Large-Scale Circuit Configuration for Flexible Sensory-Motor Decisions

R.L. van den Brink[1]* K. Hagena[1], N. Wilming[1], P.R. Murphy[1,2], C. Büchel[3], & T.H. Donner[1,4]*

[1]Section Computational Cognitive Neuroscience, Department of Neurophysiology and Pathophysiology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany.
[2]Trinity College Institute of Neuroscience and School of Psychology, Trinity College Dublin, Dublin, Ireland.
[3]Institute for Systems Neuroscience, University Medical Center Hamburg-Eppendorf, Hamburg, Germany
[4]Bernstein Center for Computational Neuroscience, Charité Universitätsmedizin, Berlin, Germany.

* To whom correspondence should be addressed: r.van-den-brink@uke.de, t.donner@uke.de

Abstract

Humans and non-human primates can acquire, and rapidly switch between, arbitrary rules that govern the mapping from sensation to action. It has remained unknown if and how the brain configures large-scale sensory-motor circuits to establish such flexible information flow. Here, we developed an approach that elucidates the dynamic configuration of task-specific sensory-motor circuits in humans. Participants switched between arbitrary mapping rules for reporting visual orientation judgments during fMRI. Rule switches were either signaled explicitly or inferred by the participants from ambiguous cues, and we used behavioral modeling to reconstruct the time course of their belief about the active rule. In both contexts, patterns of correlations between ongoing fluctuations in feature-specific activity across visual and action-related brain regions tracked participants' belief about the active rule. These rule-specific, intrinsic correlation patterns broke down on error trials and predicted individuals' model-inferred internal noise. Our findings indicate that internal beliefs about task state are instantiated in specific large-scale patterns of selective, correlated neural variability.

Introduction

Perceptual decisions transform an internal representation of the state of the environment into motor output in a flexible, context-dependent manner (Mante et al., 2013; Shadlen and Kiani, 2013). Substantial progress has been made in uncovering the neural computations that underlie such transformation when the task-relevant mapping from sensation to action is stable. One major insight has been that the transformation often entails the accumulation of fluctuating sensory input into an evolving decision variable encoded in action-related brain regions (Bogacz et al., 2006; Gold and Shadlen, 2007; Donner et al., 2009; Hanks et al., 2015; Wilming et al., 2020; Murphy et al., 2021).

In the laboratory, as well as in real life, perceptual decisions often entail an arbitrary association of specific features of the sensory input with specific features of the motor output (Sakai, 2008; Shadlen and Kiani, 2013). Consider the mapping rule that governs a typical perceptual choice task: report vertical or horizontal stimulus orientations of a visual grating by button presses with the left or right hand, respectively (Fig. 1a; Mapping rule 1). These sensory-motor associations are arbitrary because there are no fixed anatomical pathways from sensory to motor brain regions that would predispose toward these associations over their complements (Mapping rule 2 in Fig. 1a). Yet, primate behavior precisely tracks switches between distinct sensory-motor mapping rules (Gold and Shadlen, 2003; Tsetsos et al., 2015; Purcell and Kiani, 2016). Thus, the brain can rapidly configure distinct task-specific circuits that support the corresponding feature-specific information flow from sensation to action (Shadlen & Kiani, 2013).

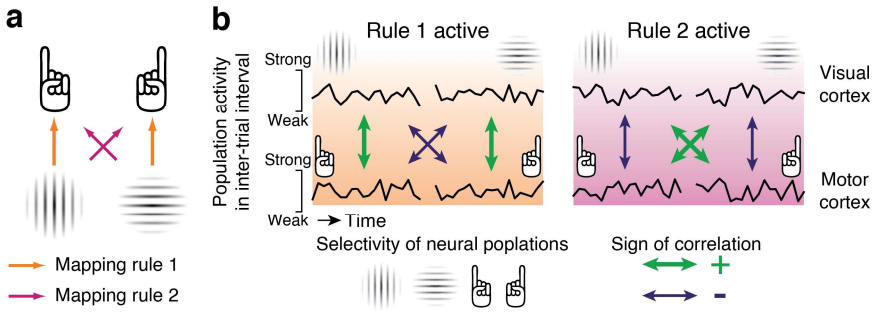


Fig. 1. Sensory-motor mapping rule, task-relevant functional networks, and correlated neural variability. a. Example sensory-motor mapping rule in a visual orientation discrimination task. Arrows indicate information flow from sensory input to motor output. **b.** Predicted patterns of correlated intrinsic variability between orientation-selective populations of visual cortical neurons (top) and action-selective populations of motor cortical neurons (bottom). Time courses indicate ongoing (i.e., not task-evoked) fluctuations in activity pooled across each population of neurons occurring during the intervals between trials of the primary orientation discrimination task. The structure of these intrinsic correlations reflects the structure of the large-scale circuits implementing the task, rather than the rule-dependent coupling of stimulus- and action-selective population responses during task execution.

How does the brain configure the neural circuits for flexible sensory-motor decisions, particularly in natural settings, in which the mapping rules may change unpredictably? Previous research into these questions has focused on the encoding of rule information within individual, associative regions of the brain, such as prefrontal cortex (Wallis et al., 2001; Stoet and Snyder, 2004; Haynes et al., 2007; Sakai, 2008; Bode and Haynes, 2009; Bennur and Gold, 2011; Zhang et al., 2013; Cole et al., 2016). It has remained unknown how rule information shapes the information flow between brain regions that encode the stimuli and motor actions for the primary task in a feature-specific manner, as is required for establishing the mapping rules at hand (Fig. 1a).

One possibility is that neural populations in association cortex act as local ‘switches’ that are activated by the stimulus responses in sensory cortical regions, and then route the sensory signals to action-related areas in a rule-dependent fashion (Cocuzza et al., 2020; Kikumoto and Mayr, 2020; Ito et al., 2022), without altering the connectivity between the sensory and action-related regions. Alternatively, the rule may directly shape the configuration of functional connections between sensory and action-related regions in a selective and dynamic fashion. Such a dynamic modulation of sensory-motor connectivity could be mediated by short-term plasticity (Fusi et al., 2007) and/or selective top-down projections from rule-coding brain regions (Miller and Cohen, 2001). The latter large-scale circuit configuration scheme has not yet been tested because inferring task-specific functional networks of stimulus-, action-, and rule-selective neural populations has been challenging.

We developed an approach to overcome this challenge and elucidate the dynamic configuration of large-scale circuits for flexible sensory-motor decisions. We asked human participants to perform a visual orientation discrimination task with varying sensory-motor mapping rules. Under the above distributed reconfiguration scheme, rule information should sculpt the feature-selective connectivity of sensory-motor networks. This, in turn, should translate into selective patterns of co-fluctuations of ongoing (i.e., intrinsically generated rather than evoked by stimuli or actions) activity (Fig. 1b) (Gerstein and Perkel, 1969; Büchel and Friston, 2000; Vincent et al., 2007). We tested this prediction by analyzing the structure of intrinsic correlations between stimulus- and action-selective patterns of ongoing fMRI activity expressed within sensory and motor brain areas. We further reasoned that the dynamics of these task-specific correlations should reflect the evolution of the agent’s internal belief about the appropriate sensory-motor association. Behavioral modeling enabled us to reconstruct participants’ hidden beliefs about the active rule at any given moment. The patterns of correlated feature-selective neural variability reliably tracked those beliefs and predicted participants’ behavioral performance, linking the structure of correlated variability to cognitive computation and establishing its functional significance. Taken together, our results indicate that the dynamics of spontaneous correlated variability of stimulus- and action-selective population codes across the brain tracks the context-dependent configuration of large-scale circuits for flexible sensory-motor decisions.

Results

We alternated the stimulus-response mapping rule (SR-rule) of a primary perceptual choice task (Fig. 1a). Under each rule, signals from vertical and horizontal orientation columns in primary visual cortex (V1) (Bonhoeffer and Grinvald, 1991) need to be routed to distinct neural populations in primary motor cortex (M1) that control left- or right-hand movements, respectively (Yousry et al., 1997). In one set of fMRI runs (see Materials and Methods), the SR-rule for the upcoming trials was explicitly instructed by visual cues (SR-rule cue; Fig. 2a), which defined mini-blocks of two trials each (Fig. 2b). We call this setting ‘instructed rule’. Each trial of the primary task entailed the presentation of a large, full-contrast grating of either horizontal or vertical orientation, followed by the participants’ report of their orientation judgment. Grating stimuli were presented at long and variable intervals (Fig. 2a).

Participants were able to switch rapidly and consistently between the rules (Fig. 2c), reaching choice fractions near-fully consistent with the active rule on the first trial after the switch. Consequently, all participants’ overall accuracy level was close to ceiling (Fig. 2d). This high performance was expected because the task entailed no uncertainty about the active rule nor about the primary perceptual judgment (maximal contrast and orientation difference between

discriminanda). This setting yielded clear expectations about the dynamics of putative changes in patterns of correlated variability during rule switching (Fig. 1b).

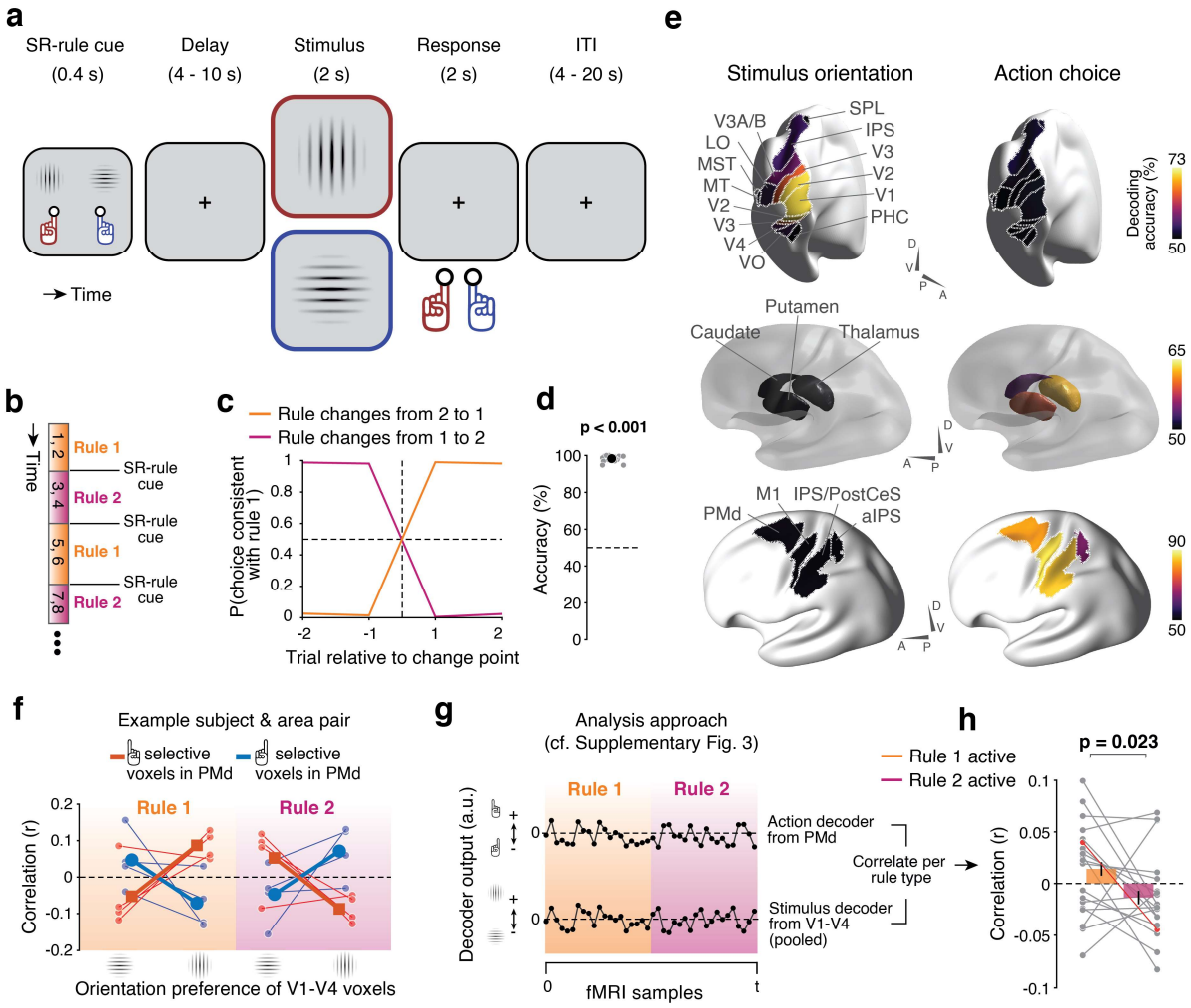


Fig. 2. Instructed rule task and assessment of feature-selective correlated variability. a, b. Behavioral task protocol in instructed rule task. **a.** Sequence of events following an SR-rule cue into the first inter-trial interval ITI of a block. See main text for details. **b.** Block structure of the task. Blocks indicate period during which one of the rules is active. Numbers in the blocks indicate trials. **c.** Group average time course of choice fractions locked to instructed rule changes. Shaded area, SEM. For many data points, this area is smaller than the line for the mean. **d.** Overall behavioral accuracy. Gray dots, individual participants; large black dot, group-average. **e.** Classification of stimulus orientation and action by decoders trained and tested on evoked responses during primary task trials. Decoding performance is evaluated for 19 pre-selected cortical and subcortical regions (pooled across hemispheres; see Materials and Methods). **f.** Selective co-fluctuation pattern for feature-selective voxel groups from example participant and region pair (V1-hV4 vs. PMd). See main text for details. Thin lines, individual task runs (N=4); thick lines, average. **g.** Final step of analysis approach. The graded outputs of stimulus and action decoders from above regions during ongoing activity are selectively correlated for the different rule blocks. See main text for details. **h.** Decoder co-fluctuations during the two rules. Thin lines, individual participants (N=19); red line, participant from panel f; bars, group average; error bars, within-participant SEM (Materials and Methods).

Rule encoding in correlated neural variability

We reasoned that if the sensory-motor mapping rule selectively sculpts the sensory-motor network that governs the primary choice task, the active rule should be reflected in feature-

specific co-fluctuations between populations of neurons in sensory and motor cortices. When participants' choices are consistent with the rule, stimulus- and action-selective evoked responses during the execution of the primary task are bound to correlate in a rule-specific fashion. Critically, our hypothesis predicts that such rule-specific correlations are also expressed in the ongoing activity fluctuations over and above the evoked responses (Fig. 1b): Maintenance of rule 1 in Fig. 1a should be reflected in co-fluctuations between vertical-preferring neural populations in visual cortex, and left-hand movement-encoding neurons in motor cortex. Horizontal-preferring visual cortical populations should similarly co-fluctuate with right-hand movement-encoding motor cortical neurons. Conversely, the other two pairs of neural populations should be decoupled or even anti-correlated (Fig. 1b, orange). When the rule then switches to rule 2, this selective pattern of correlated variability should flip (Fig. 1b, pink).

To provide an intuition of our approach, we first analyzed correlations between the fluctuations of fMRI activity of feature-selective subsets of voxels (Fig. 2f). We focused on two sets of regions implicated in orientation perception (Reynolds et al., 2000; Haynes and Rees, 2005; Kamitani and Tong, 2005) and action planning or execution (Picard and Strick, 2001; Cisek and Kalaska, 2005; Kamitani and Tong, 2005; Murphy et al., 2021), respectively: early visual cortex (V1-V4) and dorsal premotor cortex (PMd). The multi-voxel patterns of evoked fMRI responses in these regions during the trials of the primary decision task reliably encoded the stimulus orientation (for V1-V4) or the participants' action choice (for PMd, Fig. 2e and Supplementary Fig. 1).

Rule-specific correlations of stimulus- and action-selective neural populations (subsets of voxels) are bound to occur even without reconfiguration of the underlying network because participants acted in accordance with the active rule in the majority of trials (Fig. 2c). Critically, we isolated patterns of ongoing fluctuations of activity in each area (Arieli et al., 1996; Kenet et al., 2003; Fox et al., 2006) by removing the stimulus- or action-evoked responses (see Materials and Methods). This procedure effectively removed evoked responses, because it left no stimulus- or action-selective activity patterns systematically aligned in time with the trials of the primary decision task (Supplementary Fig. 2). For each area, we selected subgroups of voxels that responded preferentially to vertical or horizontal gratings (for V1-V4) or to left- or right-hand choices (PMd). We then correlated the ongoing (i.e., residual) activity fluctuations of all four possible pairs of voxel sub-groups (Materials and Methods). In the example participant in Fig. 2f, the sign of these correlations was in accordance with the prediction from Fig. 1b during rule 1: positive for vertical and left-hand coding voxels and for horizontal and right-hand coding voxels, and negative for the other two pairs (Fig. 2f, orange). This pattern flipped for rule 2 (Fig. 2f, pink).

