matching analysis
 - Data Aggregation
 - Statistical Comparison to Baseline
 - Facilitated and Suppressed Responses
 - Latency
 - Tuning Analysis
 - Overlap of Task-Relevant Units Across Formats
 - Linear Model Analysis
 - Population Analysis
 - Linear Discriminant Analysis (LDA) Within Format
 - Cross – Format Decoding
 - Cross – Time Decoding
 - PCA Trajectory Analysis
 - Procrustes Analysis
 - UMAP Visualization of Neural Representations
- **ADDITIONAL RESOURCES**

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.cell.2026.07.032>.

Received: November 21, 2025

Revised: May 23, 2026

Accepted: July 22, 2026

Published: August 11, 2026

REFERENCES

1. di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., and Rizzolatti, G. (1992). Understanding motor events: a neurophysiological study. *Exp. Brain Res.* 91, 176–180. <https://doi.org/10.1007/BF00230027>.
2. Rizzolatti, G., Fadiga, L., Gallese, V., and Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Brain Res. Cogn. Brain Res.* 3, 131–141. [https://doi.org/10.1016/0926-6410\(95\)00038-0](https://doi.org/10.1016/0926-6410(95)00038-0).
3. Gallese, V., Fadiga, L., Fogassi, L., and Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain* 119, 593–609. <https://doi.org/10.1093/brain/119.2.593>.
4. Hickok, G. (2009). Eight problems for the mirror neuron theory of action understanding in monkeys and humans. *J. Cogn. Neurosci.* 21, 1229–1243. <https://doi.org/10.1162/jocn.2009.21189>.
5. Bonini, L., and Ferrari, P.F. (2011). Evolution of mirror systems: A simple mechanism for complex cognitive functions. *Ann. N. Y. Acad. Sci.* 1225, 166–175. <https://doi.org/10.1111/j.1749-6632.2011.06002.x>.
6. Rizzolatti, G., and Sinigaglia, C. (2016). The mirror mechanism: A basic principle of brain function. *Nat. Rev. Neurosci.* 17, 757–765. <https://doi.org/10.1038/nrn.2016.135>.
7. Bonini, L., Rotunno, C., Arcuri, E., and Gallese, V. (2022). Mirror neurons 30 years later: implications and applications. *Trends Cogn. Sci.* 26, 767–781. <https://doi.org/10.1016/j.tics.2022.06.003>.
8. Kraskov, A., Dancause, N., Quallio, M.M., Shepherd, S., and Lemon, R.N. (2009). Corticospinal Neurons in Macaque Ventral Premotor Cortex with Mirror Properties: A Potential Mechanism for Action Suppression? *Neuron* 64, 922–930. <https://doi.org/10.1016/j.neuron.2009.12.010>.
9. Caggiano, V., Fogassi, L., Rizzolatti, G., Thier, P., and Casile, A. (2009). Mirror neurons differentially encode the peripersonal and extrapersonal space of monkeys. *Science* 324, 403–406. <https://doi.org/10.1126/science.1166818>.
10. Maranesi, M., Livi, A., and Bonini, L. (2017). Spatial and viewpoint selectivity for others' observed actions in monkey ventral premotor mirror neurons. *Sci. Rep.* 7, 8231. <https://doi.org/10.1038/s41598-017-08956-1>.
11. Orban, G.A., Lanzilotto, M., and Bonini, L. (2021). From observed action identity to social affordances. *Trends Cogn. Sci.* 25, 493–505. <https://doi.org/10.1016/j.tics.2021.02.012>.
12. Lake, B.M., Ullman, T.D., Tenenbaum, J.B., and Gershman, S.J. (2017). Building machines that learn and think like people. *Behav. Brain Sci.* 40, e253. <https://doi.org/10.1017/S0140525X16001837>.
13. Schwettmann, S.E., Tenenbaum, J.B., and Kanwisher, N. (2019). Invariant representations of mass in the human brain. *eLife* 8, e46619. <https://doi.org/10.7554/eLife.46619>.
14. Fusi, S., Miller, E.K., and Rigotti, M. (2016). Why neurons mix: High dimensionality for higher cognition. *Curr. Opin. Neurobiol.* 37, 66–74. <https://doi.org/10.1016/j.conb.2016.01.010>.
15. Zhang, C.Y., Aflalo, T., Revechikis, B., Rosario, E.R., Ouellette, D., Pouratian, N., and Andersen, R.A. (2017). Partially Mixed Selectivity in Human Posterior Parietal Association Cortex. *Neuron* 95, 697–708.e4. <https://doi.org/10.1016/j.neuron.2017.06.040>.
16. Zhang, C.Y., Aflalo, T., Revechikis, B., Rosario, E., Ouellette, D., Pouratian, N., and Andersen, R.A. (2020). Preservation of partially mixed selectivity in human posterior parietal cortex across changes in task context. *eNeuro* 7, ENEURO.0222-19.2019. <https://doi.org/10.1523/ENEURO.0222-19.2019>.
17. Rigotti, M., Barak, O., Warden, M.R., Wang, X.J., Daw, N.D., Miller, E.K., and Fusi, S. (2013). The importance of mixed selectivity in complex cognitive tasks. *Nature* 497, 585–590. <https://doi.org/10.1038/nature12160>.
18. Maranesi, M., Lanzilotto, M., Arcuri, E., and Bonini, L. (2024). Mixed selectivity in monkey anterior intraparietal area during visual and motor processes. *Prog. Neurobiol.* 236, 102611. <https://doi.org/10.1016/j.pneurobio.2024.102611>.
19. Andersen, R.A., and Cui, H. (2009). Intention, Action Planning, and Decision Making in Parietal-Frontal Circuits. *Neuron* 63, 568–583. <https://doi.org/10.1016/j.neuron.2009.08.028>.
20. Snyder, L.H., Batista, A.P., and Andersen, R.A. (1997). Coding of intention in the posterior parietal cortex. *Nature* 386, 167–170. <https://doi.org/10.1038/386167a0>.
21. Andersen, R.A., and Buneo, C.A. (2002). Intentional maps in posterior parietal cortex. *Annu. Rev. Neurosci.* 25, 189–220. <https://doi.org/10.1146/annurev.neuro.25.112701.142922>.
22. Gallego, J.A., Perich, M.G., Miller, L.E., and Solla, S.A. (2017). Neural Manifolds for the Control of Movement. *Neuron* 94, 978–984. <https://doi.org/10.1016/j.neuron.2017.05.025>.
23. Fu, Z., Beam, D., Chung, J.M., Reed, C.M., Mamelak, A.N., Adolphs, R., and Rutishauser, U. (2022). The geometry of domain-general performance monitoring in the human medial frontal cortex. *Science* 376, eabm9922. <https://doi.org/10.1126/science.abm9922>.
24. Bernardi, S., Benna, M.K., Rigotti, M., Munuera, J.R., Fusi, S., and Salzmann, C.D. (2020). The Geometry of Abstraction in the Hippocampus and Prefrontal Cortex. *Cell* 183, 954–967.e21. <https://doi.org/10.1016/j.cell.2020.09.031>.
25. Chivukula, S., Aflalo, T., Zhang, C., Rosario, E.R., Bari, A., Pouratian, N., and Andersen, R.A. (2025). Population encoding of observed and actual somatosensations in the human posterior parietal cortex. *Proc. Natl. Acad. Sci. USA* 122, e2316012121. <https://doi.org/10.1073/pnas.2316012121>.
26. Zhao, Z., and Schieber, M.H. (2025). Progressively shifting patterns of co-modulation among premotor cortex neurons carry dynamically similar signals during action execution and observation. *eLife* 13, RP94165. <https://doi.org/10.7554/eLife.94165>.

27. Mukamel, R., Ekstrom, A.D., Kaplan, J., Iacoboni, M., and Fried, I. (2010). Single-neuron responses in humans during execution and observation of actions. *Curr. Biol.* 20, 750–756. <https://doi.org/10.1016/j.cub.2010.02.045>.
28. Aflalo, T., Zhang, C.Y., Rosario, E.R., Pouratian, N., Orban, G.A., and Andersen, R.A. (2020). A shared neural substrate for action verbs and observed actions in human posterior parietal cortex. *Sci. Adv.* 6, eabb3984. <https://doi.org/10.1126/sciadv.abb3984>.
29. Vargas-Irwin, C.E., Feldman, J.M., King, B., Simeral, J.D., Sorice, B.L., Oakley, E.M., Cash, S.S., Eskandar, E.N., Friehs, G.M., Hochberg, L.R., et al. (2018). Watch, Imagine, Attempt: Motor Cortex Single-Unit Activity Reveals Context-Dependent Movement Encoding in Humans With Tetraplegia. *Front. Hum. Neurosci.* 12, 450. <https://doi.org/10.3389/fnhum.2018.00450>.
30. Rastogi, A., Vargas-Irwin, C.E., Willett, F.R., Abreu, J., Crowder, D.C., Murphy, B.A., Memberg, W.D., Miller, J.P., Sweet, J.A., Walter, B.L., et al. (2020). Neural Representation of Observed, Imagined, and Attempted Grasping Force in Motor Cortex of Individuals with Chronic Tetraplegia. *Sci. Rep.* 10, 1429. <https://doi.org/10.1038/s41598-020-58097-1>.
31. Dushanova, J., and Donoghue, J. (2010). Neurons in primary motor cortex engaged during action observation. *Eur. J. Neurosci.* 31, 386–398. <https://doi.org/10.1111/j.1460-9568.2009.07067.x>.
32. Tkach, D., Reimer, J., and Hatsopoulos, N.G. (2007). Congruent activity during action and action observation in motor cortex. *J. Neurosci.* 27, 13241–13250. <https://doi.org/10.1523/JNEUROSCI.2895-07.2007>.
33. Vigneswaran, G., Philipp, R., Lemon, R.N., and Kraskov, A. (2013). M1 corticospinal mirror neurons and their role in movement suppression during action observation. *Curr. Biol.* 23, 236–243. <https://doi.org/10.1016/j.cub.2012.12.006>.
34. Jiang, X., Saggar, H., Ryu, S.I., Shenoy, K.V., and Kao, J.C. (2020). Structure in Neural Activity during Observed and Executed Movements Is Shared at the Neural Population Level, Not in Single Neurons. *Cell Rep.* 32, 108148. <https://doi.org/10.1016/j.celrep.2020.108148>.
35. Dekleva, B.M., Chowdhury, R.H., Batista, A.P., Chase, S.M., Yu, B.M., Bowering, M.L., and Collinger, J.L. (2024). Motor cortex retains and reorients neural dynamics during motor imagery. *Nat. Hum. Behav.* 8, 729–742. <https://doi.org/10.1038/s41562-023-01804-5>.
36. Molenberghs, P., Cunnington, R., and Mattingley, J.B. (2012). Brain regions with mirror properties: A meta-analysis of 125 human fMRI studies. *Neurosci. Biobehav. Rev.* 36, 341–349. <https://doi.org/10.1016/j.neurobiorev.2011.07.004>.
37. Logothetis, N.K., Pauls, J., Augath, M., Trinath, T., and Oeltermann, A. (2001). Neurophysiological investigation of the basis of the fMRI signal. *Nature* 412, 150–157. <https://doi.org/10.1038/35084005>.
38. Filimon, F., Rieth, C.A., Sereno, M.I., and Cottrell, G.W. (2015). Observed, Executed, and Imagined Action Representations can be Decoded From Ventral and Dorsal Areas. *Cereb. Cortex* 25, 3144–3158. <https://doi.org/10.1093/cercor/bhu110>.
39. Lanzilotto, M., Maranesi, M., Livi, A., Ferroni, C.G., Orban, G.A., and Bonini, L. (2020). Stable readout of observed actions from format-dependent activity of monkey's anterior intraparietal neurons. *Proc. Natl. Acad. Sci. USA* 117, 16596–16605. <https://doi.org/10.1073/pnas.2007018117>.
40. Freedman, D.J., and Ibos, G. (2018). An Integrative Framework for Sensory, Motor, and Cognitive Functions of the Posterior Parietal Cortex. *Neuron* 97, 1219–1234. <https://doi.org/10.1016/j.neuron.2018.01.044>.
41. Ibos, G., and Freedman, D.J. (2016). Interaction between Spatial and Feature Attention in Posterior Parietal Cortex. *Neuron* 91, 931–943. <https://doi.org/10.1016/j.neuron.2016.07.025>.
42. Duncan, J. (2013). The Structure of Cognition: Attentional Episodes in Mind and Brain. *Neuron* 80, 35–50. <https://doi.org/10.1016/j.neuron.2013.09.015>.
43. Erez, Y., and Duncan, J. (2015). Discrimination of visual categories based on behavioral relevance in widespread regions of frontoparietal cortex. *J. Neurosci.* 35, 12383–12393. <https://doi.org/10.1523/JNEUROSCI.1134-15.2015>.
44. Corbetta, M., and Shulman, G.L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nat. Rev. Neurosci.* 3, 201–215. <https://doi.org/10.1038/nrn755>.
45. Brainard, D.H. (1997). The Psychophysics Toolbox. *Spat. Vis.* 10, 433–436. <https://doi.org/10.1163/156856897X00357>.
46. de Cheveigné, A. (2020). ZapLine: A simple and effective method to remove power line artifacts. *Neuroimage* 207, 116356. <https://doi.org/10.1016/j.neuroimage.2019.116356>.
47. Kronland-Martinet, R., Morlet, J., and Grossmann, A. (1987). Analysis of Sound Patterns through Wavelet transforms. *Int. J. Pattern Recognit. Artif. Intell.* 1, 273–302. <https://doi.org/10.1142/S0218001487000205>.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
Dataset and analysis code	CaltechDATA	https://doi.org/10.22002/6bhj2-vy860
Software and algorithms		
MATLAB R2023A	MathWorks	https://www.mathworks.com/
Psychophysics Toolbox	Brainard ⁴⁵ ; Pelli, 1997; Kleiner et al. 2007	http://psychtoolbox.org/
Central Suite	Blackrock Neurotech	https://blackrockneurotech.com/
Other		
Neuroport System	Blackrock Neurotech	https://blackrockneurotech.com/
Tobii Eye Tracker	Tobii AB	https://www.tobii.com/

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Participants

Participants JJ and RD are right-handed males (ages 55 and 40) enrolled in a brain-machine interface (BMI) clinical trial (clinicaltrials.gov identifier: NCT01958086), approved by the institutional review boards of Caltech, Casa Colina Hospital, and UCLA. Participant JJ sustained a C4-C5 level spinal cord injury that occurred approximately 3 years prior to enrollment. He retains voluntary control of the eyes, head, and shoulders. Participant RD sustained a C3-C4 level spinal cord injury approximately 5 years prior to enrollment. He retains similar control of the eyes, head, and shoulders, and shows weak residual movements of the wrists and thumbs. Both participants were clinically stable at the time of participation. Presurgical functional MRI confirmed task-related activation in regions near the planned implant sites. Both participants were male; the influence of sex and gender on the results cannot be determined from this study and should be considered a limitation to generalizability.

METHOD DETAILS

Electrophysiological Recordings

Participant JJ was implanted in 2018 with two 96-channel Utah arrays targeting the left precentral gyrus (denoted MC/JJ) and superior parietal lobule (SPL) (denoted SPL/JJ). Participant RD was implanted in 2023 with four 64-channel Utah arrays targeting the left precentral gyrus (denoted MCM/RD, MCL/RD), SPL (denoted SPL/RD), and supramarginal gyrus (SMG). The SMG array in RD did not exhibit reliable task-related activity across sessions and was therefore excluded from all analyses. Signal quality and unit isolation were comparable to other arrays; however, activity in this region did not exhibit consistent task-related modulation across sessions. Given the anatomical and functional heterogeneity of SMG, the sampled subregion may not have overlapped with sites previously implicated in action execution or observation. All arrays were 4×4 mm with 400 μ m interelectrode spacing (Blackrock Microsystems). Neural signals were recorded using a 128-channel neural signal processor (NeuroPort System, Blackrock Neurotech). Multi-unit activity (MUA) was sampled at 30 kHz and high-pass filtered at 750 Hz. Action potentials were detected using a threshold of -3.5 times the root mean square (RMS) of the high-pass filtered (250 Hz full bandwidth signal). Local field potentials (LFPs) were recorded continuously at a sampling rate of 1 kHz.

Experimental setup

Experiments were conducted at Casa Colina Hospital and Centers for Healthcare for Participant RD, and at home for Participant JJ. In all sessions, participants remained seated in their motorized wheelchairs with their hands resting prone on a flat surface, in a well-lit room. A 30-inch LCD monitor was positioned directly in front of them. Stimulus presentation was controlled using the Psychophysics Toolbox⁴⁵ for MATLAB (MathWorks).

Action Intention and Observation Task

The experimental design consisted of a primary task performed under two cognitive conditions: intention and observation. In both conditions, participants viewed identical visual stimuli—a virtual hand performing actions—on a monitor positioned directly in front of

them. Stimulus timing and content were matched across conditions. Eye position was continuously monitored using a Tobii eye tracker to confirm that participants attended to the stimuli, although no fixation requirement was imposed.

In the intention block, participants were instructed to internally generate the cued action using the specified hand while looking at the screen. Participant JJ remained physically still but volitionally intended the instructed action. Participant RD, who retained partial motor function, was instructed to overtly perform the cued action. Although his movements did not always replicate the precise kinematics of the video stimuli due to physical limitations, we allowed naturalistic execution and instructed him to maintain consistency in the type of response across trials and sessions. In the observation block, participants passively viewed the same videos while remaining still. They were instructed to observe the actions without intending any movement.

Each trial began with a 0.5-second inter-trial interval (ITI), followed by a 0.5-second hand cue. The hand cue consisted of a static image of a virtual left or right hand holding a small parallelogram object between the thumb and index finger, indicating the instructed effector. This was followed by a 1-second symbolic action cue, presented as an overlaid arrow indicating both the action type and direction. Arrow shape specified the action: straight for sliding, curved for lifting, and circular for rotating. The endpoint of the arrow indicated the direction of movement (leftward or rightward). For sliding and lifting, this corresponded to a horizontal displacement of the object to the left or right, while for rotation it corresponded to a clockwise or counterclockwise turn. For consistency, we refer to this factor abstractly as left/right direction throughout the manuscript. The disappearance of the symbol served as the go cue. Immediately afterward, the static frame of the hand transitioned into a 1-second video showing the hand performing the cued action in the indicated direction. The task followed a fully crossed 2 (hand) $\times$ 3 (action) $\times$ 2 (direction) design, resulting in 12 unique conditions. Each participant completed 12 repetitions per condition, per block, in each session. Participant JJ completed five sessions, yielding a total of 421 units in MC and 326 units in SPL, aggregated across sessions. Participant RD completed six sessions, yielding 441 units in MCM, 479 units in MCL, and 532 units in SPL.

Control Experiment Without Observed Movement During Intention

To control for the possibility that cross-format effects were driven by shared visual input between intention and observation, participant JJ completed three additional sessions in which intention trials were performed without viewing the moving hand stimulus. The intention condition followed the same structure as the main task: each trial began with a 0.5 s inter-trial interval (ITI), followed by a 0.5 s hand cue and a 1 s symbolic action cue indicating hand, action type, and movement direction. After the go cue, instead of transitioning into a movement video, the display remained a static image of the hand and object for 2 s, and the participant internally generated the instructed action while viewing this static frame. In the observation condition, each trial began with a 0.5 s ITI, after which the same action videos used in the main experiment were presented at a slower playback speed (2 s duration instead of 1 s) without preceding hand or symbolic cues. Participants passively viewed the videos without intending movement. The task retained the same fully crossed 2 (hand) $\times$ 3 (action) $\times$ 2 (direction) structure as the main experiment. Across the three sessions, this control experiment yielded 270 units in MC and 230 units in SPL.

Dissociation Task Variants

To examine how task demands and cue congruency shape action representations, Participant RD completed two additional experiments with an observation phase that is dissociated from the instruction phase. These experiments were performed only with RD, who retained partial motor function and was able to overtly execute the instructed actions. This was essential to ensure correct task performance in the dissociation paradigm, where precise execution of the cued action—despite incongruent visual input—was required. To reduce complexity, we fixed the direction of movement to rightward and used only two action types: *slide* and *rotate*, performed with either the left or right hand. This yielded a 2 (action) $\times$ 2 (hand) design. For rotation, the rightward direction corresponded to a clockwise turn, while for sliding it corresponded to a rightward displacement.

Each trial began with a 0.5-second inter-trial interval (ITI) displaying a fixation cross. This was followed by:

- a 0.5-second hand instruction screen, indicating the instructed effector (left or right hand),
- a 1-second symbolic cue, showing an overlaid arrow that specified the instructed action type (slide or rotate),
- and a 1-second go phase, during which a video of a hand performing an action was displayed.

The action in the video could be congruent or incongruent with the instructed action. Specifically, the trial could present one of four cue–video pairings:

1. Fully congruent (same action and hand),
2. Action incongruent (same hand, different action),
3. Hand incongruent (same action, different hand), or
4. Action and hand incongruent (different action and hand).

The participant was instructed to perform the cued action with the specified hand concurrently with the video playback, regardless of what was shown.

Dissociation: No-Probe Task

This version included two blocks: an intention block and an observation block. In the intention block, the participant executed the instructed action while watching the video. In the observation block, the participant passively viewed the same videos, with no movement or intention. No response was required after the trial. This allowed us to dissociate encoding of instructed versus observed actions.

Dissociation: Probe Task

This version was identical in structure with the no – probe intention block but added a 1-second probe screen immediately after the video. On each trial, the probe queried either the hand or the action shown in the video. Two icons were presented (left vs. right hand, or slide vs. rotate), and the participant was required to make a saccade to the correct icon to report what they had observed. This response was recorded using the eye-tracking system. By forcing explicit recall of the observed action or effector, while the participant simultaneously performed the instructed action, this task ensured that both streams of information were behaviorally relevant.

QUANTIFICATION AND STATISTICAL ANALYSIS

Spatial Distribution of Gaze

To visualize spatial gaze patterns during the experiment, we computed two-dimensional heatmaps of gaze position separately for the intention and observation blocks, combining data across all runs and sessions. Gaze coordinates from all trials were concatenated, and two-dimensional histograms were computed using a fixed grid (100×100 bins) spanning the full range of observed horizontal and vertical gaze values. The resulting gaze density distributions were smoothed using interpolation for visualization, yielding continuous heatmaps that reflect the spatial concentration of gaze over the course of the experiment. To assess whether trial-by-trial variations in eye movements could account for neural variance or encode task-related information, a four-dimensional gaze feature vector was constructed for each trial during the video epoch (1.5–2.5 s after stimulus onset). This feature vector consisted of the mean horizontal and vertical gaze position (μ_x , μ_y) and the standard deviation of horizontal and vertical gaze position (σ_x , σ_y). These gaze features were used for both encoding and decoding analyses performed independently within each session, array, and task format (intention and observation). For the encoding analysis, ordinary least squares linear regression was used to predict the trial-wise mean firing rate of each unit from the z-scored gaze feature matrix and an intercept term. Statistical significance of the variance explained (R^2) was assessed using a 1,000-iteration permutation test in which the trial order of the neural responses was randomly shuffled relative to the gaze feature matrix. Resulting p-values were corrected for multiple comparisons using the Benjamini–Hochberg false discovery rate procedure (FDR, $\alpha = 0.05$). To assess population-level effects, true and permuted R^2 values were pooled across sessions within each format and compared using two-sided Wilcoxon signed-rank tests. For the decoding analysis, the same gaze feature matrix was used to classify action type, effector, and movement direction using five-fold cross-validated linear discriminant analysis (LDA). Decoding accuracy was defined as the mean classification accuracy across folds. Statistical significance was assessed using a 1,000-iteration permutation test with shuffled trial labels, followed by Benjamini–Hochberg false discovery rate correction (FDR, $\alpha = 0.05$).

Single Unit Activity (SUA) and Multiunit Activity (MUA)

Each detected waveform consisted of 48 samples (1.6 ms total), including 10 samples before threshold crossing and 38 samples after. Single-unit and multiunit activity were sorted using Gaussian mixture modeling applied to the first three principal components of the waveform shapes.¹⁵ In addition to spike assignments, the sorting procedure provided a quality factor ranging from 1 to 4, determined by (1) the percentage of interspike intervals shorter than 3 ms, (2) the signal-to-noise ratio of the mean waveform, (3) the projection distance between clusters, (4) the modified coefficient of variation of the interspike intervals (CV2), and (5) the isolation distance of each cluster. Units with a quality factor of 1 or 2 were considered well-isolated, whereas those with a factor of 3 or 4 were classified as multiunit activity.¹⁵ After spike sorting, net average responses were computed in 100 ms bins by subtracting baseline activity (–500 to 0 ms before stimulus onset) from the post-stimulus activity (0 to 3000 ms after stimulus onset) on a trial-by-trial basis.

Local Field Potential (LFP)

Line noise was suppressed using a combined spectral and spatial filtering approach that preserves underlying neural signals while attenuating non-neural components.⁴⁶ The data were then high-pass filtered above 2 Hz using a zero-phase infinite impulse response (IIR) filter. Trials in which the broadband signal amplitude exceeded two standard deviations from the session mean were excluded from further analysis. High-gamma (60–120 Hz) power was estimated using Morlet wavelet convolution with a 7-cycle resolution.⁴⁷ To minimize edge artifacts introduced by filtering and wavelet convolution, the first and last 100 ms of each trial were discarded. Power was normalized within each trial by dividing the time–frequency power at each frequency by the mean power at that frequency during the 500 ms pre-stimulus baseline.

Cross-session unit matching analysis

To assess the potential re identification of the same single units across recording days, we quantified similarity of units recorded on the same electrode between pairs of sessions in RD. For each unit, we extracted the mean spike waveform and the inter spike interval distribution. Mean waveforms were normalized by subtracting their mean and scaling to unit norm, and waveform similarity for a given pair was computed as the dot product between normalized waveforms. Similarity of inter spike interval distributions was quantified using a two sample Kolmogorov Smirnov test. For each session pair, we constructed all within electrode unit pairs across the two sessions and computed waveform similarity and the Kolmogorov Smirnov p value for each pair. To estimate chance levels, we generated a null distribution by repeatedly sampling unit pairs from different electrodes across the two sessions and computing the same two metrics. For each within electrode pair, we obtained an empirical p value for waveform similarity as the proportion of null pairs with equal or larger waveform similarity, and an empirical p value for inter spike interval similarity as the proportion of null pairs with equal or smaller Kolmogorov Smirnov p value. We combined the two empirical p values using Fisher's method and classified a pair as a putative match if the joint p value was below 0.05. For each session pair, we report the percentage of within electrode pairs classified as matches. For analyses in which cross-session duplicates were excluded, putative matches were resolved by retaining only the earliest recorded instance of a unit in chronological order. Specifically, when a unit in a later session was classified as a match to a unit recorded previously on the same electrode, the later instance was removed from subsequent analyses, and only the first occurrence was retained.