To quantify these correlations in a more compact fashion across the group of participants and all pairs of areas (Fig. 2g, Supplementary Fig. 3), we trained stimulus- and action decoders on the response patters evoked specifically during trials of the primary task (Fig. 2e). We then applied these decoders to the fluctuations of ongoing (i.e., residual) activity patterns that occurred throughout the recording (Materials and Methods and Supplementary Fig. 3), thus estimating the spontaneous fluctuations of feature-selective population activity within each brain region. For example, in V1, the resulting time series reflected the extent to which multi-voxel patterns spontaneously tended towards vertical or horizontal orientations (Kenet et al., 2003) (Supplementary Fig. 3). We then correlated stimulus and action decoder outputs, separately for intervals corresponding to the two rules, shifted by the hemodynamic delay (Materials and Methods, Fig. 2g).

Given the way we constructed the decoders for stimulus (positive prediction for vertical and negative prediction for horizontal) and action (positive prediction for left hand and negative prediction for right hand; Supplementary Fig. 3 and Materials and Methods), our hypothesis (Fig. 1b) predicted that the correlation of decoder outputs in the ongoing activity should be positive when rule 1 was active and negative when rule 2 was active (Fig. 2g). This is what we found for the output of the V1-V4 stimulus decoder and the PMd action decoder, both in our example participant (see Fig. 2h, red line) as well as in the whole group (Fig. 2h, gray lines and colored bars). Thus, the feature-selective co-fluctuations of ongoing activity in early visual cortex and PMd flipped sign depending on the active rule.

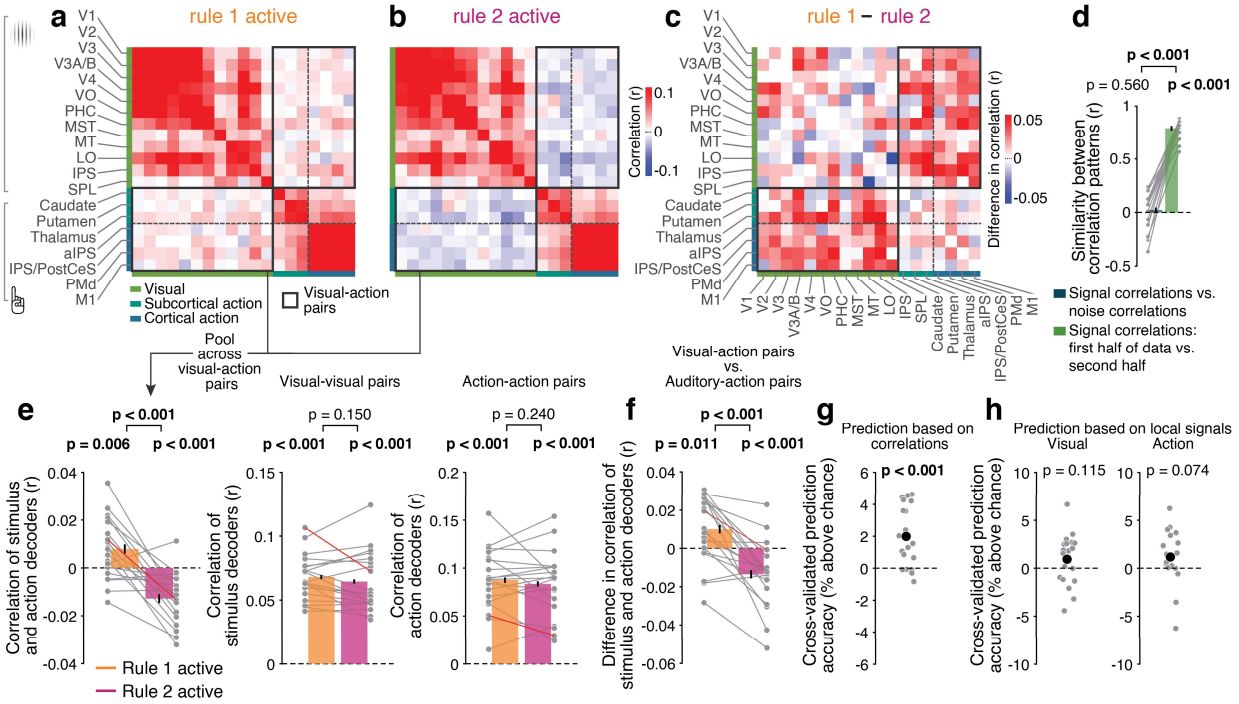


Fig. 3. Rule encoding in correlations of stimulus and action codes across sensory-motor networks. a, b. Patterns of correlation between all pairs of decoders for stimulus orientation (visual cortical areas) and for action (subcortical and action-related cortical areas). The gray rectangle contains all pairs of stimulus and action decoders. **a.** Correlations during rule 1. **b.** Correlations during rule 2. **c.** Difference matrix between rule 1 and rule 2. **d.** Similarity of correlation patterns (across region pairs) of decoder outputs for trial-evoked responses ('signal correlations'; Materials and Methods) and for the intrinsic fluctuations ('noise correlations', as used in main analyses). Gray lines, individual participants; bars, group average; error bars, within-participant SEM (Materials and Methods). **e.** Left: Correlations from panels a and b, collapsed across all pairs in gray rectangle. Right: Correlations from panels a and b, collapsed across all visual-visual pairs and action-action pairs (excluding values on diagonals). Gray lines, individual participants; red line, participant from panel e; bars, group average; error bars, within-participant SEM (Materials and Methods). **f.** Correlations from panel d (left), relative to the correlations of the graded outputs of a stimulus orientation decoder that was trained on auditory cortex (Supplementary Fig. 6.), and action decoder. **g.** Cross-validated accuracy of rule prediction from time-resolved coupling between stimulus and action decoders. Gray dots, individual participants; black dot, group mean. **h.** As e, but rule predictions based on either stimulus decoder outputs or action decoder outputs, rather than their correlations. N=19 for all panels.

This pattern was evident for most individual pairs of stimulus- and action-encoding regions (Fig. 3): The correlation between stimulus and action decoder outputs (gray rectangles in Fig. 3a-c), was positive during rule 1 for most region pairs (Fig. 3a) and negative under rule 2 for most region pairs (Fig. 3b), leading to a clear difference between pairs (Fig. 3c). This difference was consistent across stimulus-action decoder pairs, but not for pairs of visual decoders or for pairs

of action decoders (compare cells inside vs. outside of gray rectangle). Correspondingly, collapsing the correlations across all pairs of stimulus-encoding and action-encoding regions (cells inside gray rectangle) yielded a robust positive correlation for rule 1 and negative correlation for rule 2, with a clear difference between the two (Fig. 3e).

The intrinsic decoder output correlations ('noise correlations') exhibited a spatial pattern across region pairs (gray rectangle in Fig. 3c) that was distinct from the correlation pattern of the feature-specific evoked responses ('signal correlations'; Fig. 3d and Supplementary Fig. 4; Materials and Methods). Furthermore, the observed rule-related pattern of decoder correlations was present even when restricting the analysis to segments of data in the inter-trial intervals (Supplementary Fig. 5a,b). Together with our effective removal of evoked responses (Supplementary Fig. 2), these observations are inconsistent with the notion that the measured correlations were driven by trial-evoked responses.

Such a pattern of correlations was not evident when collapsing correlations across pairs of visual cortical areas or pairs of action-encoding areas alone (Fig. 3e, middle, right). Furthermore, when substituting the visual cortical regions for a set of control regions centered on primary auditory cortex, no significant (and substantially weaker) correlation between stimulus and action decoder outputs was present (Fig. 3f; Supplementary Fig. 6). These analyses highlight the specificity of the result presented in Fig. 3e for the co-fluctuations of population codes for stimulus and action.

Having established the specificity of the rule-related patterns of intrinsic correlations for spontaneous co-fluctuations of stimulus and action codes, we next tested if these patterns, measured in a given segment of data, robustly predicted the currently active rule. We trained a logistic regression model to make a prediction about the active rule at individual time points in held-out sections of data. The model contained two distinct sets of regressors: (i) the correlations of stimulus and action decoder outputs from the visual-action pairs of regions; and (ii) the local stimulus and action decoder outputs themselves (Materials and Methods). We computed a time-variant estimate of the co-fluctuation of stimulus and action decoder outputs (Materials and Methods), and trained and tested the model in a 10-fold cross-validation procedure (Materials and methods). The overall prediction accuracy of this model was above chance for most individual participants (group mean: 1.92 % above chance; range: -1.45 – 6.65). Critically, the accuracy of predictions based only on the correlation component was also above chance for most individuals and the group average (Fig. 3g). By contrast, rule prediction was not possible based on the outputs of either the stimulus or the action decoders alone (Fig. 3h).

In sum, the active rule was encoded in patterns of spontaneous co-fluctuations of stimulus and action codes distributed across the sensory-motor network of brain areas. These selective and dynamic patterns of correlated variability could be interpreted as a signature of the ongoing circuit configuration for the primary perceptual choice task (see Discussion).

Inferring volatile sensory-motor mapping rules under uncertainty

Because in our previous task the active rule was explicitly instructed without ambiguity (Fig. 2a, b), we assumed that participants' internal belief about the appropriate sensory-motor association had close-to-maximal certainty. In many real-world situations, however, this belief is

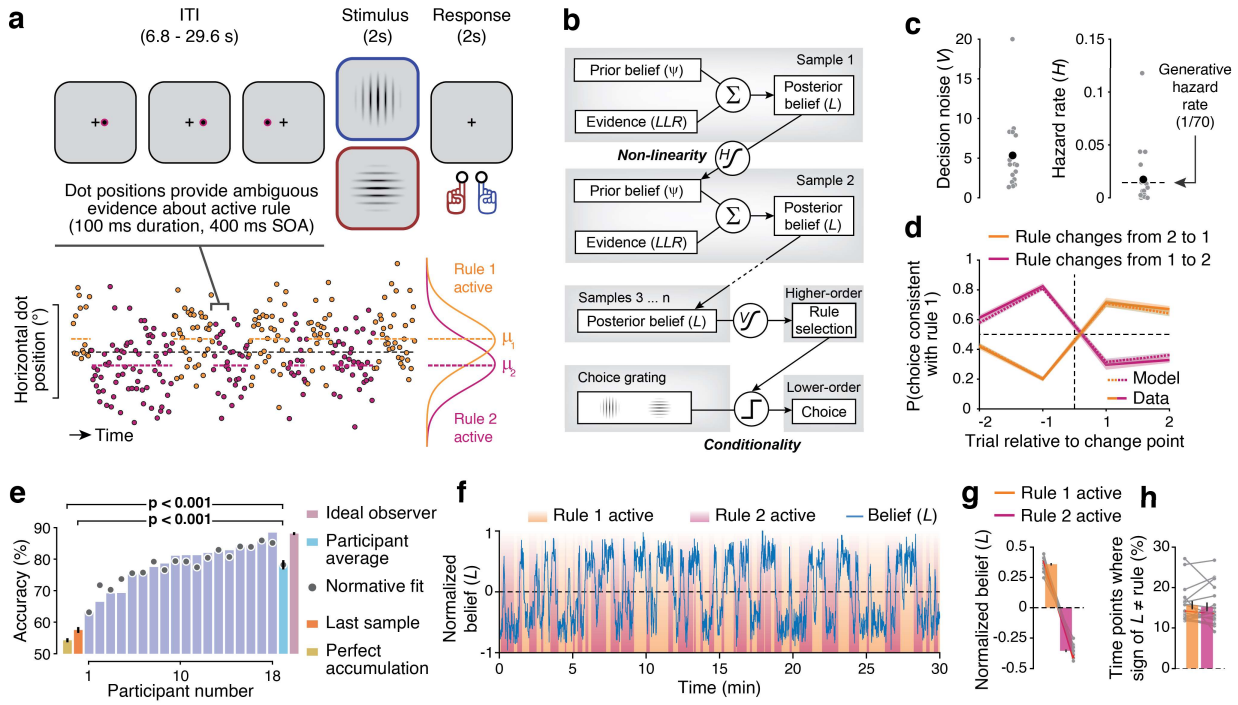


Fig. 4. Tracking the evolution of belief about volatile sensory-motor mapping rule. a. Sequence of events in an example trial of the primary perceptual choice task. Noisy samples of sensory evidence (horizontal dot positions) were presented every 400 ms during the ITI. Each sample was drawn from one of two generative distributions shown at the bottom right, producing the example stream on the bottom left. The generative distribution also governed the active rule and could switch unpredictably between any two samples with a probability of 1/70. ITI, inter-trial interval. **b.** Schematic of normative model for rule inference and selection, coupled with the primary orientation discrimination task. See main text for details. **c.** Parameters of the best fitting model. Gray dots, individual participants; black dot, group average. **d.** Fraction of choices consistent with rule 1 surrounding changes in true (generative) rule, as produced by participants (solid lines) and best-fitting normative model (dashed lines). Shaded area, SEM. **e.** Overall accuracy of participants, best-fitting normative model, and alternative models. Error bars, SEM. **f.** Example time course of posterior belief (L) estimated from best-fitting model for one participant during 30 selected minutes of an experimental session. **g.** Belief averaged across data segments corresponding to true (generative) rules. Gray lines, individual participants; red line, participant from panel f; bars, group average; error bars, within-participant SEM (Materials and Methods). **h.** Percentage of time points at which the sign of belief (L) was inconsistent with the active rule. Gray lines, individual participants; bars, group average; red line, participant from panel f; error bars, within-participant SEM (Materials and Methods). N=18 for all panels.

subject to uncertainty; it needs to be learned from noisy and incomplete information (Miller and Cohen, 2001; Durstewitz et al., 2010). We thus developed a new variant of the task to probe the evolving formation of an internal belief about the appropriate sensory-motor mapping under uncertainty. In this task, the sensory-motor mapping rule was volatile (i.e., could undergo hidden changes) and had to be inferred from noisy sensory evidence presented in the inter-trial intervals (Fig. 4a). The evidence samples were the positions of small dots presented at a high rate in a narrow range left or right from the central fixation mark (Materials and Methods). Each dot position was drawn from one of two overlapping Gaussian distributions, which corresponded to the active rule at that time. The active rule (and thus the generative distribution) could switch from one sample to the next with a low probability (hazard rate; 1/70). The trials of the primary perceptual choice task were the same as before (orientation discrimination judgement), and randomly interspersed in the evidence stream, again with long and variable intervals (Fig. 4a).

This task, called ‘inferred rule’ in the following, required participants to continuously select the currently active rule by accumulating the noisy evidence (higher-order decision) and apply the

selected rule to report their decision for orientation judgment (lower-order decision). Because the generative state for the higher-order decision could undergo unpredictable and hidden state changes, perfect (i.e., lossless) accumulation of all the evidence is suboptimal, and optimal performance requires accumulating the evidence in an adaptive fashion that strikes a balance between stable evidence accumulation and sensitivity to change points (Glaze et al., 2015; Piet et al., 2018; Murphy et al., 2021). We asked the same participants to perform this task, which they did reasonably well (Fig. 4 d,e), also reliably switching between the rules when required by the task (Fig. 4d). Note that even the ideal observer using the normative strategy for this task (Fig. 4b; see below) and equipped with exact knowledge of the environmental hazard rate would only perform around 88% correct on this task when given the exact same evidence sequences as presented to the participants (Fig. 4e, purple).

We fitted participants' behavior with a normative model of the adaptive evidence accumulation process. The model casts this process as dynamic belief updating (Fig. 4b) (Glaze et al., 2015): Each sample of sensory evidence (expressed as log likelihood ratio, *LLR* of support for one vs. the other rule) was combined with a prior belief (also expressed as log-odds, Ψ) to form a posterior belief (L) for one over the other rule. The posterior was passed on to the next updating step to become the new prior that was then combined with the next evidence sample, and so on. Critically, the transformation of posterior into subsequent prior was non-linear, with a shape controlled by the agent's estimate of the environmental hazard rate (subjective hazard rate H; Fig. 4c). This feature rendered the model adaptive, sensitive to different levels of environmental volatility, in line with observations of human behavior across different experimental settings (Glaze et al., 2015; Murphy et al., 2021; Weiss et al., 2021). In our version of the model, this adaptive inference process for the higher-order decision (rule selection) was coupled with a lower-order decision about the grating orientation that governed the behavioral choice (Fig. 4b). Our model assumed that noise corrupted the higher-order rule selection (V, Materials and Methods), but not the lower-order orientation judgment (Fig. 4b). We based this assumption on the near-perfect performance of the same participants in the instructed rule task (Fig. 2d), indicating that the orientation judgment was too easy for noise at the level of the lower-order decision to affect performance. We fitted the model to the participants' choice behavior with H and V as the only free parameters.