Data Aggregation

Unless otherwise specified, each quantification and statistical analysis was conducted at the level of individual arrays and additionally repeated after pooling motor cortex arrays across participants (MCM and MCL in RD, MC in JJ) into a single MC group and pooling SPL arrays across participants into a single SPL group. For spiking data in RD, analysis was also repeated separately for multiunit and single-unit activity and after excluding putative cross session matches identified in the unit matching procedure.

Statistical Comparison to Baseline

To identify task-modulated activity, we compared post-stimulus responses to baseline on a per-condition basis. For high-gamma analyses, this was performed at the channel level, for spiking activity, at the unit level. Neural responses (spike rates or high-gamma power) were binned into overlapping 200 ms windows with a 100 ms step size, starting at stimulus onset. For each bin, a paired t -test compared the binned activity to the mean activity during the 500 ms pre-stimulus baseline across trials. Statistical significance was assessed using a Bonferroni-corrected threshold ($\alpha = 0.05 / \text{number of bins}$). A channel or unit was classified as significantly responsive to a condition if it exhibited at least three consecutive time bins with significant deviation from baseline.

Facilitated and Suppressed Responses

For units classified as significantly responsive to a given condition (see [Statistical Comparison to Baseline](#)), we further determined the direction of modulation. For each unit and condition, we computed the mean activity during a baseline window from -0.5 to 0 s and during a task window from 1 to 2 s. The trial-wise difference between task and baseline activity was calculated and averaged across trials to obtain a single signed modulation value per unit and condition. Units with positive modulation values were classified as facilitated, whereas units with negative values were classified as suppressed. For each condition, we quantified the proportion of facilitated and suppressed responses relative to the total number of significant units.

Latency

For each condition and each significantly responsive unit, response latency was defined as the center of the first time bin within the earliest sequence of three consecutive bins showing significant modulation relative to baseline. This latency reflects the earliest consistent deviation from baseline activity. To assess regional differences in response timing, we compared latency distributions across brain areas using one-way ANOVAs, conducted separately for each participant or region for the pooled analysis, and for each task format (intention and observation). Each unit contributed a single latency value per condition, and group-level comparisons tested for significant differences in peak response timing between motor and parietal regions.

Tuning Analysis

To quantify selectivity for task variables, we performed a time-resolved three-way ANOVA separately for spiking activity (multi- and single-unit) and high-gamma (HG) power. For each unit or channel and for each 500 ms time bin, we computed the mean response across time. A three-way ANOVA was then used to test for main effects of action type (rotate, slide, lift), effector (left vs. right hand), and movement direction (leftward vs. rightward), as well as their three-way interaction. For the dissociation tasks, in which movement direction was held constant, a two-way ANOVA was performed with action and hand as factors. Trial labels were parsed from condition names into categorical variables corresponding to each factor. Significance was assessed using a Bonferroni-corrected threshold ($\alpha = 0.05 / \text{number of units or channels}$). A unit or channel was classified as selectively tuned to a main effect (e.g., action type) if the corresponding factor reached significance in the absence of a significant interaction. Tuning to a combination of factors was labeled as an interaction effect. This analysis was performed independently for each time bin and brain region, and the number of significantly tuned units or channels was tracked across time for each tuning category.

Overlap of Task-Relevant Units Across Formats

To assess the distribution of task-related neural responses across formats, we identified task-responsive units independently for the intention and observation blocks. A unit was classified as task-responsive if it exhibited a significant increase in activity relative to baseline for at least one task condition (see [Statistical Comparison to Baseline](#)). This analysis was performed separately for each brain region and format. We then quantified the degree of overlap between formats by computing, for each region, the number of units responsive exclusively during intention, exclusively during observation, or during both. These distributions were used to evaluate the extent to which task-related neural activity was shared or format-specific across the two conditions. Differences in the distribution of response categories across regions were assessed using a chi-square test of independence. For visualization, we constructed heatmaps of condition-averaged activity for units in each category. For each unit, responses were averaged across trials separately for intention and observation. A single representative condition was selected per unit based on maximal absolute modulation in the relevant format, or on the combined modulation across formats for units responsive in both.

Linear Model Analysis

We classified units based on their selectivity to task variables across observation and intention conditions. We implemented a structured model comparison framework similar to that described by Chivukula et al. (2025).²⁵ For each neuron, firing rates were averaged over a fixed task window (1–2 s after trial onset) for each of the 12 observation and 12 intention conditions. The analysis was applied to all recorded units and was not restricted to those identified as task-responsive in the sliding-window baseline test, as here we were interested in modulation according to task variables, which may be more subtle. The 24 condition-averaged responses were then combined into a single dataset for that unit, and a series of linear regression models were fit to predict neural responses based on different combinations of experimental factors (hand used, action type, movement direction) and format-specific terms (intention vs. observation).

- (1) a null model including only a constant term (*Unselective*)
- (2) a fully shared model including action, hand, and direction as predictors with the same weights across formats (*Invariant*)
- (3–5) models with shared tuning to a specific task variable: action, hand, or direction (*Action, Hand, or Direction*)
- (6) an additional model in which “action” was redefined as six distinct action × direction combinations (*Action (6)*)
- (7) a fully shared model with all main effects and interactions (*Mixed*)
- (8) a format-specific model with separate parameters for intention and observation (*Idiosyncratic*)
- (9–10) models including task features only for intention or only for observation trials (*Single format*)

Model performance was assessed using five-fold stratified cross-validation based on condition labels, and each unit was assigned to the model with the highest cross-validated R^2 value. To determine whether the observed R^2 exceeded chance, we implemented a permutation test in which the neural responses were randomly shuffled across trials while keeping the design matrix fixed. For each unit, we computed a null distribution of R^2 values (1,000 permutations) and derived a one-tailed p value based on the proportion of null R^2 values exceeding the true R^2 . p values were corrected for multiple comparisons across units using the Benjamini–Hochberg false discovery rate (FDR) procedure ($q = 0.05$). Units were considered selective if they met all of the following criteria: (1) cross-validated $R^2 > 0.01$, (2) permutation-derived p value < 0.05 after FDR correction, and (3) the best-fitting model was not the null model. Units that did not meet all three criteria were classified as unselective. Differences in the distribution of best fitting models across regions were assessed using a chi-square test of independence.

Population Analysis

Representational Similarity Analysis

To evaluate the structure of neural representations across intention and observation, we performed a cross-validated representational similarity analysis (RSA), separately for each task variable: action type (3 levels), effector/hand (2 levels), and movement direction (2 levels). All analyses were conducted independently per brain region. RSA was computed both *within format* (intention–intention and observation–observation) and *across format* (intention–observation), using the same framework across all comparisons. Neural responses were extracted from a fixed 1–2 s post-stimulus window, and each trial was reshaped into a single feature vector by concatenating the time and unit dimensions. For each of 500 random splits, trials were divided into halves within each condition. Condition-averaged activity vectors were computed independently for intention and observation trials, and Pearson correlations were calculated between all pairs of vectors across formats, resulting in a condition-by-condition cross-format RSA matrix per split. These matrices were averaged to obtain a final similarity matrix. To test whether the observed structure reflected meaningful condition-specific information, we compared the similarity between matched and unmatched conditions by computing the mean diagonal and off-diagonal values of each RSA matrix. The difference between these values (diagonal – off-diagonal) served as a measure of structure strength. To generate null distributions, we repeated the same procedure after shuffling condition labels independently within each format prior to trial splitting. Statistical significance was assessed by comparing the observed difference to the empirical null distribution from shuffled data, separately for the cross-format matrix and for each within-format matrix (intention and observation). In addition, we computed the full RSA matrices and permutation-matched nulls for each format and task variable. Mean diagonal values were extracted as a measure of within-condition reliability and representational consistency. To quantify the variability of diagonal similarity values across conditions, we additionally computed the coefficient of variation (CV; standard deviation divided by

the mean) across the 12 diagonal values of each RSA matrix. Lower CV values reflected more uniform representational similarity across conditions, whereas higher CV values reflected greater condition-specific variability.

To assess the robustness of the RSA results, we conducted control analyses using four alternative time windows (0–1 s, 0.5–1 s, 1.5–2.5 s, and 2–3 s). For each window, we recomputed RSA matrices for intention, observation, and cross-format comparisons, and then quantified similarity to the main analysis (1–2 s window) by computing the Pearson correlation between the lower triangles (excluding the diagonal) of each matrix and the corresponding 1–2 s RSA matrix. This analysis provided a direct measure of how stable the representational structure was across different temporal windows.

Linear Discriminant Analysis (LDA) Within Format

We assessed whether neural population activity encoded task-relevant information within each format, by training LDA classifiers using a time-resolved, cross-validated decoding framework. The analysis was performed separately for each brain area and each classification level. For the main task, classifiers were trained to distinguish among: (1) the 12 fully crossed task conditions (hand $\times$ action $\times$ direction), and (2) individual task variables: action type (3 levels), effector/hand (2 levels), and movement direction (2 levels). Trials were relabeled accordingly, and each classification problem was evaluated independently. For the dissociation tasks, decoding was performed in two complementary ways. First, classifiers were trained to distinguish among all 16 trial types, defined by the 2 (action) $\times$ 2 (hand) $\times$ 4 (conflict type) design. Second, to separately assess encoding of the instructed and observed actions, trials were relabeled based on the 2 (action) $\times$ 2 (hand) combinations corresponding to either the instructed cues or the observed video, and decoding was performed independently for each. For the dissociation tasks in RD, this within-format decoding analysis was additionally repeated separately for MUA and SUA to verify that population-level effects were not driven by signal type. For all analyses, neural responses were binned in 100 ms steps, and decoding was performed using non-overlapping 200 ms windows by averaging adjacent bins. Ten-fold cross-validation was used. Within each fold, principal component analysis (PCA) was applied to the training data across all time bins to reduce dimensionality. The number of components was selected to capture 95% of the variance, with an upper limit of 50 components. Both training and test data were projected into this reduced-dimensional space. LDA classifiers were trained on the reduced training data and evaluated on the held-out test set for each time window. Classification accuracy was computed per fold and averaged across folds to yield time-resolved performance curves for each brain region and classification level. For each region, the time window with the highest decoding accuracy was identified, and confusion matrices were computed at that time point. Confusion matrices were normalized by row and visualized as percent classification accuracy. To evaluate the consistency of decoding performance across sessions, we additionally ran the same decoding analysis separately within each session. This was done for each of the three individual task variables (action, hand, direction), and the maximum decoding accuracy per session was extracted. This allowed us to assess whether the results observed in the concatenated analysis were driven by any session-specific peaks or drops in performance.

Cross – Format Decoding

We evaluated whether neural population representations generalized across formats, by training LDA classifiers on trials from one format (intention or observation) and testing them on the other. This analysis was performed separately for each brain region and each task variable: action type (3 levels), effector/hand (2 levels), and movement direction (2 levels). Decoding was conducted in both directions: training on intention and testing on observation, and vice versa. Neural activity was binned in 100 ms steps, and decoding was performed per time bin. For each time bin in the training format, PCA was applied to the full training dataset (all time bins and trials), and the top 50 components were retained. Training and test data were projected into this common PCA space. An LDA classifier trained on each individual training time bin was then evaluated on all test time bins, yielding a full train $\times$ test time bin decoding accuracy matrix. This matrix captures temporal generalization and reveals whether neural representations aligned across different time points between formats. To assess statistical significance, we performed permutation testing by randomly shuffling test set labels across 1000 iterations and recomputing decoding accuracy. For each train–test pair, a p-value was computed as the proportion of permuted accuracies greater than or equal to the observed accuracy. Significance thresholds were corrected for multiple comparisons using Bonferroni adjustment. Cross-format decoding was first computed after pooling motor cortex arrays across participants (MCM and MCL in RD, MC in JJ) into a single MC group and pooling SPL arrays across participants into a single SPL group, and was then repeated at the level of individual arrays. For spiking data in RD, decoding was additionally repeated separately for MUA and SUA, and after excluding putative cross session matches identified in the cross-session unit matching analysis in RD. To summarize cross format generalization at the array level, we computed an integrated decoding strength metric for each individual array and decoding direction. Analyses were restricted to train–test pairs within a post cue window from 0 to 2.5 s. Within this window, we extracted the decoding accuracy matrix and the corresponding binary significance mask. For each array and direction, we computed (i) the mean decoding accuracy across significant train–test pairs only and (ii) the fraction of significant train–test pairs within the window. Integrated decoding strength was defined as the product of these two terms, yielding a single value per array and decoding direction that jointly reflects both effect magnitude and the temporal extent of significant cross format generalization. To assess whether integrated decoding strength exceeded chance levels, we generated permutation derived null distributions by recomputing the integrated decoding strength metric for each of the 1000 label shuffled decoding iterations. For each array and decoding direction, statistical significance was quantified as the proportion of permuted integrated decoding strength values greater than or equal to the observed value, yielding a one sided empirical permutation p value.

Cross – Time Decoding

To evaluate the temporal stability of neural representations, we performed cross-time decoding within the probe variant of the dissociation task, separately for each brain area and for classifiers trained on the instructed or video actions. The analysis followed the same preprocessing and cross-validation procedures described above. For each array, neural activity was binned in 100 ms steps, and non-overlapping 200 ms windows were created by averaging adjacent bins. Within each fold of a ten-fold cross-validation, principal component analysis (PCA) was computed on the training data across all time bins, and both training and test data were projected into the same reduced-dimensional space using the coefficients derived from the training set (95% cumulative variance threshold, maximum 50 components). LDA classifiers were then trained to discriminate either the instructed or the video actions at each time window and tested on all other time windows, generating a two-dimensional matrix of decoding accuracy (train $\times$ test time). This procedure was repeated for all folds, and decoding accuracies were averaged across folds to yield cross-temporal generalization matrices for each array. The diagonal of the matrix reflects standard within-time decoding, while off-diagonal values quantify the degree to which representations generalize across time. Sustained off-diagonal accuracy indicates temporally stable population codes, whereas narrow, diagonal patterns correspond to dynamic, time-specific encoding.

PCA Trajectory Analysis

To visualize the temporal evolution of neural population activity during each task, we performed PCA separately for each brain region and task variable (action type, effector/hand, or movement direction). Only units classified as task-relevant in both formats based on the linear model analysis (i.e., not unselective or single format) were included. Neural responses were extracted from a time window spanning -0.5 to 2.5 s relative to trial onset, using 100 ms binning. The PCA was applied jointly to intention and observation data. For each trial, neural activity was grouped by condition (e.g., action identity) and format (intention or observation). Trials were averaged within each condition and format, resulting in a set of condition-averaged response matrices of size time $\times$ units with the number of matrices equal to the number of task levels (e.g., 3 for action, 2 for hand) multiplied by 2 formats. These matrices were concatenated across time and condition, and the resulting (conditions $\times$ time) $\times$ units matrix was used as input to PCA. The top three principal components were retained, and each condition's time series was projected into this low-dimensional space. The resulting trajectories were plotted separately for intention and observation to compare their temporal geometry. In addition, we quantified the variance explained by the top three components in three ways: based on the PCA decomposition of the combined dataset, and separately for intention and observation by computing the variance captured within each format after projection into the shared PCA space.

Procrustes Analysis

To compare the geometry of neural trajectories between formats, we applied Procrustes analysis to the condition-averaged neural trajectories in PCA space. This analysis was performed separately for each brain region and task variable (action type, effector/hand, or movement direction). For each condition within a task variable (e.g., “lift” or “right hand”), we extracted one trajectory from the intention format and one from the observation format, both represented in the space defined by the top three principal components computed from the combined dataset. Procrustes alignment was then used to align the 3D trajectories of the two formats. This transformation computes the optimal translation, rotation, and isotropic scaling that minimizes the Frobenius norm between the two trajectories. A Procrustes distance was computed for each condition, providing a measure of geometric dissimilarity between formats. Lower values indicate greater similarity in the temporal structure of neural activity between intention and observation for that condition. To compare regions, Procrustes distances were pooled across arrays within MC and SPL and statistically compared using Wilcoxon rank-sum tests. As a control, we repeated the PCA and alignment procedure using all recorded units, and computed Procrustes distances for all brain regions and task variables.

UMAP Visualization of Neural Representations

To visualize the structure of trial-level neural activity in a low-dimensional space, we applied Uniform Manifold Approximation and Projection (UMAP) separately for each brain region. Only task-relevant units (identified via linear model analysis) were included. Neural responses were averaged over a 1 – 2 s window following stimulus onset. Trials from the intention and observation formats were concatenated and projected jointly. UMAP was performed using the cosine distance metric with 15 nearest neighbors and a minimum distance of 0.1 . The resulting two-dimensional embeddings were used to visualize the structure of population activity across task conditions and formats. To visualize the spatial distribution of each condition in the UMAP space, we computed a separate 2D ellipse for each condition and format combination. The center of each ellipse was defined as the empirical centroid of the UMAP coordinates for that group (i.e., the mean across all trials). The shape and orientation of the ellipse were derived from the empirical 2D covariance matrix of the points. We used the eigenvectors and eigenvalues of this matrix to construct an ellipse representing the major and minor axes of the group's spread. This approach provides an interpretable summary of the distribution of trials in the reduced space, allowing visual assessment of separation or alignment between formats.

ADDITIONAL RESOURCES

This study is part of a registered clinical trial ([ClinicalTrials.gov](https://clinicaltrials.gov), NCT01958086; <https://clinicaltrials.gov/study/NCT01958086>).

Cell, Volume 189

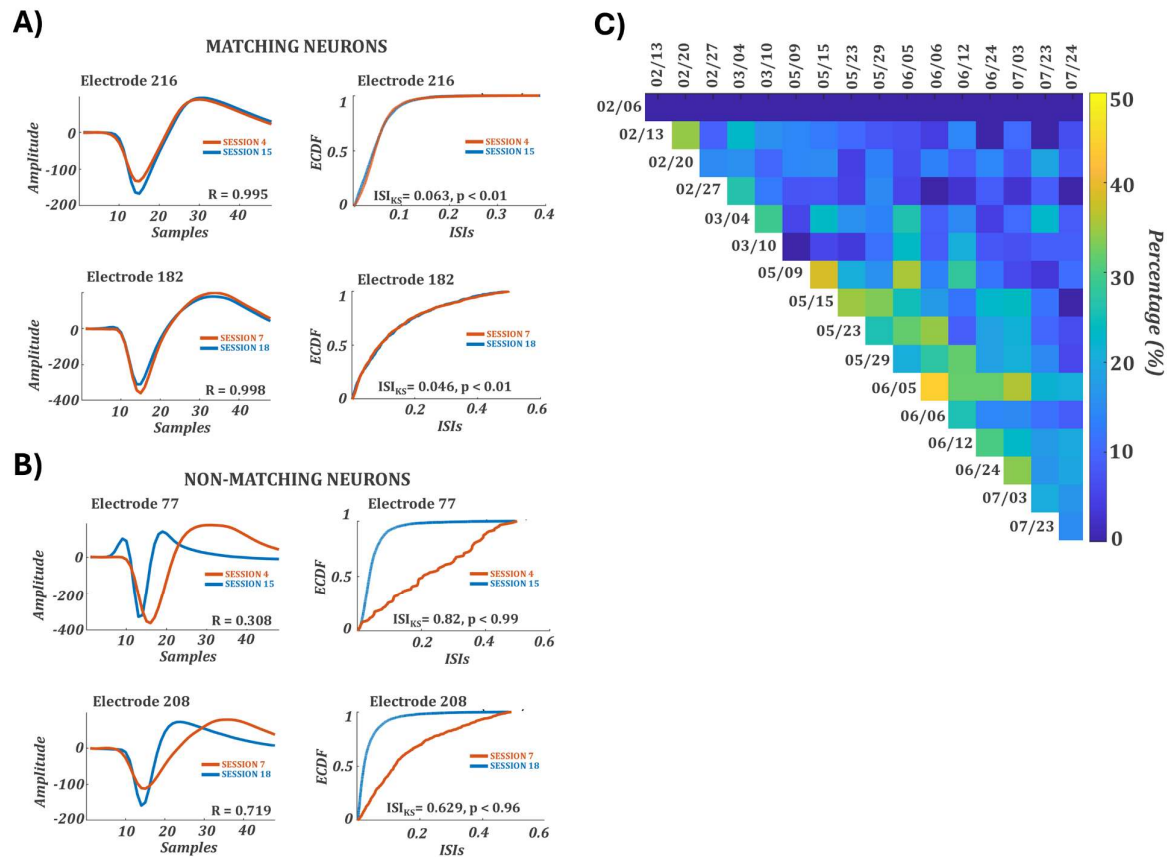
Supplemental information

**Hierarchical and context-dependent encoding
of actions in human posterior
parietal and motor cortex**

Vasiliki Bougou, Jorge Gamez, Emily R. Rosario, Charles Liu, Kelsie Pejsa, Ausaf Bari, and Richard A. Andersen

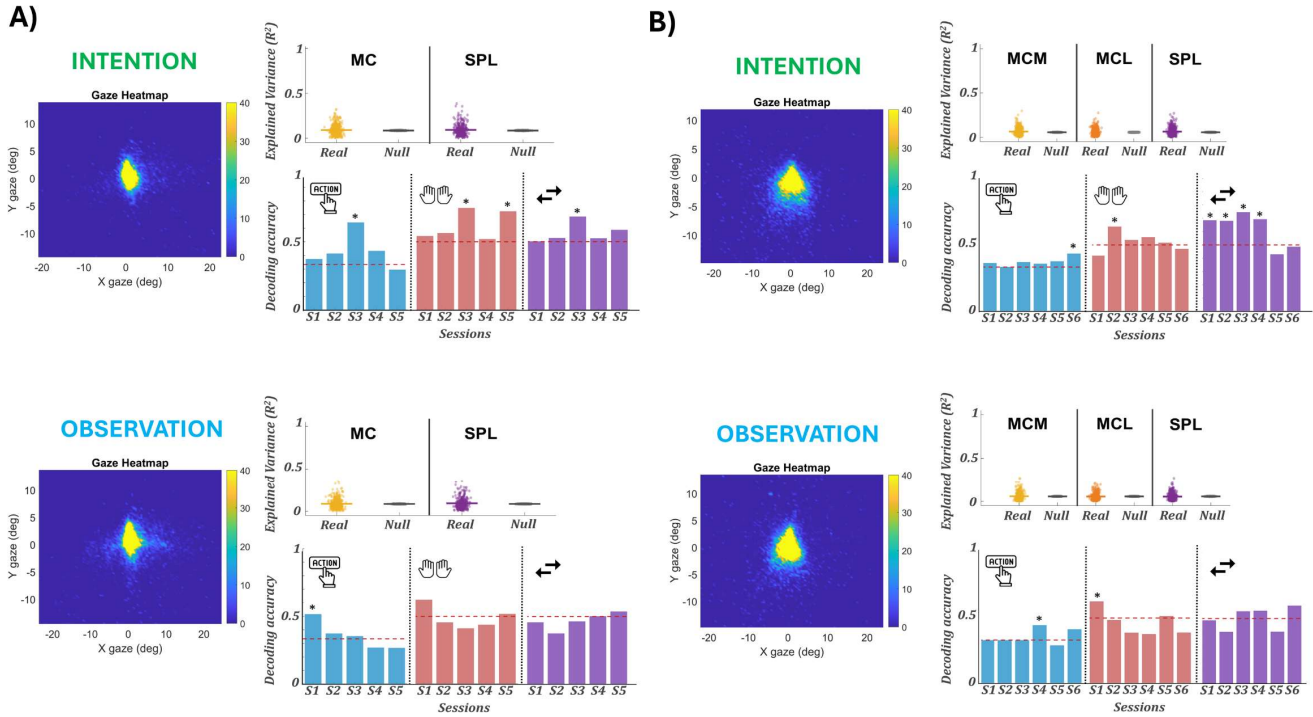
Supplementary Information:

Figure S1: Stability of Single-Unit Recordings Across Sessions, Related to STAR Methods.



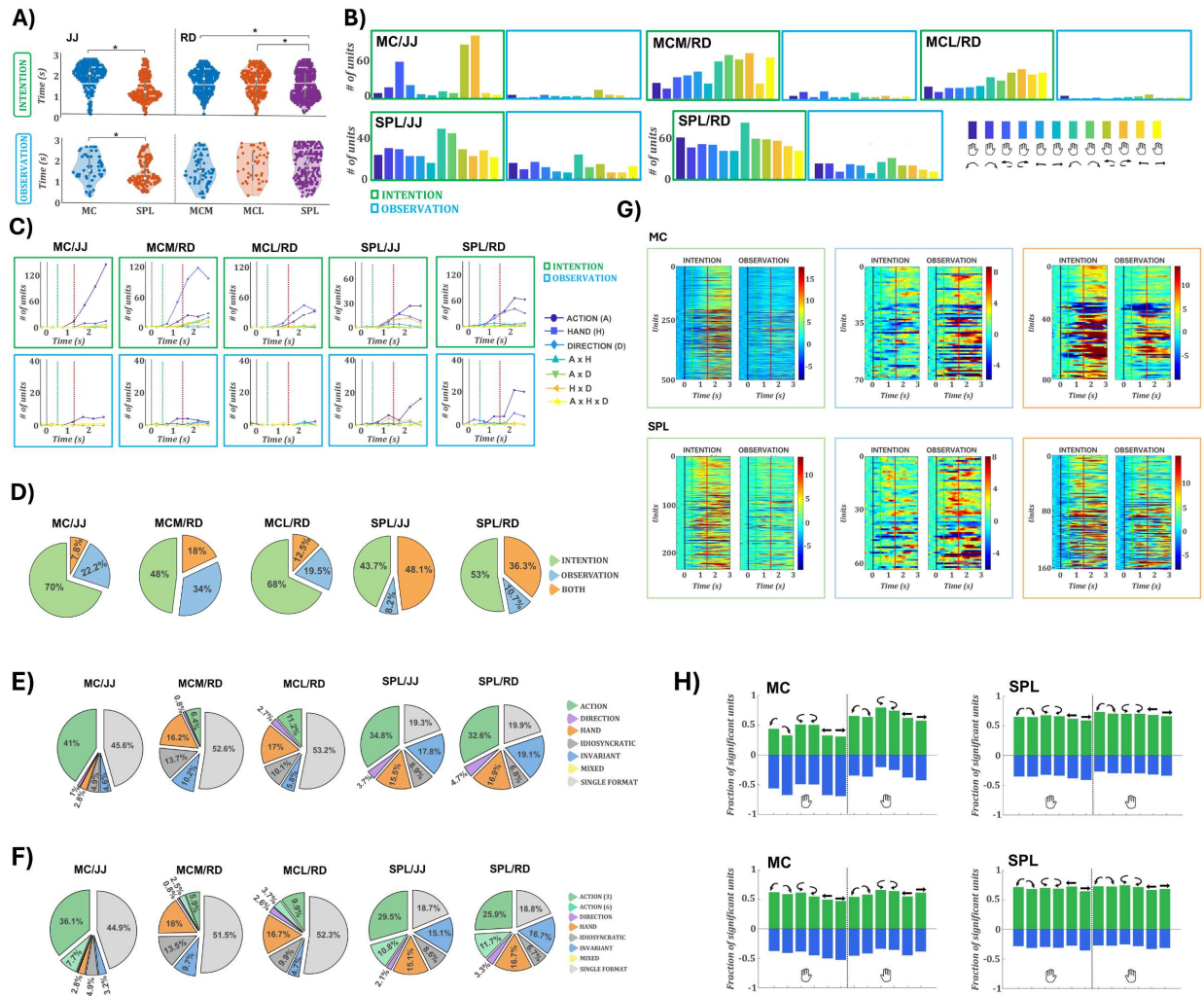
(A) Examples of matching neurons recorded from the same electrode across two sessions, as identified from the cross-session unit matching analysis. Left panels show average spike waveforms aligned to threshold crossing for two different recording days (orange and blue). Correlation coefficients (R) quantify waveform similarity across sessions. Right panels show cumulative inter-spike interval (ISI) distributions for the same units across sessions; Kolmogorov–Smirnov statistics and p-values are indicated. High waveform correlations and overlapping ISI distributions demonstrate stable unit isolation across days. **(B)** Examples of non-matching neurons recorded from the same electrode across sessions. Waveform shapes differ substantially between sessions (lower R values), and ISI distributions show marked divergence, consistent with distinct units recorded on different days. **(C)** Percentage of electrodes exhibiting stable unit matches across pairs of recording dates. Each cell indicates the proportion of electrodes classified as matching (based on waveform similarity and ISI criteria) for a given session pair. Warmer colors indicate higher percentages of stable matches.