The model fit the data well for most participants (Fig. 4d,e), and it matched the data better than alternative, heuristic strategies for all participants (Fig. 4e). A model variant that perfectly accumulated all samples across the run without information loss, as well as a model that selected the rule using only the last-seen sample, both resulted in significantly poorer performance than both the fitted normative model and the participants' behavior (Fig. 4d; $p < 0.001$ in all comparisons). Furthermore, similar to findings from previous work (McGuire et al., 2014), the computational variables derived from the model, specifically the magnitude (absolute value) of the posterior belief, the magnitude of sensory evidence, and change point probability (the likelihood that a change in rule has just occurred given the sensory evidence and belief) (Murphy et al., 2021), all co-varied with fMRI activity across widespread cortical regions (Supplementary Fig. 7g-i).

As expected, the time-varying posterior belief (L) of the fitted model tracked the active rule in all participants (Fig. 4e,f). Yet, the active rule and belief also deviated on a substantial fraction of time points (Fig. 4h). Participants never had direct access to the active rule. So their model-inferred belief, rather than the truly active rule, was the quantity that should drive their behavior in principle, and likely also did govern behavior in practice (Fig. 4d,e). In the following, we thus

used the model-derived belief time courses to interrogate the patterns of fMRI-based estimates of correlated neural variability across the sensory-motor network.

Correlated neural variability tracks beliefs about sensory-motor mapping rule

In a first approach, analogous to our previous analysis of the instructed rule task (Fig. 3), we grouped all time points within the task runs into periods where the belief parameter favored rule 1, or where belief favored rule 2, and compared the patterns of co-fluctuations of stimulus and action codes across the sensory-motor network between these data segments (Fig. 5a-c).

Again, we found a consistent difference in the sign of correlations between stimulus and action decoders between the distinct rule beliefs (Fig. 5a-c,e). As for the instructed rule task, this pattern of intrinsic correlations was (i) dissimilar from the one of feature-specific evoked responses (Fig. 5d), (ii) absent when substituting the visual cortical regions for a set of auditory control regions (Fig. 5f; Supplementary Fig. 6h-k), and (iii) preserved when only assessed in the inter-trial intervals (Supplementary Fig. 5c). Indeed, the rule-related patterns of intrinsic correlations across region pairs were similar in both behavioral contexts (instructed and inferred rule, Fig. 5j)

In a second approach, we interrogated the relationship between belief and the coupling of stimulus and action patterns at a finer level of granularity, exploiting the continuous nature of both variables. We again estimated the co-fluctuations of stimulus and action codes in a time-variant fashion (Materials and Methods), and correlated these estimates with the time courses of model-derived belief (see Fig. 5g for an example region pair and participant). Doing so with a cross-validated multivariate regression model (Materials and Methods) yielded significant prediction from the correlations of all the stimulus- and action-decoder pairs (Fig. 5h). Again, as for the categorical prediction in the instructed rule task (Fig. 3g,h), belief prediction in the inferred rule task was not possible based on the outputs of either the stimulus or the action decoders alone (Fig. 5i).

In sum, the results shown in Fig. 5 replicate all main results from the instructed rule task (Fig. 3) in independent data and a different behavioral context, and extend them to the prediction of continuous, graded, and internal beliefs about active rules.

Functional significance of rule-specific correlation patterns

How did the different features of neural activity in the sensory-motor network assessed here relate to participants' overt behavioral performance? Different from the instructed rule task, the inferred rule task was challenging and yielded a substantial fraction of errors (Fig. 4e). Action decoding from patterns of evoked fMRI responses during the primary task was robustly above chance for both correct and error trials, for PMd (Fig. 6a,b) and other action-coding regions (Supplementary Fig. 8). Decoding was more reliable on correct choices (Fig. 6b, Supplementary Fig. 8), consistent with previous findings for choice accuracy (Harvey et al., 2012) or confidence (Kiani and Shadlen, 2009; Wilming et al., 2020). Patterns of evoked responses in V1-V4 also enabled robust decoding of the grating orientation for both correct and error trials, here with no difference between the two behavioral categories (Fig. 6c).

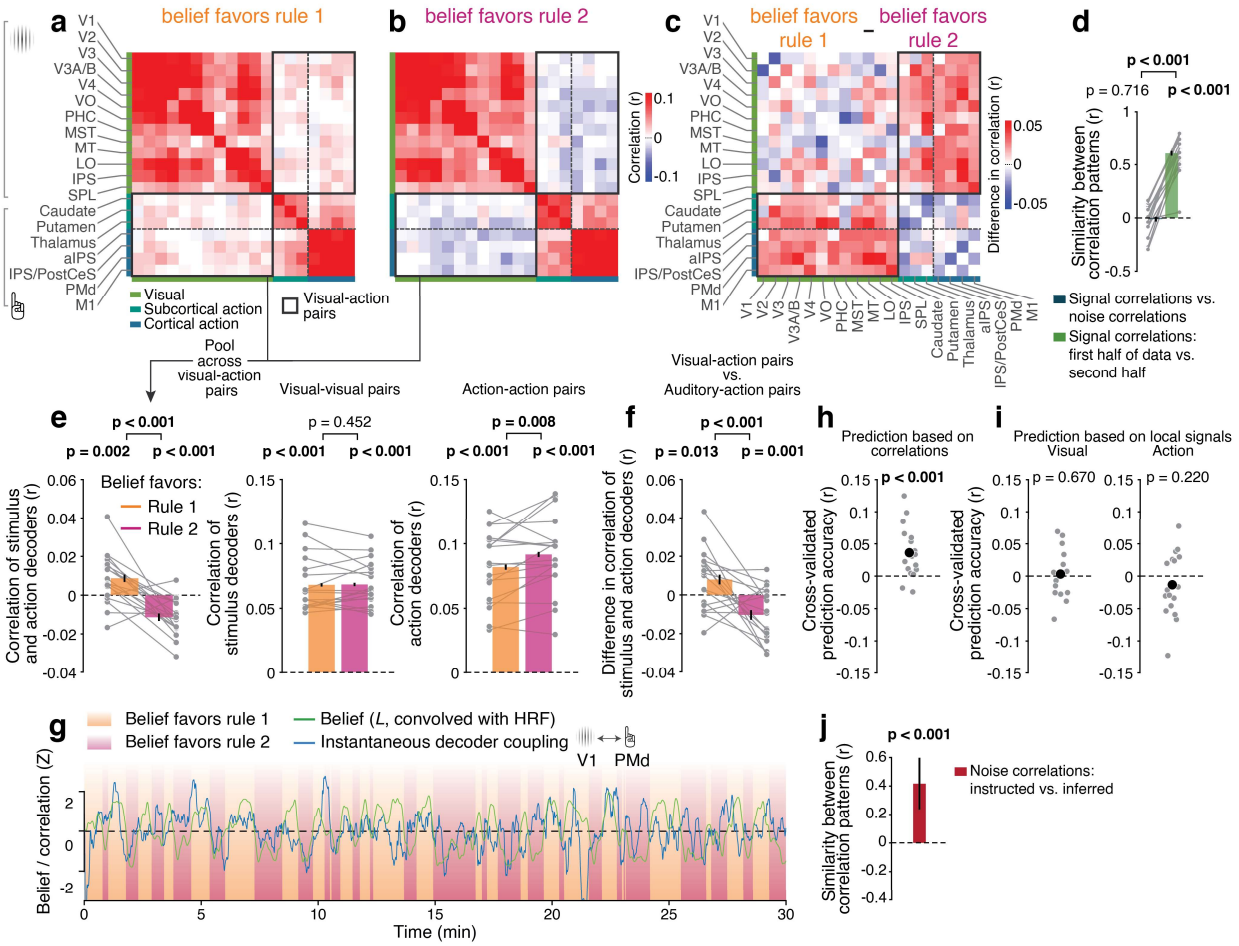


Fig. 5. Rule belief encoding in correlations of stimulus and action codes across sensory-motor networks. a-c. Same as Fig. 3a-c, but data segments for correlation based on the sign of model-derived belief (i.e., whether belief favors rule 1 or 2), rather than based on active (generative) rule. **d.** Similarity of correlation patterns of decoder outputs for trial-evoked responses ('signal correlations'; Materials and Methods) and for intrinsic fluctuations ('noise correlations'). Gray lines, individual participants; bars, group average; error bars, within-participant SEM (Materials and Methods). **e.** Left: Correlations from panels a and b, collapsed across all pairs in gray rectangle. Right: Correlations from panels a and b, collapsed across all visual-visual pairs and action-action pairs (excluding values on diagonals). Gray lines, individual participants; bars, group average; error bars, within-participant SEM (Materials and Methods). **f.** Correlations from panel d (left), relative to the correlations of the graded outputs of a stimulus orientation decoder that was trained on auditory cortex (Supplementary Fig. 6.), and action decoder. **g.** Example time courses of belief (convolved with hemodynamic response function, HRF), and temporally resolved correlation (Materials and Methods) between graded stimulus decoder output from V1 and graded action decoder output from PMd. The correlation of decoder outputs is smoothed for visualization purposes only. **h.** Cross-validated prediction of belief from the time-resolved coupling between stimulus and action decoders. Gray dots show the individual participants, the black dot shows the group mean. **i.** As f, but rule predictions based on either local stimulus decoder outputs or action decoder outputs, rather than their correlations. N=18 for all panels. **j.** Similarity of the patterns of intrinsic decoder co-fluctuations in instructed and inferred rule tasks (gray rectangles in Fig. 3c and Fig. 5c). Error bars, 95th particle of permuted null distributions.

While the patterns of intrinsic co-fluctuation between stimulus- and action-codes were robustly expressed on correct trials (Fig. 6d, left; compare with Fig. 3e), these patterns broke down on error trials (Fig. 6d, middle). Consequently, there was a marked difference between the correlation patterns on correct and error trials (Fig. 6d, right). Our analysis ensured that the lower number of incorrect trials relative to correct trials could not account for the absence of the effect during errors (Materials and Methods). Furthermore, the Bayes factor provided

‘substantial’ evidence (Wetzels and Wagenmakers, 2012) for the null-effect of rule on the correlation pattern on error trials (Fig. 6d). Notably, the observation that stimulus and action decoding from evoked responses (Fig. 6b-c) was well above chance on error trials implies that the coupling of evoked responses flipped sign on errors compared to correct trials (because a given stimulus was followed by the opposite choice on errors compared to correct trials). This flip in coupling of evoked responses on errors contrasts with the absence of rule-specific correlated, intrinsic fluctuations on errors (Fig. 6d).

Errors on the task could originate from several sources. One possible source is noise corrupting the primary orientation judgment, specifically variability of sensory responses (Arieli et al., 1996). Near-perfect performance in instructed rule (Fig. 2d) and indistinguishable orientation decoding from early visual cortex for correct and error trials (Fig. 6c) indicates that this noise source was negligible, as reflected in our model (Fig. 4b). Another source is the uncertainty about the active rule inherent to the task: Even a bias- and noise-free ideal observer used the wrong rule in about 12% of trials when exposed to the same evidence streams as participants (Fig. 4e, purple). A third source is suboptimality of the higher-order decision, in the form of noise or biases corrupting the evidence accumulation and/or resulting rule selection (Drugowitsch et al., 2016; Murphy et al., 2021).

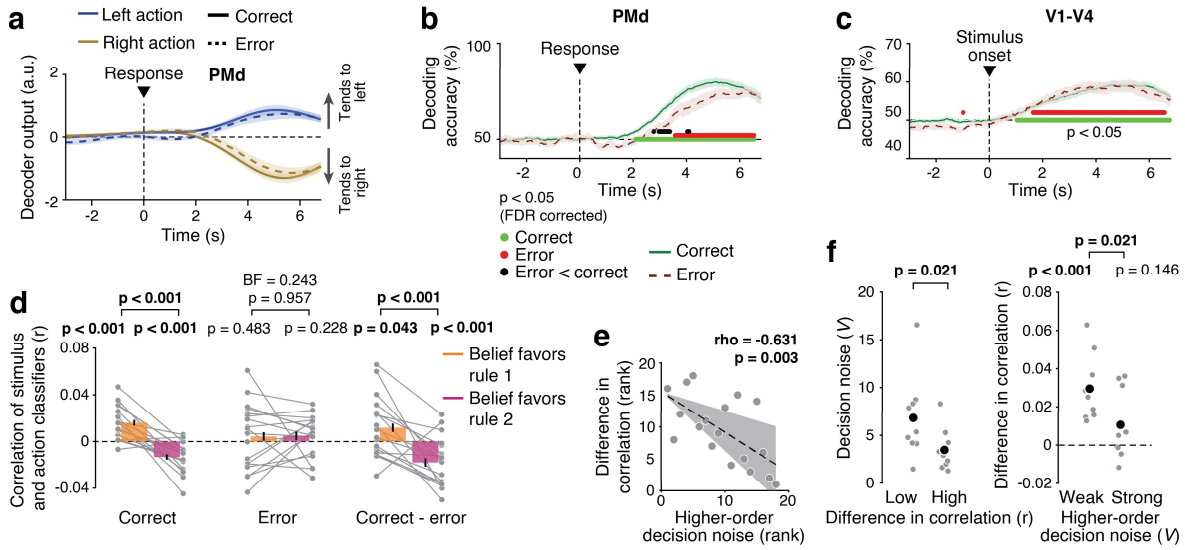


Fig. 6. Relation of correlated variability patterns to behavioral performance. a. Trial-averaged graded action decoder output in PMd time-locked to choice, for correct and error trials. **b.** Accuracy of action choice decoding from evoked responses in PMd. **c.** Accuracy of stimulus orientation decoding from evoked responses in V1-V4. b-d: Shaded areas, between-participant SEM. **d.** As Fig. 5e left panel, but separately for data segments surrounding correct and error trials. BF: Bayes Factor. **e.** Across-participant correlation between the model-derived noise estimates and the difference in the co-fluctuations of stimulus and action codes between the two belief states across all trials. Shaded area, 95% CI (bootstrapped) around the least squares regression line. **f.** Median split of participants’ model-derived noise estimate, based on the difference in the co-fluctuations of stimulus and action codes between belief states, and vice-versa. Gray dots, individual participants; black dot, group average. N=18 for all panels.

Based on these considerations, we examined the relationship between individual differences in performance, and the rule-related change of correlation patterns in the sensory-motor network. While all participants performed near-perfectly in instructed rule (Fig. 2d), they exhibited varying degrees of behavioral suboptimalities in inferred rule (Fig. 4d; range of individual accuracies:

63% to 88%). In our model, deviations of accuracy from the ideal observer were accounted for by the individual noise parameter *V* (Fig. 4b,c). Indeed, individual *V* estimates correlated strongly with the differences in the co-fluctuations of stimulus and action codes between the two belief states (favoring rule 1 versus rule 2; Spearman's rho = -0.631, p = 0.003; Fig. 6e). This was the case irrespective of whether the individual choices were consistent or inconsistent with the model-inferred belief at that moment, indicating that the relationship between correlated variability and internal noise was a stable individual trait, rather than only expressed when choices were actually decoupled from the belief state (Supplementary Fig. 9). Dividing participants based on the difference in correlations (between belief favoring rule 1 versus rule 2; median split) showed a clear difference in *V* between the sub-groups (Fig. 6f, left). The reverse median split, based on *V*, showed a significant correlation difference only for the low-*V* sub-group, but not for the high-*V* subgroup (Fig. 6f, right). The difference in correlated decoder output also reliably predicted *V* in held out participants (correlation between predicted and observed *V*: r = 0.523, p = 0.026). These results implicate the observed flexibility of intrinsic correlation patterns as a trait that governs the individual level of suboptimality of the higher-order rule selection process.

Taken together, the results suggest that the flexibility of the patterns of co-fluctuation of stimulus and action codes across the sensory-motor network limited performance on our challenging task, linking it to trial-to-trial fluctuations of performance within participants (correct vs. error), as well as across-participant variation in internal noise (*V*).

Rule and belief encoding within premotor and visual cortex

Our analyses showed that (participants' belief about) the appropriate sensory-motor mapping rule was reflected in the correlations between local population codes for stimulus and action in the sensory-motor network (Figs. 3f and 5g), but not in these local population codes themselves (belief; Figs. 3g and 5h). The latter is expected because the decoders used to estimate these population codes were each trained on the stimuli or actions only, which were, by design, orthogonal to the task-relevant stimulus-response association specified by the rule.