Figure S2: Eye Movement Control Analysis, Related to STAR Methods.



(A) Participant JJ. **(B)** Participant RD. Top panels (green) correspond to the intention format and bottom panels (blue) to the observation format. Heatmaps show the 2D distribution of gaze positions (in degrees of visual angle) aggregated across all trials, demonstrating stable and centrally maintained fixation across participants and task formats. Color scale indicates gaze density. Scatter plots show the distributions of real and permuted R^2 values obtained from the gaze-to-neural regression analysis, pooled across sessions for each array and format. Bar plots show gaze-based decoding accuracies for action type, effector, and movement direction across sessions using trial-wise gaze features extracted from the eye-tracking data. Dashed lines indicate chance level, and asterisks denote significant decoding after FDR correction ($p < 0.05$).

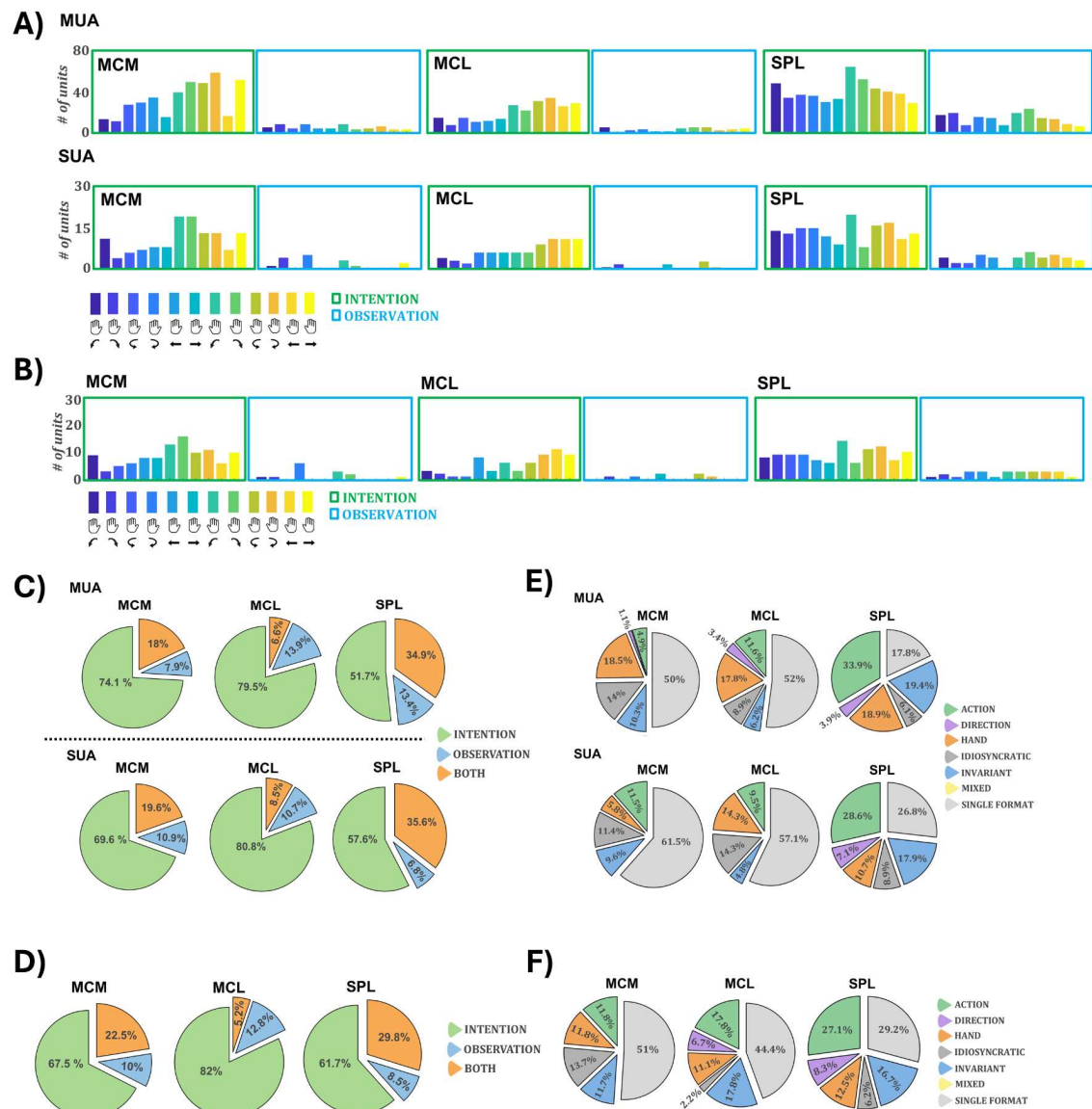
Figure S3: Per-Array Single-Unit Properties and Selectivity Across Intention and Observation, Related to Figure 2.



A) Violin plots of onset latencies across all arrays for intention (top) and observation (bottom). Asterisks indicate statistically significant differences. In JJ, SPL units show significantly earlier responses than MC units in both formats. In RD, SPL responses are earlier than those in both MCM and MCL during intention, but not during observation. **(B)** Number of significantly responsive units per condition (Bonferroni-corrected t-test against baseline, ≥ 3 consecutive bins). Green and blue outlines correspond to intention and observation blocks, respectively. MC/JJ and MCM/RD arrays show robust responses during intention but few during observation. In contrast, SPL arrays show substantial responses in both formats. **(C)** Three-way ANOVA tuning profiles over time (factors: action, hand, direction). Top: intention (green outline); bottom: observation (blue outline). Lines represent main effects and all two- and three-way interactions. In JJ, MC is strongly tuned to action. MCM/RD and MCL/RD are predominantly tuned to hand, followed by action. SPL in both participants shows dominant action tuning and secondary hand tuning. Interaction terms

are rare. During observation, only SPL arrays preserve action tuning. **(D)** Format selectivity: percentage of units with significant responses in intention only (green), observation only (blue), or both (orange). These percentages reflect response presence but not tuning congruency. **(E)** Linear model classification: units are assigned to their best-fitting model based on cross-validated R^2 . Categories include shared action (green), shared hand (orange), shared direction (purple), invariant (blue), mixed (yellow), idiosyncratic (dark gray), and single-format (light gray). Unselective units are not shown. SPL arrays in both participants contain the fewest single-format units and the highest proportion of shared action models. **(F)** Distribution of best-fitting linear models across brain areas including the action (6) model. Units are assigned to their best-fitting model based on cross-validated R^2 . **(G)** Heatmaps showing condition-averaged, baseline-normalized activity of units in MC (top row; pooled motor cortex arrays) and SPL (bottom row; pooled across participants). Colored outlines indicate unit classification: green, units responsive during intention only; blue, units responsive during observation only; orange, units responsive in both formats. For each unit, activity is shown during intention (left) and observation (right) for a representative condition capturing its strongest modulation. Each row corresponds to one unit; color indicates signed modulation relative to baseline across time. **(H)** Top row: Proportion of facilitated (green) and suppressed (blue) responses for each condition during intention, computed relative to the total number of significant units in pooled motor cortex arrays (MCM and MCL in RD, MC in JJ; left) and pooled SPL arrays (RD and JJ; right). Bottom row: Same analysis for observation.

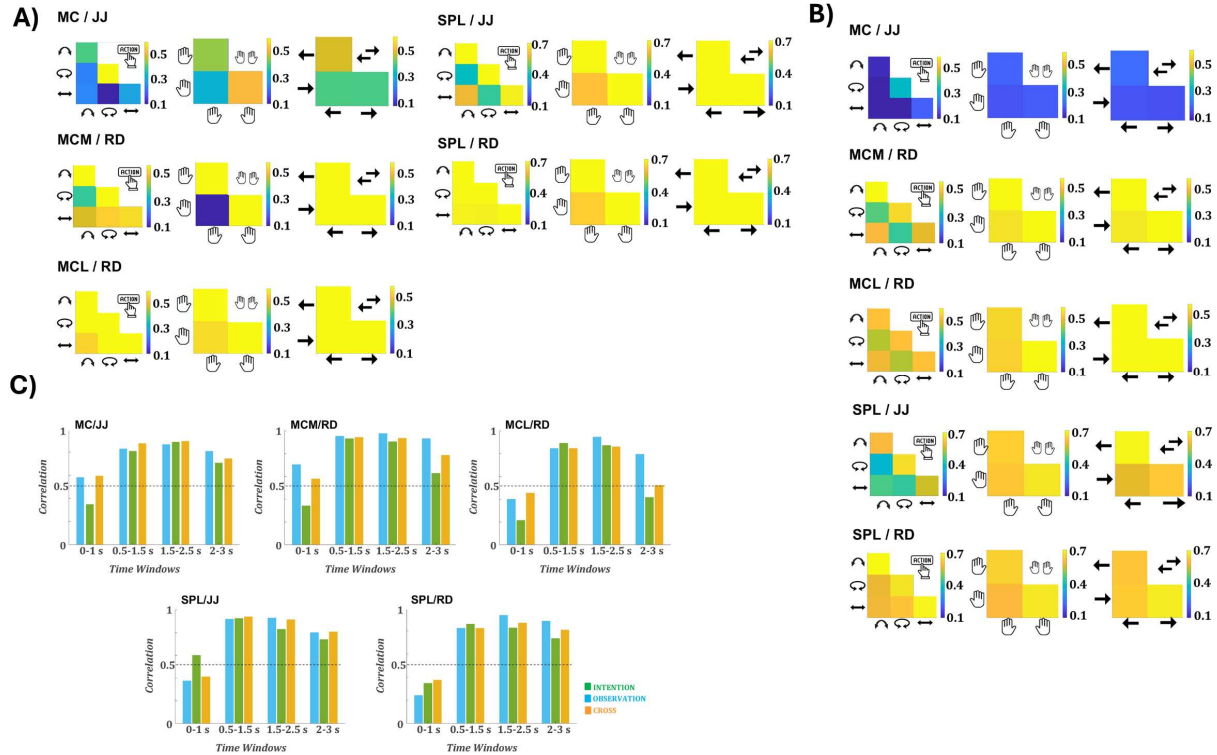
Figure S4: Single-Unit Selectivity Across Intention and Observation for MUA and SUA and After Excluding Cross-Session Matches, Related to Figure 2.



(A) Number of significantly responsive units per condition for *MUA* (top row) and *SUA* (bottom row) in RD (Bonferroni-corrected t-test against baseline, ≥ 3 consecutive bins). Green and blue outlines indicate intention and observation blocks, respectively. **(B)** Number of significantly responsive units per condition in RD after removing putative cross-session matches. **(C)** Format selectivity for *MUA* (top row) and *SUA* (bottom row): percentage of units with significant responses in intention only (green), observation only (blue), or both (orange). **(D)** Format selectivity after removing putative cross-session matches. **(E)** Linear model classification for *MUA* (top row) and *SUA* (bottom row): units are assigned to their best-fitting model based on cross-validated R^2 . Categories include shared action (green), shared hand (orange), shared direction

(purple), invariant (blue), mixed (yellow), idiosyncratic (dark gray), and single-format (light gray). Unselective units are not shown. **(F)** Linear model classification after removing putative cross-session matches.