We, therefore, also trained a separate set of decoders to identify putative representations of (belief about) the rule in local activity patterns in a large set of regions covering the entire cortex (Materials and Methods). Local patterns of fMRI activity afforded robust decoding of (belief about) the active rule during both the instructed rule and the inferred rule runs in three regions: PMd (area 6d), posterior parietal cortex (area 5L, only left hemisphere), and primary visual cortex (V1, Fig. 7a-c). Sub-cortical regions did not yield significant decoding of rule in either the instructed or inferred rule task (all p-values > 0.075).

Encoding of the instructed rule was more confined, restricted to the above three regions and two prefrontal cortical areas (SFL and 10v, Fig. 7a,c), whereas encoding of rule belief (inferred rule) was widespread across many regions (Fig. 7b,c), similar to what was observed in studies using larger and more rule sets (Ito et al, 2022). The more confined effects in the instructed rule task may be due to the smaller demand in tracking and implementing the active rule in this task (Miller & Cohen, 2001) and/or the long delays without any additional rule cues in that condition (Fig. 2a), which may approach the limits of persistent cortical activity. However, the observation that rule encoding was robust in some cortical areas, including even V1, in that condition (Fig. 7c), argues against the latter possibility. Encoding of SR-mapping rules in posterior parietal and frontal cortex is consistent with primate physiology (Wallis et al., 2001; Stoet and Snyder, 2004;

Sakai, 2008; Bennur and Gold, 2011), but the task-general encoding of rule information in V1 may be unexpected (but see Zhang et al., 2013; Siegel et al., 2015).

One concern is that the V1 effect may have been confounded by sensory responses to the visual cues that signaled the SR-mapping rule in an unambiguous (instructed rule) or ambiguous (inferred rule) fashion. This concern is difficult to address for the inferred rule task, because the time course of cue positions was similar to the belief state time courses after convolution with the hemodynamic response function. However, for instructed rule, the scenario makes two predictions that were readily testable: (i) If responses evoked by the SR-cue drove the rule decoding in V1, then training and testing the decoder on cue-evoked responses should yield at least as precise rule decoding from V1; (ii) removing cue-evoked responses should impair rule decoder performance. Both predictions were falsified (Fig. 7d, Supplementary Fig. 10), indicating that genuine rule information was present in persistent V1 activity.

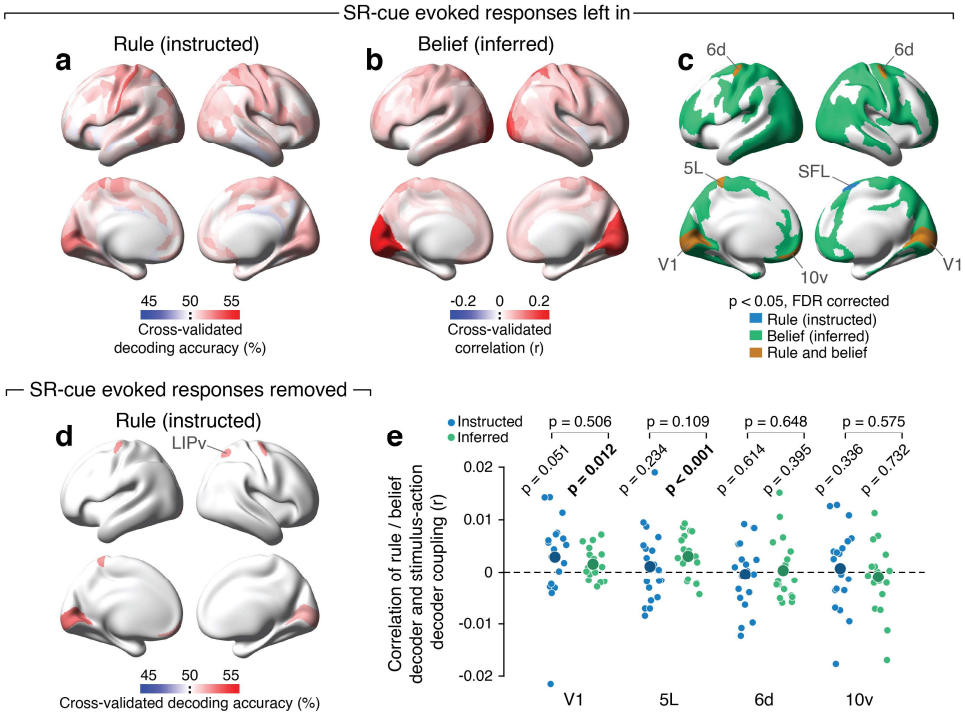


Fig. 7. Local encoding of rule or rule belief across cortical areas. All maps show cross-validated classification accuracies. **a.** Active rule (instructed rule task), decoded by means of support vector machine from continuous time-series data encompassing ITIs as well as choice trials (Materials and Methods). **b.** Graded rule belief (inferred rule task), decoded by means of support vector regression from the continuous time-series data (Materials and Methods). **c.** Overlay of areas with significant classification accuracy from **a** and **b**. **d.** Significant ($p < 0.05$, FDR corrected) active rule decoding (instructed rule task), after removing responses evoked by the SR-mapping cue. **e.** Correlations between rule (instructed) or belief (inferred) decoder output, and time-variant stimulus-action decoder coupling. Small dots, individual participants; Large dots, group average. ROI labels on the x-axis mark the regions on which the rule / belief decoders were trained.

The dynamics of the rule codes in V1 and parietal cortex (area 5L) also correlated with the dynamics of instantaneous coupling between stimulus and action codes (Fig. 7e). We measured the dynamics of the local rule code in terms of the continuous output of the rule belief decoder (inferred rule) or rule decoder (instructed rule; Materials and Methods). For both V1 and 5L, we found a significant correlation for the inferred rule task (V1: mean r = 0.002, p = 0.012, 5L: mean

r = 003, $p < 0.001$), and a trend for V1 in the instructed rule task (mean r = 0.003, p = 0.051). Taken together, our findings point to the interplay between local rule codes within parietal and visual cortex and distributed rule information expressed in the coupling between stimulus and action codes.

No reflection of rule in correlations of region-average fMRI signals

Our analyses of correlations between decoder outputs were based on spontaneous fluctuations of fMRI signals (Arieli et al., 1996; Fox and Raichle, 2007), over and above the evoked responses during trials of the primary choice task as well as occurring during the inter-trial intervals. A large body of neuroimaging work has used spontaneous signal fluctuations to infer intrinsic networks of brain regions and assess their dependence on neuromodulatory or behavioral state, including sensory-motor mapping rules (Fox and Raichle, 2007; Vincent et al., 2007; Honey et al., 2009; Heinzle et al., 2012; Hipp et al., 2012; van den Brink et al., 2016; van den Brink et al., 2018; Lurie et al., 2019; van den Brink et al., 2019; Zamani Esfahlani et al., 2020; Pfeffer et al., 2021). In this previous work, correlations were computed between fMRI signals from single voxels or from the voxel-average per region, thus lacking the feature-specific information entailed in the local activity patterns (Haynes and Rees, 2006; Kriegeskorte and Bandettinin, 2007).

In contrast, by correlating feature-specific signals contained in the multivariate activity patterns within each area (i.e., continuous decoder outputs) between regions (Supplementary Fig. 3), our analysis approach quantified 'informational connectivity' (Coutanche and Thompson-Schill, 2013). A final set of control analyses tested if that aspect was indeed necessary to detect rule information in the correlation patterns. We repeated the comparison between correlations for each active rule (instructed rule) or belief about the active rule (inferred rule) as before, only this time using the voxel-average time series of the same ROIs. This did not yield any difference between rules in the inter-regional correlations (Supplementary Fig. 11). Thus, inter-regional correlations of average activity were uninformative about the active rule, highlighting the specificity of the rule encoding for correlations of graded decoder outputs.

Discussion

A large body of work has focused on the costs of switching between different 'task sets' (Sakai, 2008), measured in terms of behavioral performance reductions, and/or the recruitment of many brain regions (Monsell, 2003). Another line of work has assessed the information about such task rules that is present within individual brain regions (Wallis et al., 2001; Stoet and Snyder, 2004; Haynes et al., 2007; Sakai, 2008; Bode and Haynes, 2009; Bennur and Gold, 2011; Zhang et al., 2013; Cole et al., 2016). Yet, which and how brain circuits are configured for a given task set has remained elusive (Shadlen & Kiani, 2013). Some evidence points to hub regions in association cortex that encode conjunctions of the current stimulus and the rule that is active during a given trial, which may implement a transient activation of local rule-dependent switches on the way from sensory to motor areas (Cocuzza et al., 2020; Kikumoto and Mayr, 2020; Ito et al., 2022). Other theoretical work raises the possibility of a dynamic reconfiguration of the task-relevant networks of sensory and motor regions themselves (Miller & Cohen, 2001; Fusi et al, 2007). Evaluating such schemes of large-scale information flow across the brain requires the simultaneous assessment of feature-specific codes in stimulus-, action-, and rule-encoding brain regions.

Here, we developed an approach to track the ongoing reconfiguration of the task-specific circuit, and relate this reconfiguration to the dynamics of participants' beliefs about the appropriate stimulus-action association. The approach is based on the assumption, well established in neurophysiology (Gerstein and Perkel, 1969; Büchel and Friston, 2000), that the architecture of functional networks is manifested in the correlation structure of spontaneous neural activity. Previous neuroimaging work has exploited this principle to infer coarse-grained functional networks of brain regions (Fox & Raichle, 2007; Vincent et al, 2007; Honey et al, 2009; Hipp et al, 2013) Our approach merged this principle with the assessment of feature-specific population codes that are expressed in the fine-grained activity patterns within each brain region. We found that our approach was indeed critical for revealing the encoding of (beliefs about) the required sensory-motor mapping in intrinsic co-fluctuations across brain areas. The latter was a distributed form of rule information, which co-existed, and interacted, with rule-specific activity patterns within individual brain regions (Fig. 7).

Our findings provide new empirical constraints for theoretical studies of the neural circuit mechanisms underlying context-dependent sensory-motor transformations (Fusi et al., 2007; Mante et al., 2013; Rigotti et al., 2013; Saez et al., 2015; Mastrogiuseppe and Ostojic, 2018; Yang et al., 2019). This theoretical work has so far focused on the dynamics of neural population activity within local cortical regions. Our approach shifts the focus on correlations *between* feature-selective patterns of spontaneous activity expressed in distinct areas of the sensory-motor network. This property of collective neural dynamics has not yet been explored in computational analyses of flexible input-output mapping. One study, unrelated to sensory-motor mapping, has analyzed correlations within and between sensory and motor modules of a recurrent neural network that was trained on a visual evidence accumulation task (Pinto et al., 2019). Our findings may inspire future studies to train similar large-scale and modular networks to solve rule switching problems. The correlation patterns between the sensory and action codes in these various modules can be analyzed with the same approach we present here.

Our study also constitutes a key extension of recent work showing that humans (Glaze et al., 2015; Filipowicz et al., 2020; Murphy et al., 2021; Weiss et al., 2021) and rats (Piet et al., 2018) can accumulate perceptual evidence in volatile environments in an approximately normative fashion. In all these studies, the behavioral tasks required basic perceptual judgments about the state of the sensory environment. By contrast, in our task, the accumulation process informed a higher-order decision: the selection of an internal task model (i.e., the appropriate mapping-rule) used to report the outcome of a lower-order perceptual decision. The observation that the same normative process can explain human behavior also in our context generalizes these previous findings to higher-order cognition, and establishes the adaptive nature of the human cognitive machinery for inference in an uncertain world.

Two previous studies have shown that humans and non-human primates can effectively implement hierarchical decision processes in volatile environments (Purcell and Kiani, 2016; Sarafyazd and Jazayeri, 2019). In these studies, participants needed to infer possible changes in the sensory-motor mapping rule from negative choice outcomes after a challenging perceptual judgment that was subject to uncertainty. Participants' behavior could be explained by models that integrated previous choice outcomes with the expected accuracy of perceptual judgments to infer rule switches. The findings from our inferred rule task extend the principle of hierarchical decisions (here: about rule and stimulus orientation) to a setting that requires a different inference strategy: In our task trial-by-trial outcomes were not available (performance feedback was only provided at the end of each run) and there was no uncertainty about the primary task. Instead of integrating outcomes over trials, participants integrated noisy sensory

cues on a dimension orthogonal to the features of the primary choice task. In this setting, participants were able to effectively track the changing mapping rules and apply them to the report of the primary judgment. The coupling of integration processes at different hierarchical levels seems to be a key principle used by the brain to generate adaptive behavior in uncertain environments.

Other work has studied the neural bases of switching between sensory-motor mappings with inherent asymmetries (Heinzle et al., 2012; Sarafyazd and Jazayeri, 2019; Duan et al., 2021), such as pro- and anti-saccade tasks (Munoz and Everling, 2004). Here, the 'pro-rule' requires using a neural default pathway between matching positions within the spatial maps in sensory and action-coding brain regions, whereas the 'anti-rule' requires overriding this default pathway. By contrast, in our task, the mapping from the feature space of visual orientation to the space of motor action required a selective, and completely arbitrary, routing of signals from populations of visual cortical neurons to action-coding neural populations in downstream regions, eliminating any asymmetry between the rules. Recent work has begun to illuminate the local circuit mechanisms underlying the switches between asymmetric rules in prefrontal cortex (Sarafyazd and Jazayeri, 2019) and the brainstem (Duan et al., 2021). It may be instructive to compare large-scale patterns of rule-selective connectivity in cortical sensory-motor pathways between asymmetric and symmetric rule settings.

Our insight that patterns of correlated neural variability reflect beliefs about sensory-motor mapping rules bears analogy to task-related patterns of noise correlations within sensory cortex (Cohen and Newsome, 2008; Haefner et al., 2016; Bondy et al., 2018). In our work, the rule-specific patterns of correlated activity were expressed at the macroscale, spanning many different cortical and subcortical brain regions that make up the sensory-motor network for the primary choice task. Even so, the analogy of findings points to a common functional principle that may underlie these adaptive patterns of correlated variability at different scales – for example, top-down signaling from task-coding regions (Miller and Cohen, 2001; Sakai, 2008) or neural sampling (Fiser et al., 2010; Haefner et al., 2016).

Our findings suggest that rules that are required for the execution of basic choice tasks may be encoded in a distributed fashion, in the form of correlated neuronal variability, across sensory-motor pathways. It is tempting to speculate that correlated variability may be a general format the brain employs for encoding contextual variables: Different from sensory, motor, or even cognitive variables (e.g., value or symbolic meaning), such contextual variables do not need to be 'read out' from any downstream neural population, but rather control the information routing through sensory-motor networks.

Acknowledgements

We thank Joshua Gold, Matthew Nassar, Sander Nieuwenhuis, Srdjan Ostojic, and Stefano Panzeri for helpful discussion, and Jürgen Finsterbusch for help with preparation of MRI sequences. This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation, DO 1240/3-1, DO 1240/4-1 and 178316478 - A6 (CB), A7 (THD) & Z3 (THD).

Materials and Methods

Participants

A total of 22 participants (median age 27, range 21 – 44, 8 male) took part in our experiment. All participants gave written informed consent and the study was approved by the ethics committee of the Hamburg Medical Association. All participants were healthy individuals with normal or corrected vision recruited via the recruitment pool of the Department of Neurophysiology and Pathophysiology of the University Medical Center Hamburg-Eppendorf. Exclusion criteria included a current or past diagnosis of mental or neurological illness, use of illegal substances, above-average consumption of alcohol as well as non-compatibility with the MRI-scanner.

The experiment comprised three sessions, one behavioral training session and two sessions in the MRI-scanner. All but one participant completed all three sessions. This one participant, plus another two were excluded from MRI analyses, the latter two because of a failure to record pulse and respiration. One further participant was excluded from the inferred rule task (see below) because of an error in logging the response data. Thus, 19 (instructed rule) and 18 (inferred rule) out of 22 tested participants were included in the analyses presented here.

Participants were remunerated with 10 Euros per hour, 10 Euros for completing all three sessions, and a variable bonus, the amount of which depended on task performance across all three sessions. The maximum bonus was 30 Euros.

Behavioral tasks

Participants performed two different versions of a hierarchical decision-making choice task, which combined the selection of a changing sensory-motor (SR) mapping rule (higher-order decision) with a basic visual orientation discrimination judgment (lower-order decision). The SR-mapping rule defined the correct mapping from visual orientations to motor responses (button press with either left or right index finger; Fig. 1a). Participants needed to apply the selected rule to report their orientation judgment in order to obtain a monetary reward (5 cents per correct choice in the ‘inferred rule’ and 3.5 cents in the ‘instructed rule’ task version). The lower-order decision was the same in both versions of the task. Stimuli for this decision were large contrast gratings of either vertical or horizontal orientation (see *Stimuli and Procedure* for details).