Figure S5: Within-Format Representational Similarity Analysis (RSA) and Control Analysis for RSA Matrix Stability, Related to Figure 4.



(A) RSA heatmaps computed within intention. Columns correspond to action (left), hand (middle), and direction (right). Each matrix shows the pairwise correlation structure between condition-specific population activity vectors. **(B)** RSA heatmaps within observation. **(C)** Robustness of representational structure across time windows. Bars show correlations between the lower triangle of the 12-condition RSA matrices computed in the main analysis window (1–2 s) and alternative windows: 0–1 s, 0.5–1.5 s, 1.5–2.5 s, and 2–3 s. Colors indicate format: within observation (blue), within intention (green), and cross-format (orange). Correlation remains high for all windows except the earliest (0–1 s), confirming the stability of the representational structure over time.

Figure S6: Decoding Task Variables Within and Across Formats Per Array, Related to Figure 4.

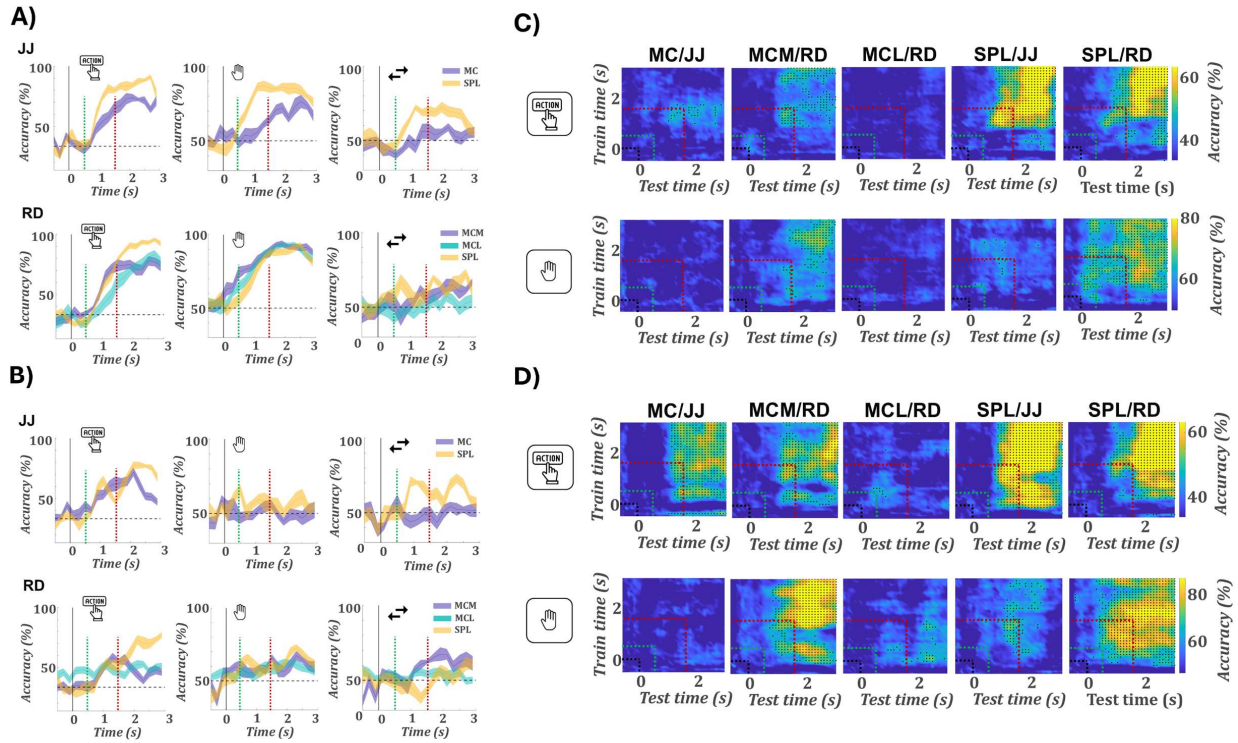
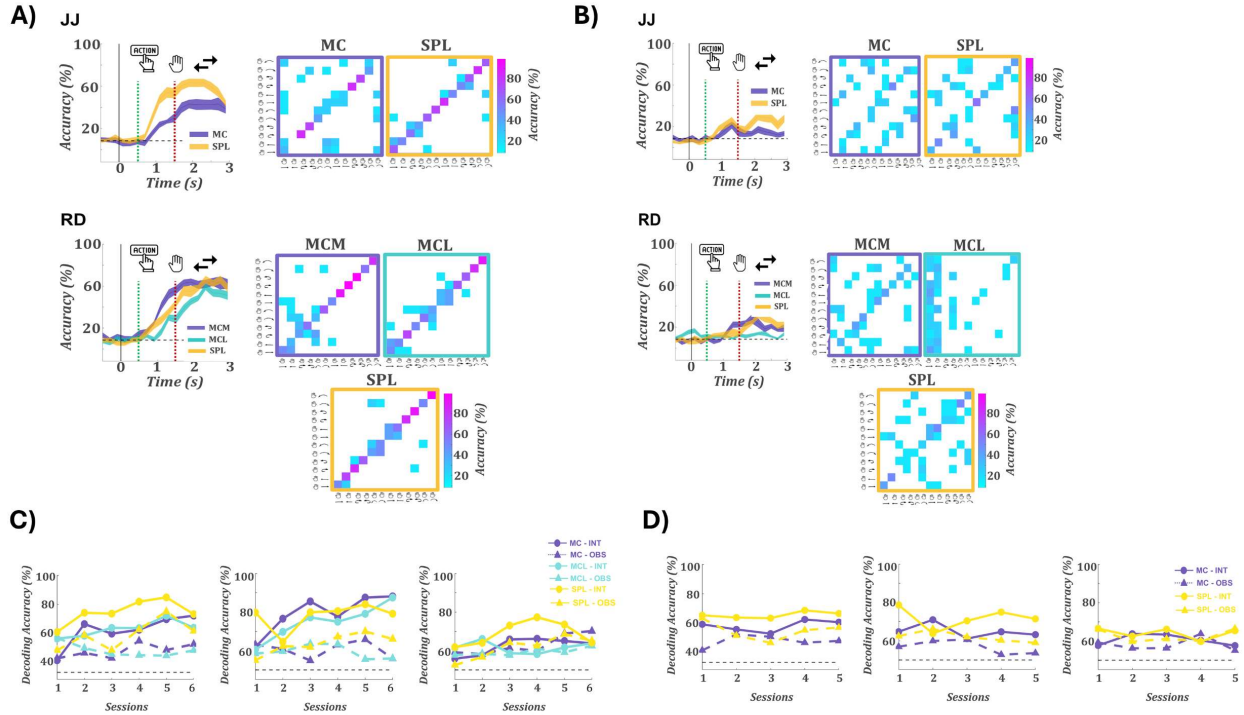


Figure S7: Decoding of All Task Conditions and Session-Wise Decoding Accuracy Within Format, Related to Figure 4.



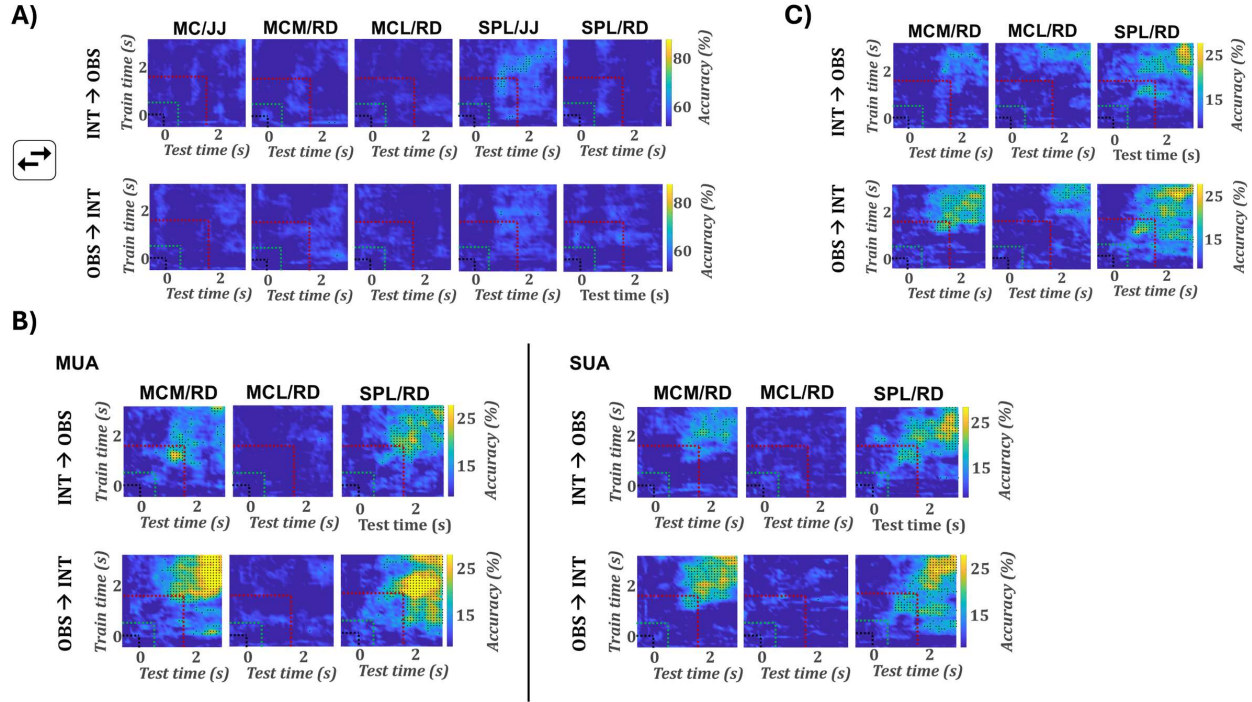
A) Intention. Left: time-resolved LDA decoding accuracy for all 12 task conditions (2 hands × 3 actions × 2 directions). Top: JJ (MC, SPL); bottom: RD (MCM, MCL, SPL). Lines show average accuracy across cross-validation folds; shaded regions indicate ±SEM across cross validations. Green and red dashed lines mark the symbolic action cue and go cue, respectively. All arrays achieve high classification accuracy, with SPL consistently performing best. Right: corresponding confusion matrices show that most misclassifications occur between different directions of the same action, indicating that action identity and hand are more robustly represented than direction.

B) Observation. Same format as A. Decoding accuracy is substantially lower across all arrays. SPL outperforms other regions in both participants, but peak accuracy remains below 40%. Confusion matrices show reduced separability between conditions, consistent with weaker observation-related encoding.

(C) Participant JJ. (D) Participant RD. Line plots show decoding accuracy for each session across all task variables: action (left), hand (middle), and direction (right). Solid lines and circle markers represent intention; dashed lines and triangles represent observation. Colors indicate recording arrays (see legend). MC/MCM/MCL are shown in purple and cyan tones; SPL is yellow. Action and hand decoding remain consistently above chance across sessions, indicating stable encoding. SPL shows the highest decoding accuracy for action during observation across nearly all sessions. Direction decoding is

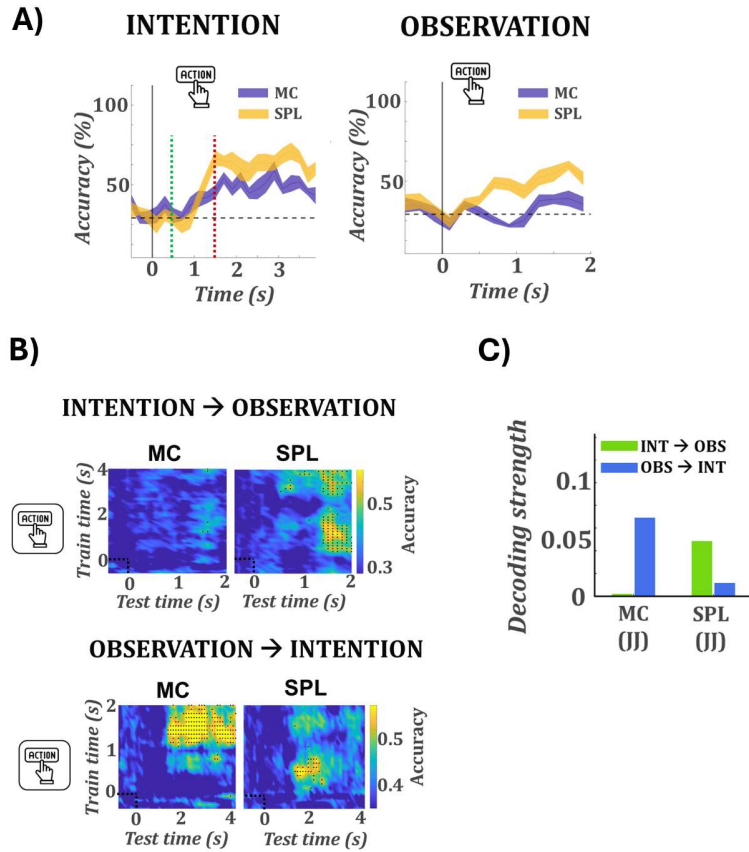
consistently poor across all arrays and formats. Overall, decoding performance is stable across sessions, particularly for action and hand representations in intention.

Figure S8: Control Cross-Decoding Analyses, Related to Figure 4.



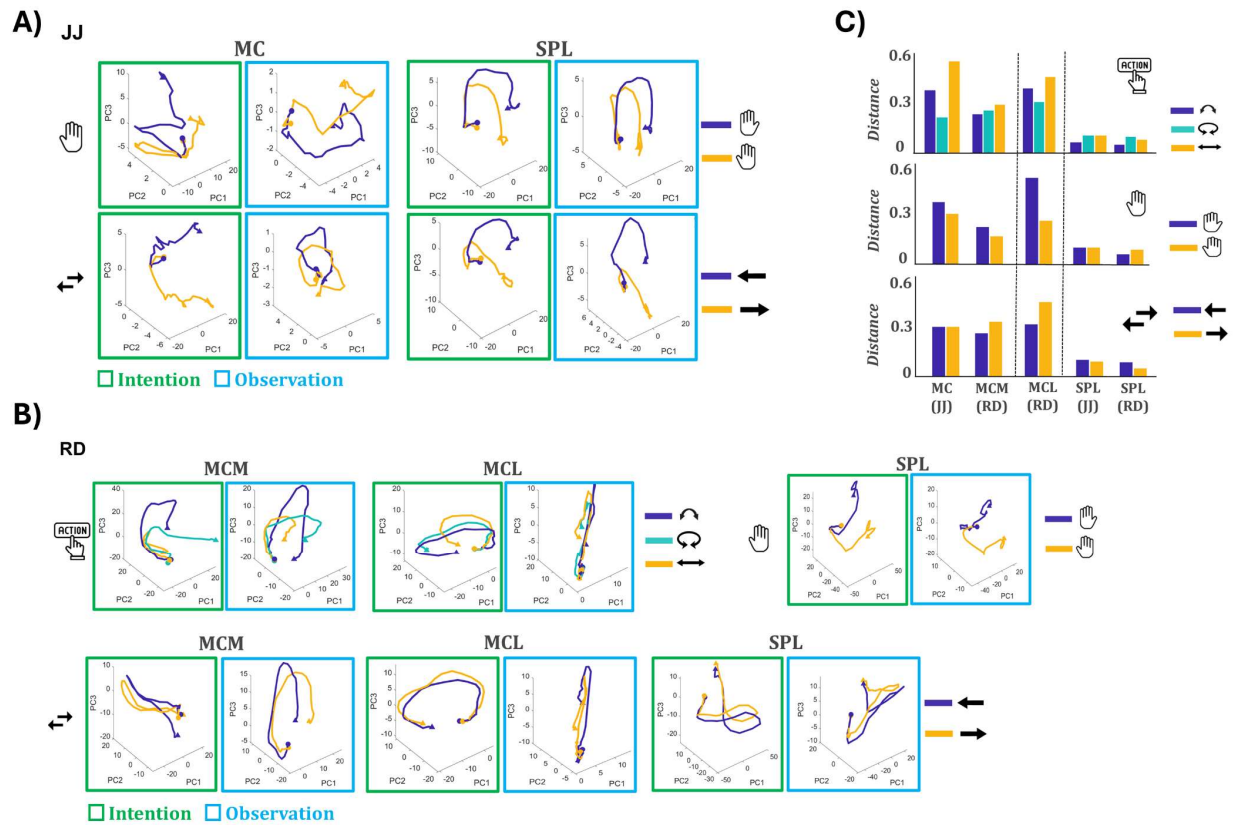
Each heatmap shows decoding accuracy across all combinations of training (y-axis) and testing (x-axis) time bins. Green and red dashed lines indicate the onset of the symbolic action cue and the go cue, respectively. Significant decoding time points (permutation test, $p < 0.05$) are marked with dots. **(A)** Cross-temporal decoding of movement direction across formats for each array. Classifiers were trained in one format and tested in the other (INT → OBS, top; OBS → INT, bottom). Direction could not be reliably cross-decoded in any array. **(B)** Control analysis: cross-decoding of the full 12 task conditions in RD, shown separately for MUA (left) and SUA (right). Both signal types show the same pattern as in the main analysis: unidirectional cross-decoding in MCM (OBS → INT only), no cross-decoding in MCL, and bidirectional cross-decoding in SPL. **(C)** Control analysis after removing putative cross-session matched units in RD. The same pattern persists, confirming that the observed effects are not driven by unit resampling.

Figure S9: Control Experiment in Participant JJ, Related to Figure 4.



(A) Time-resolved LDA decoding accuracy for action type in the control experiment in participant JJ. Left: intention block. Right: observation block. Lines show mean decoding accuracy across cross-validation folds; shaded regions indicate $\pm$ SEM. Green and red dashed lines mark the symbolic action cue and the go cue, respectively. Action type could be decoded above chance during intention in both MC and SPL, whereas during observation decoding remained robust only in SPL, consistent with the main experiment. **(B)** Cross-format decoding for action type. Top: classifiers trained on intention and tested on observation. Bottom: classifiers trained on observation and tested on intention. Each heatmap shows decoding accuracy across all combinations of training (y-axis) and testing (x-axis) time bins. Significant train–test time bin pairs (permutation test, $p < 0.05$) are marked with dots. MC shows significant cross-decoding only in the Observation $\rightarrow$ Intention direction, whereas SPL exhibits significant decoding in both directions. **(C)** Cross-format decoding strength for action type. Green bars indicate Intention $\rightarrow$ Observation decoding and blue bars indicate Observation $\rightarrow$ Intention decoding. SPL shows bidirectional cross-format decoding (weak for the Observation $\rightarrow$ Intention direction) whereas MC exhibits only unidirectional decoding in the Observation $\rightarrow$ Intention direction.

Figure S10: Additional PCA Trajectory Comparisons Across Task Variables and Control Analysis with All Units, Related to Figure 5.

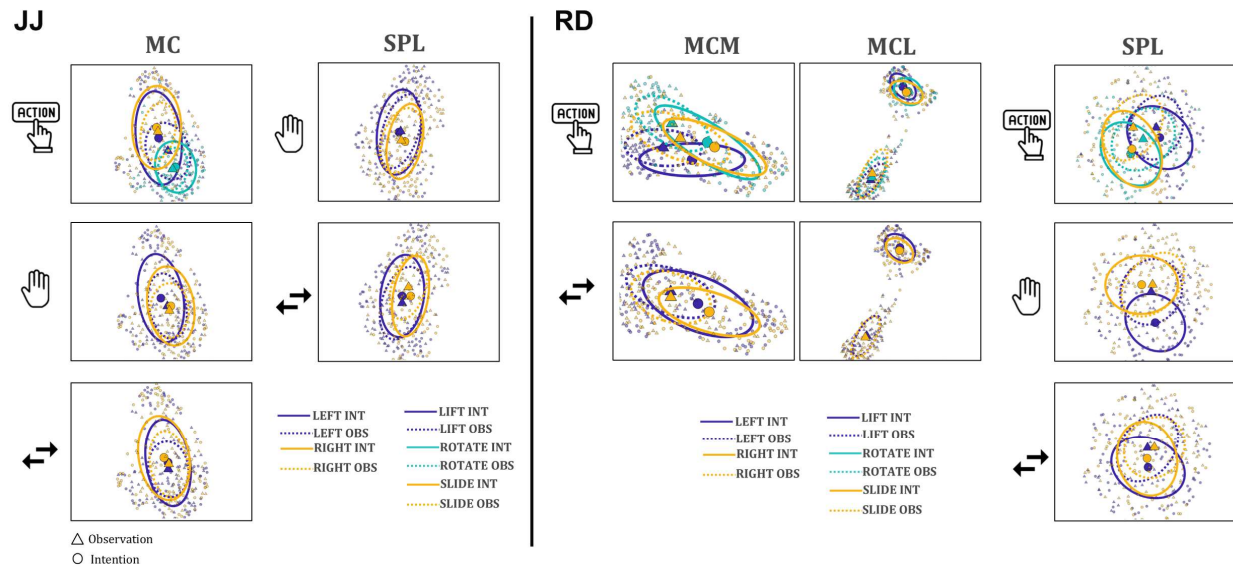


(A) PCA trajectories for hand (top) and direction (bottom) in JJ, plotted for MC and SPL. Each 3D plot shows the first three principal components. Within each region, the left panel (green outline) shows intention and the right panel (blue outline) shows observation. Trajectories in MC differ substantially between formats, while SPL trajectories appear very similar for both task variables.

(B) Same analysis for RD. Shown are action trajectories for MCM and MCL (top row), hand for SPL (top right), and direction for all three arrays (bottom row). Consistent with JJ, trajectories in MCM and MCL differ across formats, while SPL exhibits similar trajectories between intention and observation for both hand and direction.

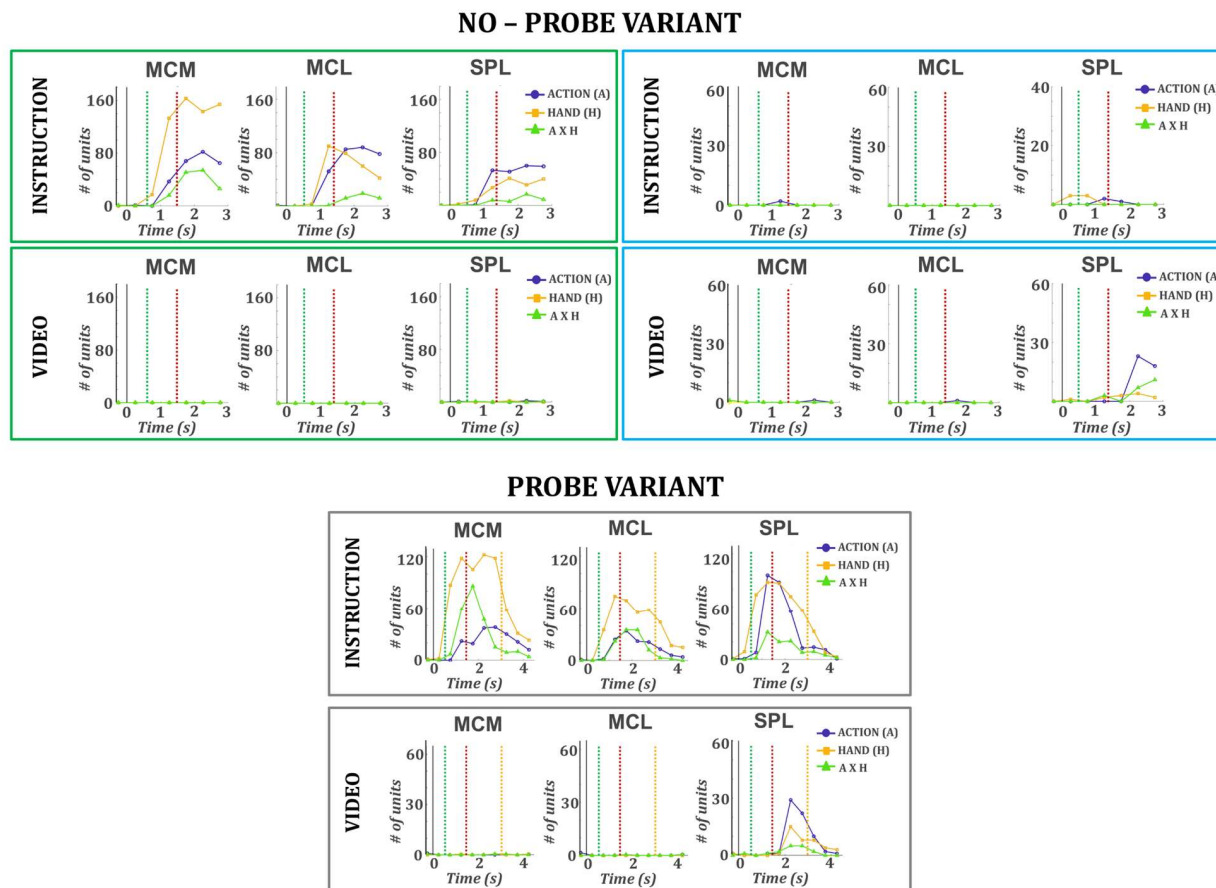
(C) Alignment analysis for action, hand, and direction when including all recorded units rather than just task-relevant ones. Results are consistent with the main analysis: SPL shows consistently low alignment distances across all variables ($d < 0.2$), confirming that the high similarity between formats in SPL is not dependent on unit selection.

Figure S11: UMAP Embeddings for All Task Variables and Cortical Arrays, Related to Figure 5.