The two different versions of the task, called ‘instructed rule’ or ‘inferred rule’, differed only in the difficulty of the higher-order decision (i.e., rule selection). During ‘instructed rule’ runs, visual SR-mapping cues, presented transiently (duration: 400 ms) during the inter-trial-intervals (ITIs) every two trials, instructed participants unambiguously that one of the two possible stimulus-response mapping rules would be active on the next pair of trials. The active rule alternated every pair of trials. Thus, in instructed rule runs, the selection of the SR-mapping rule was both unambiguous and predictable. ITIs for the lower-order decision were long and variable (uniform: 4-20 s).

In ‘inferred-rule’ runs, by contrast, the active rule had to be continuously inferred from a sequence of noisy sensory evidence samples and it could undergo hidden changes at any moment during each run. Participants monitored a rapid stream of evidence samples: small dots that were flashed briefly (100 ms, 400 ms SOA) around the horizontal meridian. The sample positions were drawn from one of two generative distributions: two Gaussians with equal

standard deviation ($\sigma_{left} = \sigma_{right}$) and different means symmetric around fixation ($|\mu_{left}| = |\mu_{right}|$). The generative distribution at any moment governed the active rule, and it could change from one sample to the next with a low probability (hazard rate) of 0.0143. At variable ITIs (uniform: 6.8 – 29.6 s), the stimulus for the lower-order decision appeared, prompting participants to report their orientation judgment. The correctness of the response depended on both the selection of the correct rule and the correct orientation judgment. Thus, in inferred rule runs, participants needed to continuously infer and select the rule that was most likely to be active at any time. To this end, they needed to integrate the noisy rule evidence over time. Finally, they needed to apply the selected rule on the next trial of the lower-order decision (orientation discrimination).

Participants were instructed at the beginning of each block, which distribution corresponded to which rule. The relationship between the two generative distributions and the response rules stayed constant across experimental sessions.

Stimuli and procedure

All stimuli were created using Matlab and the Psychophysics Toolbox Version 3 (Brainard, 1997) and presented on a medium grey background. The fixation mark, presented throughout the entire run in the center of the screen, was a white symmetric cross with a length of 0.51° of visual angle and a thickness of 0.05°. The grating stimuli for the lower-order decision were circular, achromatic Gabor patches with full contrast and truncated at an inner eccentricity of 2.5° and an outer eccentricity of 13.85°. The spatial frequency was fixed at 1.2 cycles/°, whereas orientation (the discriminandum) varied randomly from trial to trial and was either vertical or horizontal.

In the MRI-scanner, stimuli were presented on an MRI-compatible LCD-screen with a resolution of 1920x1080 pixels at a refresh rate of 60 Hz. The screen was positioned at an approximate distance of 60 cm and viewed through a surface mirror that was mounted on top of the head coil. In the training session in the psychophysics lab stimuli were presented on a VIEWPixx monitor (VPixx Technologies, Saint-Bruno, Quebec, Canada) with the same resolution and refresh rate as the monitor in the MRI-scanner. The grating stimuli spanned the full height of the projection screen. Participants reported their decisions by button presses with their left or right hand. To this end, they used two MRI-compatible button interfaces (Current Designs, Philadelphia, Pennsylvania, USA) in the scanner and a keyboard in the psychophysics lab. At the end of each run, participants received feedback regarding their overall performance in that run, in the form of (i) the percentage of correct choices, (ii) the monetary reward gained in the run, and (iii) their total monetary reward accumulated across the whole experiment up to that moment.

The two MRI sessions of each participant took place on the same testing day with a break of 105 minutes in between sessions. The training session took place 1-2 days before the MRI sessions. During the training session, participants first performed a run in which the cue that instructed the rule was continuously presented. Next, participants performed one instructed rule run, in which the participants were still informed about the correct rule before each trial, but the cue was only transiently presented, and had to be remembered upon trial presentation. Afterward, participants performed five runs of the final task version that was used during the main experiment (i.e., MRI sessions). Before the onset of each run, participants were shown a visualization of the mapping of generative distributions to rules, in order to avoid error trials resulting from false rule association.

In each MRI session, we first ran three blocks of retinotopic mapping (not used for the current article). Then, participants performed three blocks of the inferred rule task, which on average lasted 609 s (SD=0.54 s) and included 36.0 choices (SD=1.94). Thereafter, participants performed two instructed rule runs, which lasted on average 604 s (SD=6.40 s) and included on average 56.5 choices (SD=1.44). As during the training session, participants were reminded of the correspondence between generative distributions and active rule.

Data collection

All MRI data were collected with a Siemens Prisma 3T MRI scanner with a 32-channel head coil. During task performance on each of the two MRI sessions, we collected 5 runs (3 inference, 2 instructed) of T2*-weighted EPI data (Flip angle: 70; TR: 1.9 s; TE: 28 ms; FOV: 112 x 112 x 62 slices of 2 mm isotropic voxels; 328 volumes). Cardiac pulsation and breathing during task performance were recorded using a pulse oximeter and pneumatic belt. Eye position and pupil size were recorded with an EyeLink 1000. At the end of the second session, we collected a T1 high resolution anatomical scan (MPRAGE; Flip angle: 9; TR: 2.3 s; TE: 2.98 ms; FOV: 192 x 240 x 256 slices; 1 x 1 x 1 mm) for registration purposes. On both MRI sessions, we also collected B0 field homogeneity scans for field distortion correction purposes (Phase image: flip angle: 40; TR: 0.678 s; TE: 7.88 ms; Magnitude image: flip angle: 40; TR: 0.678 s; TE: 5.42 ms).

Behavioral modeling

Normative model

The normative solution for the higher-order decision problem (rule selection) in the inferred rule task was cast as a dynamic belief updating process (Glaze et al., 2015). In this model, belief *L* was expressed in units of log-odds, for one possible state (rule 1) versus the alternative possible state (rule 2). For each evidence sample X_n, the posterior belief L_n was computed by combining the log-likelihood ratio LLR_n associated with that sample, with the prior belief ψ_n. The key feature rendering the model adaptive in volatile environments (i.e., possibility of hidden state changes) was a non-linear transformation of the posterior from each updating step into the prior for the next step, dependent on the subjective hazard rate *H*:

$$L_n = \psi(L_{n-1}, H) + LLR_n \qquad (1)$$

Here, LLR_n was the logarithm of the ratio of the likelihoods of the sample X_n to be observed under each of the two states, whereby the states corresponded to the generative distributions. LLR_n could also be expressed as the difference of the logs of the likelihoods:

$$LLR_n = \log(normal\ (X_n \mid \mu_1, \sigma_1^2)) - \log(normal\ (X_n \mid \mu_2, \sigma_2^2)) \qquad (2)$$

The non-linear transformation of posterior into the next prior, $\psi(L_{n-1}, H)$ was given by:

$$\psi(L_{n-1}, H) = L_{n-1} + \log\left[\frac{1-H}{H} + \exp(-L_{n-1})\right] - \log\left[\frac{1-H}{H} + \exp(L_{n-1})\right] \qquad (3)$$

This transformation discounted the previous posterior as a function of the participant's estimate of the environmental hazard rate *H* (i.e., probability of a change in generative distribution) and

constituted the key difference to previous evidence accumulation schemes (Bogacz et al., 2006). When *H*=0, $\psi = L_{n-1}$, resulting in perfect accumulation (no discounting). When H=0.5, the three righthand terms in equation (3) cancel out, so that ψ = 0 and L_n = LLR_n (complete discounting). Thus, *H* balanced the impact of new evidence and prior belief on the current belief and thereby controlled the tradeoff between stable evidence accumulation and sensitivity to change-points. The prior was initialized at the start of each run ($\psi_1 = 0$) and then evolved throughout the run according to eq. 3.

We assumed that rule selection for each sample *n* was based on L_n corrupted by a Gaussian noise term *V*, such that the probability $\hat{q}_n$ of selecting rule 1 was computed as:

$$\hat{q}_n = \frac{1}{2} + \frac{1}{2}\mathrm{erf}\left(\frac{L_n}{\sqrt{2}V}\right) \qquad (4)$$

This rule selection probability then determined the probability of each action choice (left or right) upon presentation of a vertical or horizontal grating on trial *trl*. Specifically, the probability of choosing a left response $\hat{r}_{trl}$ was computed as:

$$\begin{aligned} \hat{r}_{trl}|vertical &= \hat{q}_{trl} \\ \hat{r}_{trl}|horizontal &= 1 - \hat{q}_{trl} \end{aligned} \qquad (5)$$

where $\hat{q}_{trl}$ corresponds to the estimated rule selection probability after accumulating the sample directly preceding the grating presentation on trial *trl*.

Model fitting

We assumed that the lower-order decision (orientation judgment) was noise-free so that behavioral performance was only limited by the higher-order decision (rule selection) and a correct observed behavioral choice on a given trial implied the latent selection of the correct rule. Furthermore, we assumed no biases other than a potentially biased internal representation of the true hazard rate. Consequently, *H* and *V* were the only free parameters in our model fits to participants' data.

Following previous work (Glaze et al., 2015; Murphy et al., 2021), we fitted the model separately to each participant's data, by minimizing the cross-entropy between the choices of that participant and the model:

$$e = -\sum_{trl}(1 - r_{trl})\log(\hat{r}_{trl}) + r_{trl}\log(\hat{r}_{trl}) \qquad (6)$$

where r_{trl} was the participant's choice on trial *trl* (left = 1, right = 0) and $\hat{r}_{trl}$ was the model choice probability on the sample *n* that corresponded to the onset of the same trial (eq. 5). The sum of *e* was minimized via particle swarm optimization. We set wide bounds on all parameters and ran 300 pseudorandomly initialized particles for 1,500 search iterations.

Having fit the model, we computed the participant- and session-specific time courses of L_n and LLR_n for analysis of the fMRI data (see *MRI data analysis* below).

Alternative models

We also computed the performance of three alternative models without noise: i) selecting the rule based on perfect evidence accumulation across the run (i.e., $H = 0$), ii) selecting the rule based on only the last evidence sample (i.e. $H = 0.5$), and iii) the 'ideal observer', which used the normative belief updating process described by eqs 1-3 with the true generative H (i.e.; $H = 1/170$). We computed the performance of these models for the same evidence streams presented to the participants.

Model-derived change-point probability

We used the model fits to compute change-point probability (CPP) associated with each evidence sample, a computational quantity that both the normative belief updating process and human participants are sensitive to in volatile decision-making contexts such as ours (Murphy et al., 2021). CPP was the posterior probability of a change having occurred, given H, the previous posterior and the new evidence sample. CPP was computed as follows (Murphy et al., 2021):

$$\mathrm{CPP}_n = \frac{H(N(X_n\, S_1|)\, p(S_{2,n-1}) + (N(X_n \mid S_2)\, p(S_{1,n-1}))}{H(N(X_n \mid S_1)\, p(S_{2,n-1}) + (N(X_n \mid S_2)\, p(S_{1,n-1})) + (1-H)(N(X_n \mid S_2)\, p(S_{2,n-1}) + (N(X_n \mid S_1)\, p(S_{1,n-1}))}$$

(7),

where S_1 and S_2 denoted the two generative distributions with mean and variance μ_1, σ_1^2 and μ_2, σ_2^2.

CPP_n was used for univariate regressions against fMRI data (see *MRI data analysis* below).

MRI data analysis

Preprocessing and physiological noise correction

We used tools from the FMRIB Software Library for preprocessing of the MRI data (Smith et al., 2004; Jenkinson et al., 2012). EPI scans were first realigned using MCFLIRT motion correction and skull-stripped using BET brain extraction. We used B0 unwarping to control for potential differences in head position each time the participant entered the scanner and resulting differences in geometric distortions in the magnetic field. The B0 scans were first reconstructed into an unwrapped phase angle and magnitude image. The phase image was then converted to units rad/s and median-filtered, and the magnitude image was skull-stripped. We then used FEAT to unwarp the EPI images in the y-direction with a 10% signal loss threshold and effective echo spacing of 0.279998. Following field-map correction, the EPI data were high-pass filtered at 100s, pre-whitened, and corrected for physiological noise using retrospective image correction (RETROICOR) (Glover et al., 2000).

RETROICOR was applied by assigning phases of the cardiac and respiratory cycles to each volume in the EPI time series, and removing them from the data. To this end, pulse oximetry (i.e., cardiac) and respiratory time series were first down-sampled from 500 Hz to 100 Hz. Next, the pulse oximetry data were bandpass filtered between 0.6 and 2 Hz, and the respiration data were low-pass filtered at 1 Hz, using a two-way FIR filter. We extracted peaks in each time series corresponding to maximum blood oxygenation and maximum diaphragm expansion, which were used to construct 34 slice-specific time-series (4th order harmonics for cardiac cycle;

4th order harmonics for respiration cycle; 2nd order harmonics for cardiac-respiration interactions; 2nd order harmonics for respiration-cardiac interactions; 1 heart rate time series; 1 respiratory volume time series). These time-series were used to estimate cardiac and respiratory effects from the EPI time series using multiple linear regression. All further analyses described below proceeded on the residuals from this regression.

Subsequently, slice time correction was applied and the EPI data were co-registered with the anatomical T1 image to 2 mm isotropic MNI space. To this end, we used FLIRT and subsequently FNIRT for maximal anatomical alignment. No spatial smoothing was applied in order to preserve high-spatial frequency information in the functional data. All critical analyses presented in this paper were applied at the level of regions of interests (ROIs, see next section).

Table 1. ROI definition.

ROI labels	Functional areas	Source
V1	Combination of dorsal/ventral parts of V1	(Wang et al., 2015)
V2	Combination of dorsal/ventral parts of V2	
V3	Combination of dorsal/ventral parts of V3	
V3A/B	Combination of V3A and V3B	
V4	Human analogue of V4	
LO	Combination of LO1 and LO2	
VO	Combination of VO1 and VO2	
PHC	Combination of PHC1 and PHC2	
MT	Human analogue of MT	
MST	Human analogue of MST	
IPS	Combination of IPS0-IPS5	
SPL	SPL1	
Caudate	Caudate nucleus	Harvard-Oxford structural Atlas (https://neurovault.org/collections/262/)
Putamen	Putamen	
Thalamus	Thalamus	
PMd	Areas 6a, and 6d	(Glasser et al., 2016)
M1	Somatotopic hand-specific aspect subregion of M1	(de Gee et al., 2017)
IPS/PostCeS	Junction of IPS/ and PostCeS	
aIPS	Anterior part of horizontal ramus of IPS	

Abbreviations. Vx, visual area X; LO, lateral occipital; VO, ventral occipital; PHC, parahippocampal; MST, middle superior temporal; MT, middle temporal; IPS, intraparietal sulcus; SPL, superior parietal lobule; PMd, dorsal premotor cortex; M1 primary motor cortex; PostCeS, postcentral sulcus.

Delineation of ROIs

We analyzed the fMRI data for a large set of ROIs, defined based on functional and anatomical properties based on two published MRI-based atlases (Wang et al., 2015; Glasser et al., 2016), our lab's own previous work (de Gee et al., 2017), and the Harvard-Oxford structural atlas (https://neurovault.org/collections/262/), the latter only for sub-cortical ROIs. The primary set of ROIs that we used for all analyses presented in the main paper, selected to span the sensory-motor pathway underlying the lower-order decision, is defined in Table 1. In addition, we used the complete multi-modal parcellation from the human connectome project (HCP-MMP1.0) (Glasser et al., 2016) for supplementary analyses (see below: *Decoding of rule or belief from local activity patterns*).

Visual cortical areas. Our primary ROI set used an established parcellation of retinotopically organized visual field maps (Wang et al., 2015) to delineate visual cortical areas, some of which were further combined into clusters. For all analyses referring to 'early visual cortex' (Figs. 2 and 6), we further combined areas V1-V4 into a single ROI.

Action-related areas. We used our previous data (de Gee et al., 2017) to define the hand-specific somatotopic subregion of M1 as well as two posterior parietal regions (non-overlapping with the parietal visual field maps), all of which exhibited robust lateralization of activity related to the planning and execution of right versus left hand button presses across a range of previous fMRI and magnetoencephalography (MEG) studies (de Gee et al., 2017; Wilming et al., 2020; Murphy et al., 2021).To this set of action-related ROIs, we added dorsal premotor cortex from the HCP-MMP1.0 parcellation, as well as the caudate nucleus, putamen, and thalamus from the Harvard-Oxford structural atlas, based on the observation that those exhibited robust hand movement-selective activity.