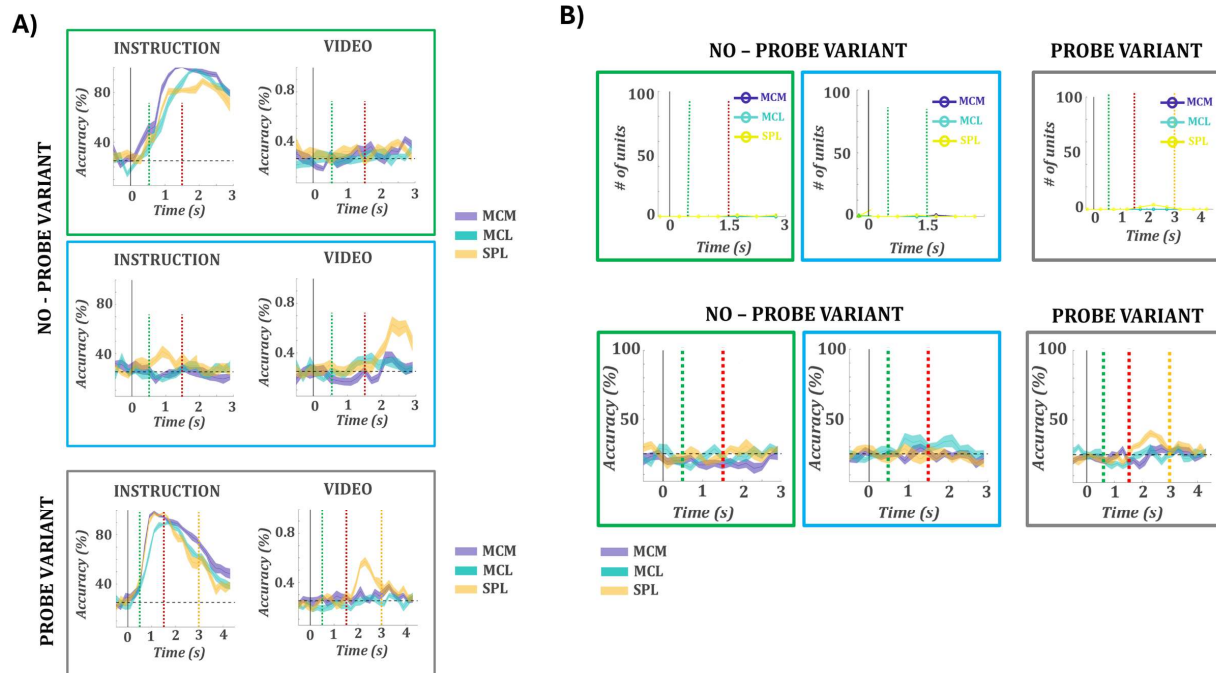
UMAP projections of neural population activity for participants JJ (left) and RD (right), shown separately for each task variable. Each point represents a trial, with triangles for observation and circles for intention. Ellipses represent 2D distributions per condition (solid for intention, dashed for observation), and colors indicate specific task conditions (e.g., lift, rotate, slide; left/right hand; direction). In MC/JJ, rotation actions show the greatest overlap between intention and observation formats. In MCL/RD, condition clusters are clearly segregated between formats across all variables. SPL exhibits relatively consistent overlap: for the majority of conditions, the centroids of intention and observation responses are close together, suggesting shared structure across formats.

Figure S12: Neural Tuning for Instructed and Observed Actions, Related to Figure 7.



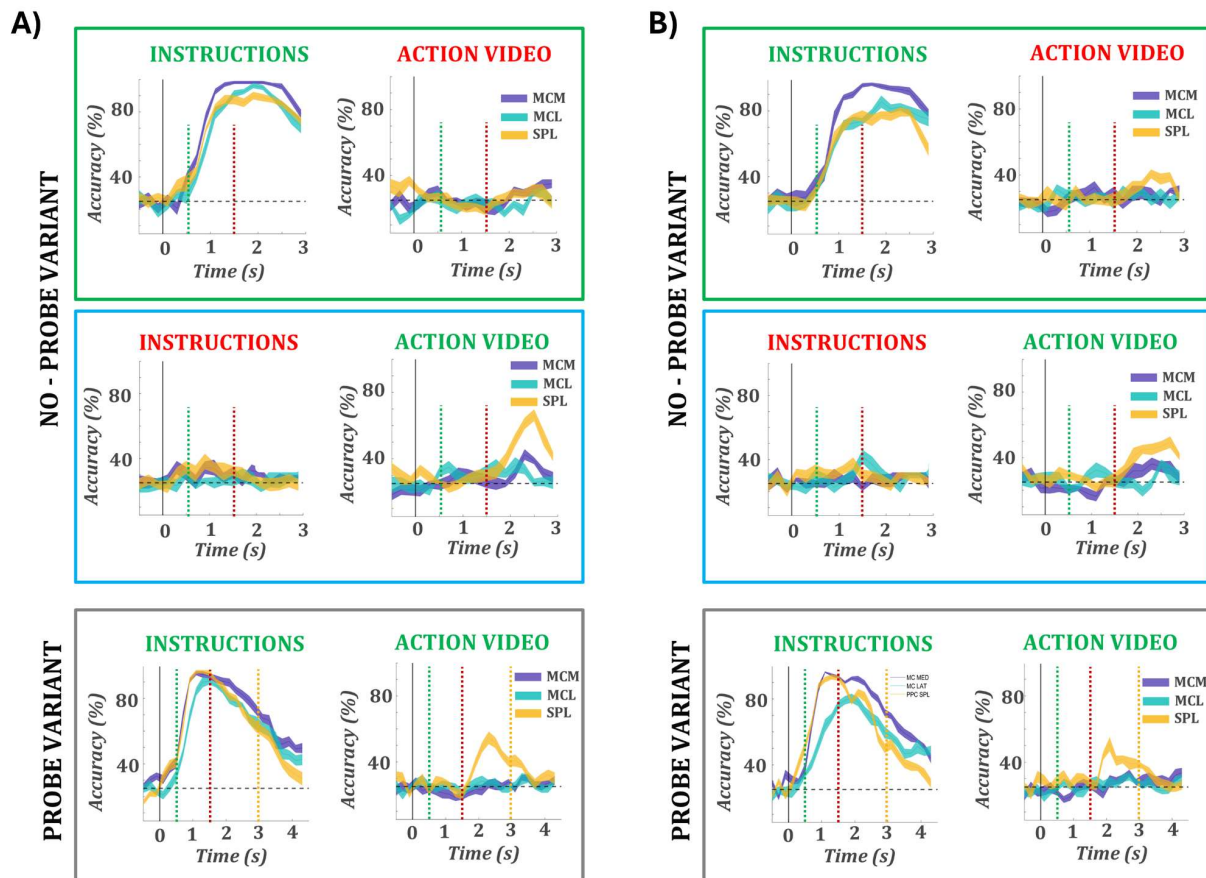
Time-resolved tuning. Green outline: no probe variant – instruction block; blue outline: no probe variant – observation block; gray outline: probe variant. Each plot shows the number of units tuned to the instructed (top row) or video (bottom row) action, hand, or their interaction in MCM, MCL, and SPL. Vertical lines denote key task events: hand cue (black), action cue (green), video onset (red), and probe onset (orange).

Figure S13: Control Analysis for Congruency in the Dissociation Tasks, Related to Figure 7.



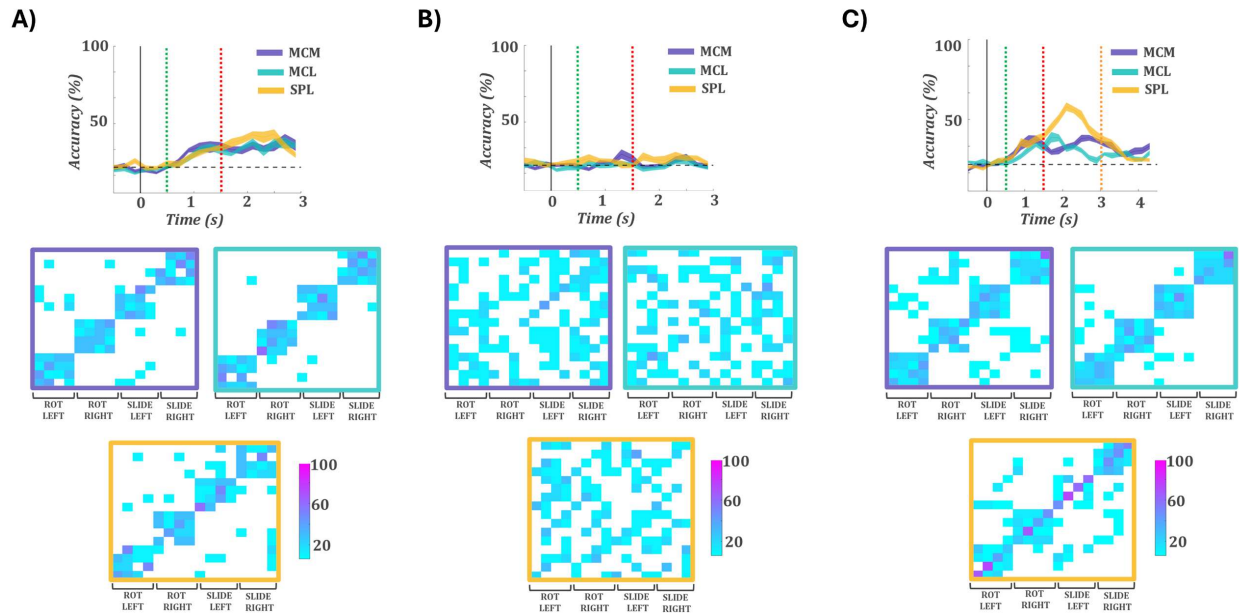
(A) Decoding analysis including only incongruent trials from both dissociation tasks. Decoding performance for the instructed (left) and video (right) actions is shown for each variant: no-probe instruction block (green outline), no-probe observation block (blue outline), and probe variant (gray outline). Colored lines indicate mean decoding accuracy over time for MCM (purple), MCL (teal), and SPL (orange); shaded areas show $\pm$ SEM across cross-validation folds. The dashed horizontal line marks chance level. Vertical lines denote key task events: hand cue (black), action cue (green), video onset (red), and probe onset (orange). **(B)** Top row: number of tuned units per array for conflict type. Bottom row: decoding accuracy for conflict type. Green outline: no-probe instruction block; blue outline: no-probe observation block; gray outline: probe variant. Shaded areas indicate $\pm$ SEM across folds; dashed horizontal line marks chance level; vertical lines as in (A).

Figure S14: Decoding in the Dissociation Task Shown Separately for MUA and SUA, Related to Figure 7.



(A) Decoding accuracy in the dissociation task using MUA. **(B)** Same analysis using SUA. Decoding performance for the instructed action (left) and video action (right) is shown for each variant: no-probe instruction block (green outline), no-probe observation block (blue outline), and probe variant (gray outline). Colored lines indicate mean decoding accuracy over time for MCM (purple), MCL (teal), and SPL (orange); shaded areas represent $\pm$ SEM across cross-validation folds. The dashed horizontal line marks chance level. Vertical lines denote key task events: hand cue (black), action cue (green), video onset (red), and probe onset (orange). Across signal types, decoding patterns are consistent.

Figure S15: Decoding of All 16 Dissociation Task Conditions, Related to Figure 7.



(A–C) Time-resolved decoding of the full set of 16 dissociation conditions (2 actions × 2 hands × 4 conflict types) in MCM (purple), MCL (teal), and SPL (orange), shown for the no-probe intention block **(A)**, no-probe observation block **(B)**, and the probe task **(C)**. Line plots show mean classification accuracy over time; shaded areas represent $\pm$ SEM across cross-validation folds. The dashed horizontal line indicates chance level ($1/16 = 6.25\%$). Vertical lines mark the hand cue (black), action cue (green), video onset (red), and probe onset (orange). Confusion matrices below each plot show classifier performance at the peak decoding time for each region.

Tables:

Table S1: Experimental Summary: Sessions, Trial Counts, and Recorded Units Per Region, Related to STAR Methods.

SUBJECT	SESSION DATE	REPS/CLASS	MC	SPL	MCM	MCL
JJ	02/06/25	12	76	72	-	-
JJ	02/13/25	12	75	70	-	-
JJ	02/20/25	12	91	62	-	-
JJ	04/03/25	12	88	62	-	-
JJ	04/10/25	12	91	60	-	-
TOTAL:			421	326		
RD	02/06/25	12	-	110	57	72
RD	02/13/25	12	-	79	79	86
RD	02/20/25	12	-	82	78	84
RD	02/27/25	12	-	86	80	100
RD	04/03/25	12	-	92	73	60
RD	04/10/25	12	-	83	74	77
TOTAL:				532	441	479

Table S2: Explained Variance of Neural Trajectories Across Regions and Formats, Related to Figure 5.

ACTION	Intention (%)	Observation (%)	Total (%)
JJ/MC	84.07	47.7	76.2
JJ/SPL	79.2	67.1	73.8
RD/MCM	68.6	59.2	66.4
RD/MCL	58.4	48.9	61.6
RD/SPL	78.2	65.2	72.8
HAND	Intention (%)	Observation (%)	Total (%)
JJ/MC	81.7	31.3	72.6
JJ/SPL	84.6	69.5	77.6
RD/MCM	93.6	64.6	88.4
RD/MCL	79.9	62.4	78.5
RD/SPL	85.7	71	80
DIRECTION	Intention (%)	Observation (%)	Total (%)
JJ/MC	79.6	34.3	71
JJ/SPL	82.5	71.8	77.1
RD/MCM	71.8	61.5	71.7
RD/MCL	79.2	59.6	77.2
RD/SPL	83.4	72.8	80

Table S3: Dissociation Task: Sessions, Trial Counts, and Recorded Units Per Region, Related to STAR Methods.

SUBJECT	SESSION DATE	REPS/CLASS	SPL	MCM	MCL
RD	05/09/25	10	88	55	41
RD	05/15/25	10	94	91	87
RD	05/23/25	10	92	77	70
RD	05/29/25	10	95	81	71
RD	06/05/25	10	80	98	84
RD	06/06/25	10	87	69	55
TOTAL:			536	471	408

Table S4: Dissociation Task with Probe: Sessions, Trial Counts, and Recorded Units Per Region, Related to STAR Methods.

SUBJECT	SESSION DATE	REPS/CLASS	SPL	MCM	MCL
RD	06/12/25	15	102	87	90
RD	06/24/25	15	94	83	50
RD	07/03/25	15	116	85	86
RD	07/23/25	15	82	76	72
RD	07/24/25	15	90	83	88
TOTAL:			484	414	386