Deconvolution of evoked fMRI responses

We used deconvolution to estimate evoked responses time-locked to specific experimental events (visual grating stimuli, button presses, or SR-mapping cues) without making any assumptions about the shape of the hemodynamic response (Dale, 1999). These trigger events were not synchronized with the EPI acquisition. Thus, we first up-sampled the fMRI data by creating a 1 x N vector S at 50 times the sampling rate, where N denoted the number of up-sampled time points. S was filled with value 1 at the time point of the trigger event, and 0 in all other positions. S was then iteratively staggered forward by one sample P times to yield an N x P design matrix X, with value 1 along the diagonals locked to stimulus onset. P was set to 396, equivalent to ~15 seconds. The estimated response R (a total of P samples) was then obtained via:

$$R = X^{+} Y \qquad (8)$$

where $^{+}$ denotes pseudoinverse. Y was an N x 1 up-sampled, linearly detrended and z-scored time series (single voxel or ROI-average, see below).

Estimation of individual hemodynamic responses functions (HRFs)

Individual HRFs were estimated non-parametrically using deconvolution (see previous section) of the evoked V1 responses to the grating stimulus presentation. For this analysis, time series were first averaged across all voxels within V1. The resulting response was baseline-corrected

by subtraction of the first sample from each sample, and it was normalized to unit height. This response estimate was used as an empirically derived HRF model that was convolved with the regressors in all general linear model (GLM) analyses reported below.

Decoding of stimulus and action from patterns of evoked responses

Decoding of stimuli (horizontal versus vertical orientation) and action (left versus right hand button press) was performed on the data of individual participants using a support vector machine (SVM) classifier with linear kernel (cf. Wilming et al., 2020). For each ROI, we selected the 300 voxels that exhibited maximum fMRI-responses to the gratings (for stimulus decoding) or maximum lateralization of fMRI-responses during button presses (action decoding). Maximal stimulus- and action-evoked fMRI responses (response lateralization for action) were determined based on a GLM of stimulus and action onto the voxel-level fMRI data (see *Whole brain GLMs* below). For stimulus decoding, voxels were pooled across the left and right hemisphere. For action decoding, the number of voxels was evenly split between the left and right hemispheres so as to evenly distribute feature-selectivity between classes.

We then used an iterative procedure to estimate single-trial response amplitudes (beta weights) for each of the selected voxels, by fitting one GLM for each trial (Mumford et al., 2012). Each of those GLMs contained two regressors: one stick function for the trial of interest, and another stick function for all remaining trials, whereby both regressor time series were convolved with the individual HRF estimate (see *Individual HRF estimation* above). Single-trial regression was performed on all runs (including both the instructed and inferred rule runs) concatenated within sessions, but not across sessions, under the assumption that the discriminant pattern may differ between sessions. The resulting trials x voxels matrix served as the data for classification.

Decoding was cross-validated with ten folds. The conditions in the training data, but not the testing data, were balanced by up-sampling the condition with fewer trials, and the training and testing data were subsequently z-scored separately. Decoding accuracy was then computed in the testing data, as the percentage of correctly classified horizontal choice gratings (or leftward actions) plus the percentage of correctly classified vertical choice (or rightward actions) gratings, divided by two. For each type of classification (stimulus orientation or action), this yielded a participant x session x fold x ROI matrix of classification accuracies, which was then averaged across sessions and cross-validation folds. We assessed its significance for each ROI using permutation testing (10,000 iterations) by comparing it to chance level (50% for both stimulus and action decoding). Similar results in terms of direction and significance were obtained using alternate numbers of included voxels (50, 75, 100, or 200).

Removal of feature-specific evoked responses

We again used deconvolution to isolate the ongoing fluctuations of each voxel's activity in order quantify the correlation structure of these fluctuations (see subsequent sections). We here use the term 'ongoing activity' as a shorthand to refer to activity unrelated to the events of the primary choice task (grating stimuli or button presses), acknowledging that these fluctuations may be partly driven by the SR-mapping cue (instructed rule) or the evidence samples for the higher-order decision (inferred rule). Responses to such events are not expected to generate activity patterns encoding one or the other stimulus orientation, or one or the other action. Thus, such responses do not confound our analyses of correlations between stimulus and action decoder outputs described in subsequent sections.

We estimated the mean orientation-specific or action-specific evoked responses and then subtracted these feature-selective evoked responses from the voxel time series. For each type of ROI (visual field maps or action-related), we thus used the procedure described above (see (*Deconvolution of evoked fMRI responses*), but now to construct two separate matrices S, one per feature (horizontal / vertical grating orientation, or left / right response hand). These two matrices were then concatenated horizontally to produce a new composite design matrix X with dimensionality N x 2P. The time series of ongoing activity (i.e., feature-specific evoked responses removed) was then computed as the residual (ε) (Dale, 1999):

$$\varepsilon = Y - X^{T+} Y(X\ Y^{T}) \qquad (9)$$

where $^{+}$ denotes pseudoinverse, ^{T} denotes transposition, and Y was the up-sampled, detrended, and z-scored voxel time series concatenated across the complete session. This procedure was repeated for both sessions.

To verify that this procedure effectively removed feature-specific evoked activity (Supplementary Fig. 2), we projected the SVM decision function onto the time series of multi-voxel patterns (voxels used for stimulus and action decoding, see above) and computed the mean decoding accuracy across trials, time-locked to grating onset (visual cortical areas) or button press (action-related ROIs). This was done once for the full voxel time series and once for the residual voxel time computed via eq. 9.

Correlation of stimulus- and action-selective voxel groups

To illustrate the rationale behind our feature-specific correlation analyses in a simple format for an example participant and area pair (Fig. 2f), we selected the 150 voxels from across early visual cortex (V1-V4) that exhibited the largest responses to horizontal gratings, and another 150 voxels from the same ROI with maximal responses to vertical gratings. Likewise, we selected 150 voxels in left PMd with maximal responses during right button presses, and 150 voxels in right PMd with maximal responses during left button presses. These response amplitudes were estimated with GLM regressing these trial events (grating orientations or actions) on the time series of all voxels within a given ROI, after convolution with the individual HRF estimate. We first removed feature-specific evoked responses (eq. 9) and then the mean activity across all selected voxels (via linear regression). The resulting voxel time series were then selectively averaged for each of the two voxel groups (i.e., across all vertical-preferring and all horizontal-preferring voxels (visual areas) or across all right-preferring and all left-preferring voxels (action-related areas).

We next created a time series that encoded active rule at each given time point in the fMRI data (rule 1 = -1; rule 2 = 1), convolved it with the individual HRFs, and took the sign of this vector to define alternating data segments corresponding to the two active rules. Separately for these two types of data segments (active rules), we finally computed the correlations between the time series of all four pairs of stimulus- and action-selective voxel groups, and tested their interaction with a repeated measures ANOVA.

Correlation of stimulus and action decoder outputs

Our main analyses computed the correlations between the graded outputs of stimulus and/or action decoders applied to the fluctuations of multi-voxel activity patterns (i.e., after removal of feature-specific evoked responses; see above). The analysis pipeline is illustrated in

Supplementary Fig. 3. For each participant, ROI, and experimental session, we projected the decision functions from the stimulus and action decoders trained on the evoked responses (see *Decoding of stimulus and action from patterns of evoked responses*) onto our estimates of ongoing multi-voxel activity patterns (see *Removal of feature-specific evoked responses*). This yielded a scalar time series that quantified the graded decoder output, which was signed and could be interpreted as the instantaneous tendency of the ongoing population activity towards one or the other stimulus orientation (visual decoders) or one or the other action (action decoders). Given the way we computed the decoders, high values indicated an ongoing activity pattern resembling responses to vertical gratings (left-hand button presses) and low values indicated an ongoing activity pattern resembling responses to horizontal gratings (rightward action).

We segmented the time series according to the (belief about) the active mapping rule. For the instructed rule runs this proceeded as described above (*Correlation of stimulus- and action-selective voxel groups*). For the inferred rule runs, we replaced the block vector for active rule by the time series of *L* from the fitted behavioral model. Again, we applied convolution with the individual HRFs and binarized by the sign of the resulting time series accounted for the hemodynamic lag. We then correlated the continuous decoder outputs across all pairs of areas, separately for the two types of data segments corresponding to the active rule (instructed rule) or the rule favored by belief (inferred rule).

Similarity between correlation patterns across visual-action ROI pairs

We used correlation analyses to quantify the similarity between patterns of correlations of decoder outputs across all visual-action ROI pairs, computed for either the intrinsic signal fluctuations ('noise correlations') as used in the main analyses, or for trial-evoked responses ('signal correlations'). We estimated the evoked responses by projecting the stimulus and action classifiers onto the voxel-level data, without residualization. The resulting decoder output time series were epoched, time-locked to stimulus onset, and separately averaged across trials for all four unique stimulus-rule combinations. For each rule, we then concatenated the resulting average decoder responses for the two stimuli, yielding two time series (one per rule) for each ROI. These time series should correlate positively between visual-action ROI pairs for rule 1, and negatively for rule 2, simply due to the rule-consistent flipping of stimulus-response associations in subjects' behavior. We finally used Pearson correlation to quantify the spatial similarity of this pattern of signal correlations (Supplementary Fig. 4) with the pattern of noise correlations (i.e., correlations of continuous decoder outputs after removal of evoked responses as used in our main analyses (gray outline in Fig. 3c and 5c)

To quantify the across-trial consistency of signal correlation patterns in an analogous fashion, we randomly sampled 50% of trials, averaged graded decoder output across trials and computed the pattern of signal correlation as described above. This pattern was then correlated with signal correlation pattern for the other 50% of trials. This was repeated 1,000 times for each per participant, to create a null distribution of expected signal correlation patterns. The average of the null distribution served as the participant-specific estimate of the across-trial similarity (correlation) of signal correlation patterns, and was close to 0.7 for both instructed rule (Fig. 3d, green) and inferred rule (Fig. 5d, green) conditions.

We finally used an analogous approach to quantify the similarity of patterns of noise correlations between the instructed and inferred rule tasks. To this end, the group average pattern of stimulus and action decoder output correlations in instructed (gray rectangle in Fig. 3c) was

correlated with the corresponding pattern from inferred (gray rectangle in Fig. 5c), and the significance of this correlation was assessed by comparing it to a null distribution that was generated by shuffling the ROI-pair labels at the single-participant level (10,000 iterations).

Time-variant correlation of decoder outputs

To estimate stimulus and action decoder correlations in a time-variant manner, we computed a measure of the correlation of graded decoder outputs at a given time point *t* as follows:

$$IC(t) = S(t)A(t) \qquad (10)$$

where *S* and *A* were the z-scored time series of the stimulus (*S*) or action (*A*) decoder outputs. Because the time series *IC*(*t*) averages to the Pearson correlation coefficient, it is a measure of the instantaneous coupling between the decoder outputs (Zamani Esfahlani et al., 2020).

Prediction of active rule or belief from decoder correlations

We used cross-validated multiple regression to test whether the fluctuations of time-variant correlations between stimulus and action decoders, and/or of these decoder outputs themselves, predicted the active rule (instructed rule) or belief about the active rule (inferred rule). The fluctuations of decoder correlations contained correlations of 84 pairs of ROIs, which tended to be similar (Figs. 3 and 5). To reduce the dimensionality of the data (i.e., the feature set for the decoder), we submitted the matrix of time-variant correlations of all stimulus-action pairs, estimated with eq. 10, to PCA. We selected the number of components used for prediction by comparing the eigenvalues λ to a theoretical noise distribution ρ (Mitra and Pesaran, 1999):

$$\rho = \frac{1}{\pi\sigma\lambda}\sqrt{(\lambda_+^2 - \lambda^2)(\lambda^2 - \lambda_-^2)} \qquad (11)$$

where:

$$\lambda_\pm^2 = 2\sigma\left[\frac{p+q}{2}\right] \pm \sqrt{pq} \qquad (12)$$

and σ was the standard deviation of λ, and *p* and *q* were the dimensions of the covariance matrix of correlated decoder outputs. A scalar-multiplied version of ρ was fit to λ. The components for which $\rho < \lambda$ were taken as 'signal' (Mitra and Pesaran, 1999), and their time series were included in the regression model. A procedure where we fixed the number (12) of selected components based on the group-average yielded identical results in terms of direction and significance of effects.

We quantified the contribution of the decoder correlations (*m* selected components, see above), stimulus decoders, and action decoders, to the prediction of the active rule (instructed rule task) or the graded belief about the rule (inferred rule). For instructed rule, we used logistic regression:

$$ln\left(\frac{p_{rule\,1}(t)}{1 - p_{rule\,2}(t)}\right) = \beta_0 + \sum_{i=1}^{m}\beta_i\, U_i(t) + \sum_{j=1}^{n}\beta_j\, S_j(t) + \sum_{k=1}^{p}\beta_k\, A_k(t) \qquad (13)$$

where *U* was the set of component time series (quantifying decoder correlations) and *i* indexed the components, *S* was the set of n stimulus decoder outputs from visual cortical areas indexed by *j*, and *A* was the set of *p* action decoder outputs from action-related areas indexed by *k*. $p_{rule\ 1}(t)$ indicated the probability that rule 1 was active at time point *t* and correspondingly for $p_{rule\ 2}(t)$. To account for the hemodynamic lag, the rule time series was convolved with the individual HRF estimate and binarized again (see *Correlation of stimulus- and action-selective voxel groups* above). For the inferred rule task, we used linear regression:

$$L(t) = \beta_0 + \sum_{i=1}^{m} \beta_i \, U_i(t) + \sum_{j=1}^{n} \beta_j \, S_j(t) + \sum_{k=1}^{p} \beta_k \, A_k(t) \qquad (14)$$

where *L(t)* was the belief time series extracted from the normative model fit for each participant, convolved with the participant-specific HRF (to account for hemodynamics).

We used a ten-fold cross-validation procedure to compute prediction accuracy. We divided the data into ten segments of equal length. These segments contained the to-be-predicted variable (HRF-convolved) and the regressor time series (i.e., the selected components as well as the local decoder outputs). For each fold, nine segments were used for training the prediction model (i.e., computing beta weights using eqs. 13 or 14), and the remaining segment was used for prediction. Prediction accuracy was quantified as the mean percentage of correctly predicted rule 1 and rule 2 (chance = 50%; instructed rule), or the correlation between predicted belief and observed belief (inferred rule).

Predictions during testing (in the held-out segment) were either obtained based on all regressors (*U, S, and A*), based on only decoder correlations (*U*), or based on either of the local decoder outputs (*S* or *A*). For each fold, the time series of all regressors, or of *U*, *S*, or *A* individually, that were measured in the testing segment, were multiplied with the beta weights estimated from the training data. The accuracies of these predictions were then calculated, and averaged across the ten folds.

Analysis of errors

Analyses of the relationship between decoder correlations and behavioral errors were only possible for the inferred rule, not the instructed rule task, because participants only made a considerable number of errors in the former condition. We computed the correlations between stimulus and action decoders separately for time windows from -6.8 to +6.8 s from the trials of the lower-order task (orientation judgments). This window corresponded to the minimum inter-trial interval and thus prevented overlap between successive trials. For this analysis, the stimulus and action decoders were trained on the evoked responses in instructed rule runs, and projected onto the inferred rule runs.

We then sorted the trials based on choice correctness (i.e. consistency of the choice with the active rule). There were between 25 and 74 (mean: 46; s.d.: 15) error trials per participant. Because there were fewer error trials, we computed decoder correlations for correct trials on a randomly sampled subset of trials that was equal in number to error trials. We repeated this procedure 1000 times and then averaged across iterations. We also sorted trials based on the consistency of the choice (and, thus, by inference: of the selected rule) with the model-derived belief parameter. There were between 10 and 55 (mean: 33; s.d.: 14) belief-inconsistent trials per participant. Finally, we compared the decoder correlations for each rule, as well as the

differences between rules, between these trial types (correct versus error, and consistent versus inconsistent with belief).

The relationship between correlated decoder output and the model decision noise parameter was assessed using a permutation-based correlation procedure. We selected time windows surrounding choices (see above), and computed the difference in decoder correlations for the two belief states (favoring rule 1 or rule 2), separately for each participant. This difference score was then correlated in turn with the individual noise parameters (V) estimated from the behavioral model fits using Spearman's rank correlation. Qualitatively identical results were obtained with using Pearson correlation.

For the purpose of prediction, we trained, on all participants except one, a linear regression model that captured the relationship between correlated decoder output and the model decision noise parameter. This regression model was then used to generate a prediction about the decision noise parameter in the held-out participant, based on correlated decoder output for that participant. This was repeated iteratively such that there was a prediction for all individual participants. The predicted value was then correlated across participants with the true model decision noise.

Decoding of rule or belief from local activity patterns

We used support vector classification (instructed rule) or support vector regression (inferred rule runs) to test if the local patterns of fMRI activity encoded the (belief about) active rule. These analyses were run for each ROI of the HCP-MMP1.0 parcellation (Glasser et al., 2016), thus covering the complete cortex. Decoding was performed in two different ways: 1) based on the evoked response to the cue that instructed the participant of the active rule (instructed rule runs), and 2) the continuous fMRI time series, non-residualized for evoked responses (both the instructed and inferred rule runs). All voxels from each ROI were included. We used the same ten-fold cross-validation procedure as used in other analyses (see sections: *Decoding of stimulus and action from patterns of evoked responses*, and *Prediction of active rule or belief from decoder correlations*). The classification accuracy of active rule in the held-out runs (instructed rule runs) and the correlation between SVR-predicted and model-estimated belief in the held-out runs (inferred rule task) was compared to chance (permutation tests, see below) and corrected for multiple comparisons using the false discovery rate (FDR; $q = 0.05$).

Correlation between rule / belief decoder output and stimulus-action decoder covariation

We correlated the continuous output of the rule decoder (instructed) or belief decoder (inferred), computed as above (section *Decoding of rule or belief from local activity patterns*), with the time-variant stimulus-action decoder coupling (computed via eq. 10). This was done for each stimulus-action ROI pair, and each ROI that showed significant prediction scores of rule or belief (orange areas in Fig. 7c). For bilateral ROIs (V1 and 6d) we first averaged the belief decoder output across the two hemispheres. Prior to computing the belief decoder output, we removed evoked responses due to the SR-cue in the instructed rule task. We averaged the correlation between belief / rule decoder output, and time-variant stimulus-action decoder output, across all stimulus-action ROI pairs to obtain a final correlation score. This was compared to chance using permutation testing (10,000 iterations).

Whole brain GLMs

We fitted two separate voxel-wise GLMs to compute statistical maps of the covariation of fMRI activity with a number of external trial variables (grating stimulus orientations, left and right button presses) or hidden computational variables extracted from the fitted behavioral model. Both GLMs contained four regressors. The first GLM contained one regressor for each unique stimulus-action combination. The second GLM contained the magnitude of belief ($|L|$) and of sensory evidence ($|LLR|$), change point probability (*CPP*), and all grating onsets (stick function). All regressors were convolved with the individual HRFs.

These GLMs were fit separately for each run, and the resulting beta weights were averaged across runs and sessions. Contrasts were tested using FSL's 'randomise' (10,000 iterations) and corrected for multiple comparisons using threshold-free cluster enhancement (Smith and Nichols, 2009).

Standard functional connectivity analysis

For a conventional functional connectivity analysis, we computed inter-regional correlations between the voxel-averaged ongoing activity of all ROIs, separately per active rule (instructed rule) or belief-favored rule (inferred rule). We then collapsed across the instructed and inferred rule runs and compared the correlations visual and action-related areas between rules, as we did for the correlations of decoder outputs from the same areas.

Statistics

Within-participant error bars

For within-participant comparisons, we discarded between-participant variance by mean-centering each participants' data prior to calculating the SEM (Cousineau, 2005). The purpose of this procedure was to ensure that the error bars depict only within-participant, across-condition, variance that was relevant for the statistical comparison.

Statistical tests

We used non-parametric permutation tests (Tibshirani and Efron, 1993) with 10,000 permutations for all statistical comparisons, unless mentioned otherwise. Specifically, single-participant correlations were compared with zero or between rules. The significance of decoding and prediction accuracies was tested by comparing the obtained percentage correct to 50% chance level (instructed rule) or the correlation coefficients between predicted and actual beliefs to zero (inferred rule).

For testing the across-participant correlation between decision noise V and correlation difference, we repeated the Spearman correlation 10,000 times, but with shuffled V values, yielding a null distribution of correlation coefficients. A p-value for the observed correlation was finally obtained by comparing that value to the null distribution.

References

Arieli A, Sterkin A, Grinvald A, Aertsen A (1996) Dynamics of ongoing activity: explanation of the large variability in evoked cortical responses. Science 273:1868-1871.

Bennur S, Gold JI (2011) Distinct representations of a perceptual decision and the associated oculomotor plan in the monkey lateral intraparietal area. The Journal of neuroscience 31:913-921.

Bode S, Haynes JD (2009) Decoding sequential stages of task preparation in the human brain. NeuroImage 45:606-613.

Bogacz R, Brown E, Moehlis J, Holmes P, Cohen JD (2006) The physics of optimal decision making: a formal analysis of models of performance in two-alternative forced-choice tasks. Psychol Rev 113:700-765.

Bondy AG, Haefner RM, Cumming BG (2018) Feedback determines the structure of correlated variability in primary visual cortex. Nature neuroscience 21:598-606.

Bonhoeffer T, Grinvald A (1991) Iso-orientation domains in cat visual cortex are arranged in pinwhieel-like patterns. Nature 253:429-431.

Brainard DH (1997) The psychophysics toolbox. Spatial Vision 10:433-436.

Büchel C, Friston K (2000) Assessing interactions among neuronal systems using functional neuroimaging. Neural Networks 13:871-882.

Cisek P, Kalaska JF (2005) Neural correlates of reaching decisions in dorsal premotor cortex: specification of multiple direction choices and final selection of action. Neuron 45:801-814.

Cocuzza CV, Ito T, Schultz D, Bassett DS, Cole MW (2020) Flexible Coordinator and Switcher Hubs for Adaptive Task Control. The Journal of neuroscience 40:6949-6968.

Cohen MR, Newsome WT (2008) Context-dependent changes in functional circuitry in visual area MT. Neuron 60:162-173.

Cole MW, Ito T, Braver TS (2016) The Behavioral Relevance of Task Information in Human Prefrontal Cortex. Cerebral cortex 26:2497-2505.

Cousineau D (2005) Confidence intervals in within-subject designs: A simpler solution to Loftus and Masson's method. Tutorials in Quantitative Methods for Psychology 1:42-45.

Coutanche MN, Thompson-Schill SL (2013) Informational connectivity: identifying synchronized discriminability of multi-voxel patterns across the brain. Frontiers in human neuroscience 7:15.

Dale AM (1999) Optimal experimental design for fMRI. Human Brain Mapping 8:109-114.

de Gee JW, Colizoli O, Kloosterman NA, Knapen T, Nieuwenhuis S, Donner T (2017) Dynamic modulation of decision biases by brainstem arousal systems. eLife 6:e23232.

Donner TH, Siegel M, Fries P, Engel AK (2009) Buildup of choice-predictive activity in human motor cortex during perceptual decision making. Current biology : CB 19:1581-1585.

Drugowitsch J, Wyart V, Devauchelle AD, Koechlin E (2016) Computational Precision of Mental Inference as Critical Source of Human Choice Suboptimality. Neuron 92:1398-1411.

Duan CA, Pagan M, Piet AT, Kopec CD, Akrami A, Riordan AJ, Erlich JC, Brody CD (2021) Collicular circuits for flexible sensorimotor routing. Nature neuroscience 24:1110-1120.

Durstewitz D, Vittoz NM, Floresco SB, Seamans JK (2010) Abrupt transitions between prefrontal neural ensemble states accompany behavioral transitions during rule learning. Neuron 66:438-448.

Filipowicz AL, Glaze CM, Kable JW, Gold JI (2020) Pupil diameter encodes the idiosyncratic, cognitive complexity of belief updating. eLife 9.

Fiser J, Berkes P, Orban G, Lengyel M (2010) Statistically optimal perception and learning: from behavior to neural representations. Trends in cognitive sciences 14:119-130.

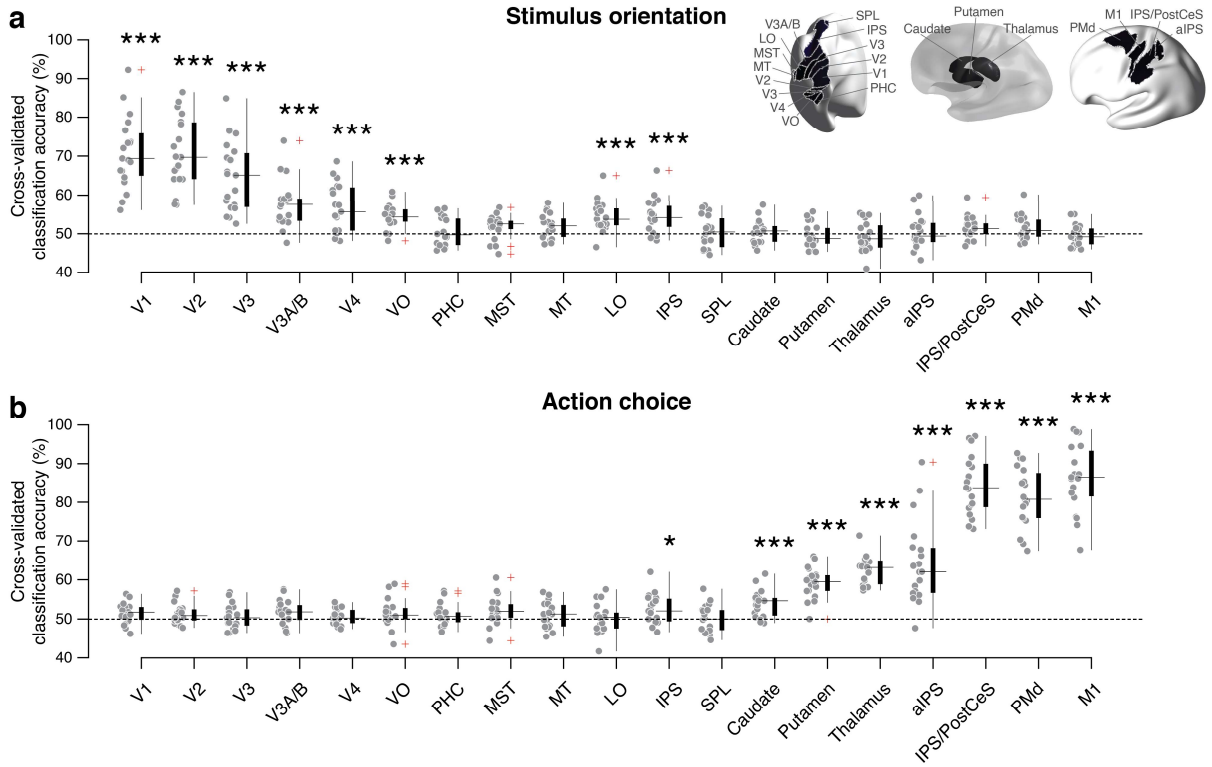
Fox MD, Raichle ME (2007) Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. Nature reviews Neuroscience 8:700-711.

Fox MD, Snyder AZ, Zacks JM, Raichle ME (2006) Coherent spontaneous activity accounts for trial-to-trial variability in human evoked brain responses. Nature neuroscience 9:23-25.
Fusi S, Asaad WF, Miller EK, Wang XJ (2007) A neural circuit model of flexible sensorimotor mapping: learning and forgetting on multiple timescales. Neuron 54:319-333.
Gerstein GL, Perkel DH (1969) Simultaneously recorded trains of action potentials: Analysis and functional interpretation. Science 164:828-830.
Glasser MF, Coalson TS, Robinson EC, Hacker CD, Harwell J, Yacoub E, Ugurbil K, Andersson J, Beckmann CF, Jenkinson M, Smith SM, Van Essen DC (2016) A multi-modal parcellation of human cerebral cortex. Nature 536:171-178.
Glaze CM, Kable J, Gold JI (2015) Normative evidence accumulation in unpredictable environments. eLife 4:e08825.
Glover GH, Li T-Q, Ress D (2000) Image-based method for retrospective correction of physiological motion effects in fMRI: RETROICOR. Magnetic Resonance in Medicine 44:162-167.
Gold JI, Shadlen MN (2003) The influence of behavioral context on the representation of a perceptual decision in developing oculomotor commands. The Journal of neuroscience 23:632-651.
Gold JI, Shadlen MN (2007) The neural basis of decision making. Annual review of neuroscience 30:535-574.
Haefner RM, Berkes P, Fiser J (2016) Perceptual Decision-Making as Probabilistic Inference by Neural Sampling. Neuron 90:649-660.
Hanks TD, Kopec CD, Brunton BW, Duan CA, Erlich JC, Brody CD (2015) Distinct relationships of parietal and prefrontal cortices to evidence accumulation. Nature 520:220-223.
Harvey CD, Coen P, Tank DW (2012) Choice-specific sequences in parietal cortex during a virtual-navigation decision task. Nature 484:62-68.
Haynes JD, Rees G (2005) Predicting the orientation of invisible stimuli from activity in human primary visual cortex. Nature neuroscience 8:686-691.
Haynes JD, Rees G (2006) Decoding mental states from brain activity in humans. Nature reviews Neuroscience 7:523-534.
Haynes JD, Sakai K, Rees G, Gilbert S, Frith C, Passingham RE (2007) Reading hidden intentions in the human brain. Current biology : CB 17:323-328.
Heinzle J, Wenzel MA, Haynes JD (2012) Visuomotor functional network topology predicts upcoming tasks. The Journal of neuroscience 32:9960-9968.
Hipp JF, Hawellek DJ, Corbetta M, Siegel M, Engel AK (2012) Large-scale cortical correlation structure of spontaneous oscillatory activity. Nature neuroscience 15:884-890.
Honey CJ, Sporns O, Cammoun L, Gigandet X, Thiran JP, Meuli R, Hagmann P (2009) Predicting human resting-state functional connectivity from structural connectivity. Proceedings of the National Academy of Sciences of the United States of America 106:2035-2040.
Ito T, Yang GR, Laurent P, Schultz DH, Cole MW (2022) Constructing neural network models from brain data reveals representational transformations linked to adaptive behavior. Nature communications 13:673.
Jenkinson M, Beckmann CF, Behrens TE, Woolrich MW, Smith SM (2012) Fsl. NeuroImage 62:782-790.
Kamitani Y, Tong F (2005) Decoding the visual and subjective contents of the human brain. Nature neuroscience 8:679-685.
Kenet T, Bibitchkov D, Dsodyks M, Grinvald A, Arieli A (2003) Spontaneously emerging cortical representations of visual attributes. Nature 425:954-956.
Kiani R, Shadlen MN (2009) Representation of confidence associated with a decision by neurons in the parietal cortex. Science 324:759-764.
Kikumoto A, Mayr U (2020) Conjunctive representations that integrate stimuli, responses, and rules are critical for action selection. PNAS 117:10603-10608.

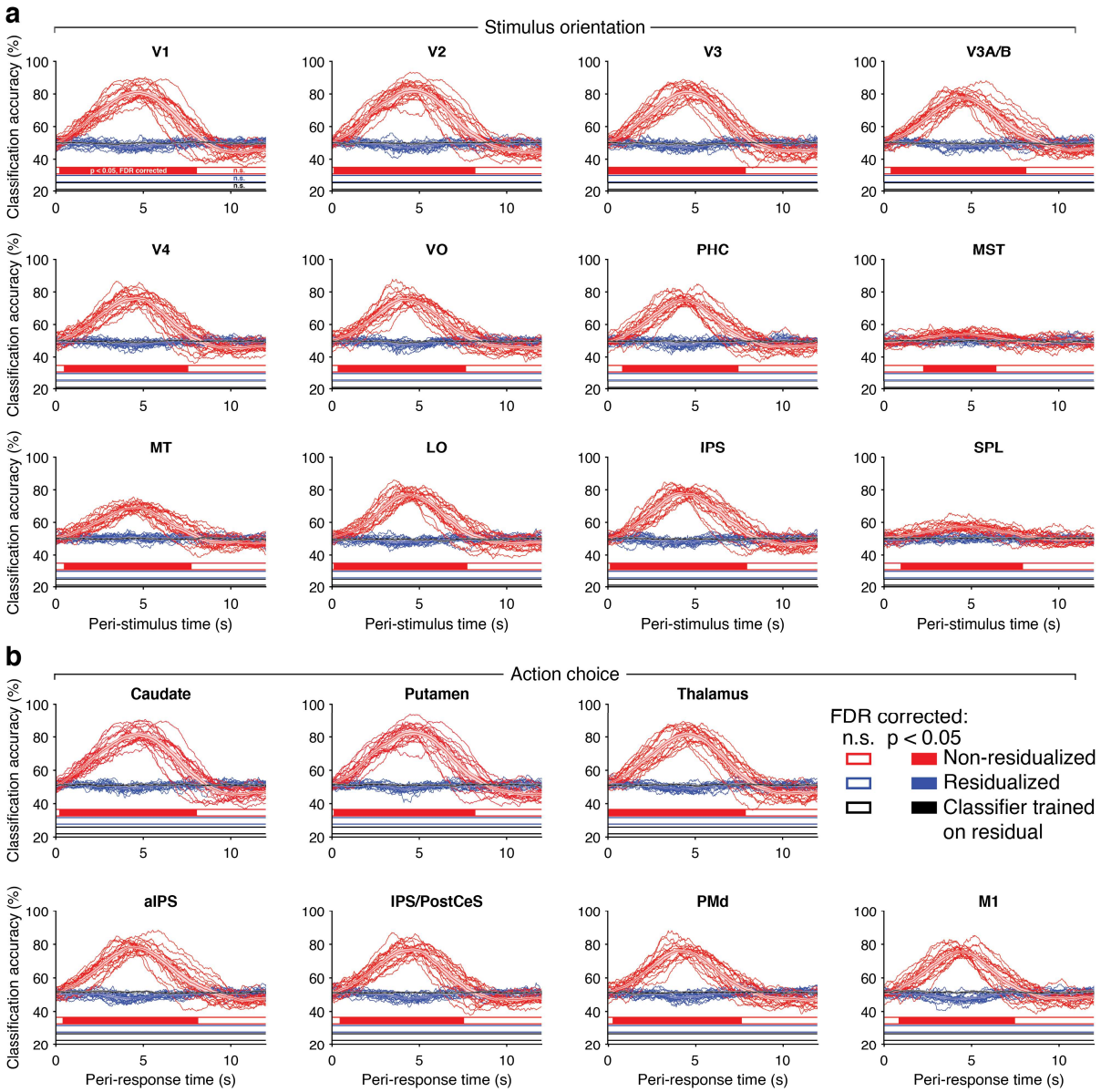
Kriegeskorte N, Bandettinin B (2007) Analyzing for information, not activation, to exploit highresolution fMRI. NeuroImage 38:649-662.
Lurie DJ et al. (2019) Questions and controversies in the study of time-varying functional connectivity in resting fMRI. Network Neuroscience 4:30-69.
Mante V, Sussillo D, Shenoy KV, Newsome WT (2013) Context-dependent computation by recurrent dynamics in prefrontal cortex. Nature 503:78-84.
Mastrogiuseppe F, Ostojic S (2018) Linking Connectivity, Dynamics, and Computations in Low-Rank Recurrent Neural Networks. Neuron 99:609-623 e629.
McGuire JT, Nassar MR, Gold JI, Kable JW (2014) Functionally dissociable influences on learning rate in a dynamic environment. Neuron 84:870-881.
Miller EK, Cohen JD (2001) An integrative theory of prefrontal cortex function. Annual review of neuroscience 24:167-202.
Mitra PP, Pesaran B (1999) Analysis of dynamic brain imaging data. Biophysical Journal 76:691-708.
Monsell S (2003) Task switching. Trends in cognitive sciences 7:134-140.
Mumford JA, Turner BO, Ashby FG, Poldrack RA (2012) Deconvolving BOLD activation in event-related designs for multivoxel pattern classification analyses. NeuroImage 59:2636-2643.
Munoz DP, Everling S (2004) Look away: the anti-saccade task and the voluntary control of eye movement. Nature reviews Neuroscience 5:218-228.
Murphy PR, Wilming N, Hernandez-Bocanegra DC, Prat-Ortega G, Donner TH (2021) Adaptive circuit dynamics across human cortex during evidence accumulation in changing environments. Nature neuroscience 24:987-997.
Pfeffer T, Ponce-Alvarez A, Tsetsos K, Meindertsma T, Gahnström CJ, van den Brink RL, Nolte G, Engel AK, Deco G, Donner TH (2021) Circuit mechanisms for the chemical modulation of cortex-wide network interactions and behavioral variability. Science Advances 7:eabf5620.
Picard N, Strick PL (2001) Imaging the premotor areas. Current opinion in neurobiology 11:663-672.
Piet AT, El Hady A, Brody CD (2018) Rats adopt the optimal timescale for evidence integration in a dynamic environment. Nature communications 9:4265.
Pinto L, Rajan K, DePasquale B, Thiberge SY, Tank DW, Brody CD (2019) Task-Dependent Changes in the Large-Scale Dynamics and Necessity of Cortical Regions. Neuron 104:810-824 e819.
Purcell BA, Kiani R (2016) Hierarchical decision processes that operate over distinct timescales underlie choice and changes in strategy. Proceedings of the National Academy of Sciences of the United States of America 113:E4531-4540.
Reynolds JH, Pasternak T, Desimone R (2000) Attention increases sensitivity of V4 neurons. Neuron 26:703-714.
Rigotti M, Barak O, Warden MR, Wang XJ, Daw ND, Miller EK, Fusi S (2013) The importance of mixed selectivity in complex cognitive tasks. Nature 497:585-590.
Saez A, Rigotti M, Ostojic S, Fusi S, Salzman CD (2015) Abstract Context Representations in Primate Amygdala and Prefrontal Cortex. Neuron 87:869-881.
Sakai K (2008) Task set and prefrontal cortex. Annual review of neuroscience 31:219-245.
Sarafyazd M, Jazayeri M (2019) Hierarchical reasoning by neural circuits in the frontal cortex. Science 364.
Shadlen MN, Kiani R (2013) Decision making as a window on cognition. Neuron 80:791-806.
Siegel M, Buschman TJ, Miller EK (2015) Cortical information flow during flexible sensorimotor decisions. Science 384:1352-1355.
Smith SM, Nichols TE (2009) Threshold-free cluster enhancement: addressing problems of smoothing, threshold dependence and localisation in cluster inference. NeuroImage 44:83-98.
Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TE, Johansen-Berg H, Bannister PR, De Luca M, Drobnjak I, Flitney DE, Niazy RK, Saunders J, Vickers J, Zhang Y, De Stefano N, Brady JM,

Matthews PM (2004) Advances in functional and structural MR image analysis and implementation as FSL. NeuroImage 23 Suppl 1:S208-219.
Stoet G, Snyder LH (2004) Single neurons in posterior parietal cortex of monkeys encode cognitive set. Neuron 42:1003-1012.
Tibshirani RJ, Efron B (1993) An introduction to the bootstrap. Monographs on statistics and applied probability 57:1-436.
Tsetsos K, Pfeffer T, Jentgens P, Donner TH (2015) Action Planning and the Timescale of Evidence Accumulation. PloS one 10:e0129473.
van den Brink RL, Nieuwenhuis S, Donner TH (2018) Amplification and Suppression of Distinct Brainwide Activity Patterns by Catecholamines. The Journal of neuroscience 38:7476-7491.
van den Brink RL, Pfeffer T, Donner TH (2019) Brainstem Modulation of Large-Scale Intrinsic Cortical Activity Correlations. Frontiers in human neuroscience 13.
van den Brink RL, Pfeffer T, Warren CM, Murphy PR, Tona KD, van der Wee NJ, Giltay E, van Noorden MS, Rombouts SA, Donner TH, Nieuwenhuis S (2016) Catecholaminergic Neuromodulation Shapes Intrinsic MRI Functional Connectivity in the Human Brain. The Journal of neuroscience 36:7865-7876.
Vincent JL, Patel GH, Fox MD, Snyder AZ, Baker JT, Van Essen DC, Zempel JM, Snyder LH, Corbetta M, Raichle ME (2007) Intrinsic functional architecture in the anaesthetized monkey brain. Nature 447:83-86.
Wallis JD, Anderson KC, Miller EK (2001) Single neurons in prefrontal cortex encode abstract rules. Nature 411:953-956.
Wang L, Mruczek RE, Arcaro MJ, Kastner S (2015) Probabilistic Maps of Visual Topography in Human Cortex. Cerebral cortex 25:3911-3931.
Weiss A, Chambon V, Lee JK, Drugowitsch J, Wyart V (2021) Interacting with volatile environments stabilizes hidden-state inference and its brain signatures. Nature communications 12:2228.
Wetzels R, Wagenmakers EJ (2012) A default Bayesian hypothesis test for correlations and partial correlations. Psychon Bull Rev 19:1057-1064.
Wilming N, Murphy PR, Meyniel F, Donner TH (2020) Large-scale dynamics of perceptual decision information across human cortex. Nature communications 11:5109.
Yang GR, Joglekar MR, Song HF, Newsome WT, Wang XJ (2019) Task representations in neural networks trained to perform many cognitive tasks. Nature neuroscience 22:297-306.
Yousry TA, Schmid UD, Alkadhi H, Schmidt D, Peraud A, Buettner A, Winkler P (1997) Localization of the motor hand area to a knob on the precentral gyrus. Brain 120:141-157.
Zamani Esfahlani F, Jo y, Faskowitz J, Byrge L, Kennedy DP, Sporns O, Betzel RF (2020) High-amplitude cofluctuations in cortical activity drive functional connectivity. PNAS 117:28393-28401.
Zhang J, Kriegeskorte N, Carlin JD, Rowe JB (2013) Choosing the rules: distinct and overlapping frontoparietal representations of task rules for perceptual decisions. The Journal of neuroscience 33:11852-11862.

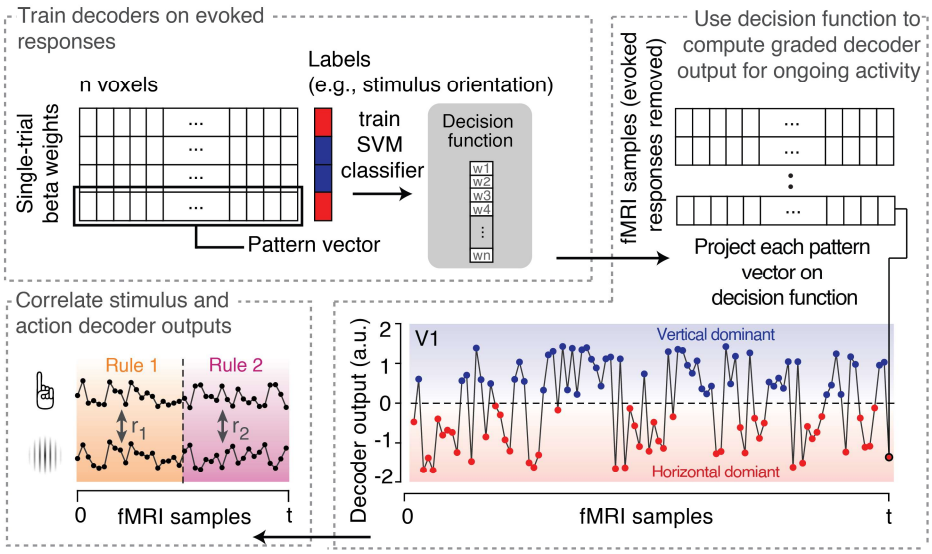
Supplementary Figures



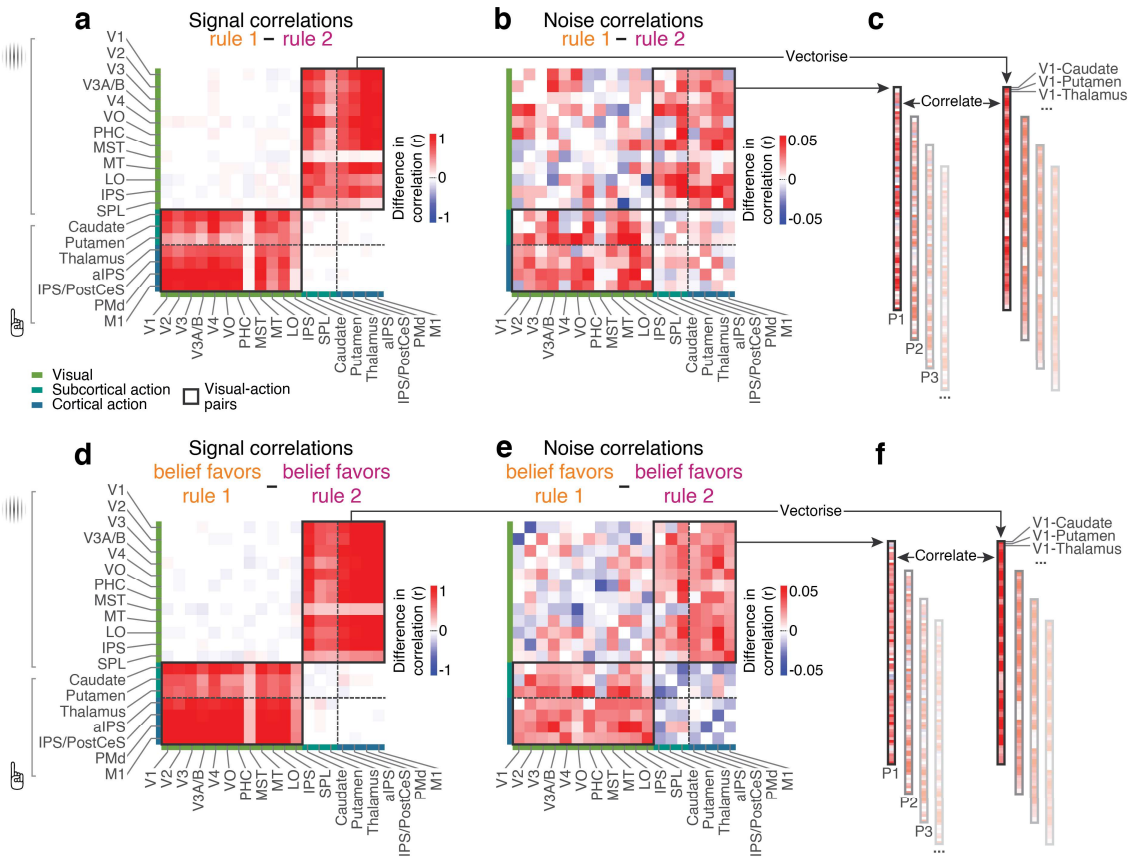
Supplementary Fig. 1. Cross-validated classification accuracy of a. stimulus orientation, and b. action choice. As expected, classification performance for stimulus orientation was highest for early visual cortex, decreased along the visual cortical hierarchy and was at chance in action-related regions. Conversely, classification performance for action choice increased along the visuo-motor pathway, and peaked in primary motor cortex. Gray dots show individual participants. The whiskers extend to the most extreme values non-outlier values of each distribution. Outliers are shown in red. Asterisks show FDR-corrected p-values: *p<0.05, **p<0.01, ***p<0.001



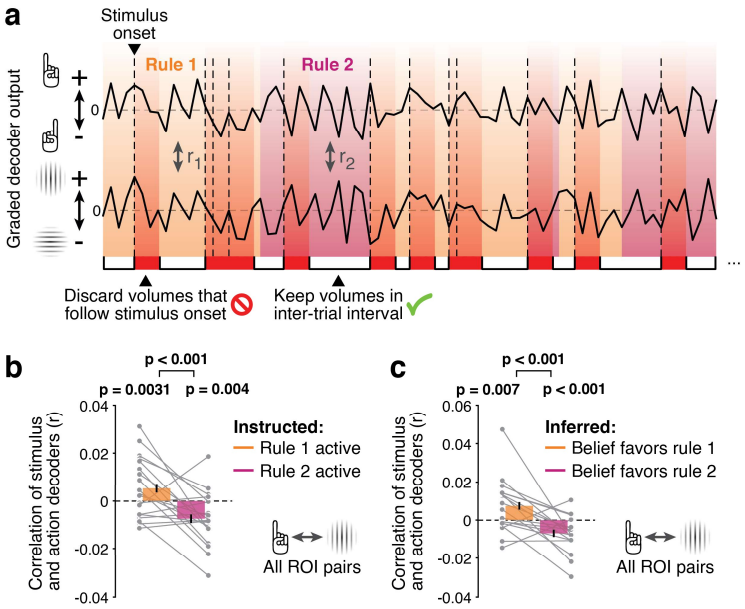
Supplementary Fig. 2. Effective removal of feature-specific evoked responses. a. Mean stimulus orientation decoding performance time-locked to grating onset for visual field maps. **b.** Mean action decoding performance time-locked to button presses for action-related ROIs. Red, evoked decoding performance computed based on full voxel time series. Blue, decoding performance of the classifiers trained on non-residualized data, tested on residualized data (i.e., time series after removal of evoked responses; see Materials and Methods). At no time point did decoding performance for residuals deviate significantly from chance in any ROI. Black, as blue, but showing decoding performance of classifiers trained and tested on residualized data (only group average). Again, at no time point did residual decoding performance deviate from chance in any ROI. Thin lines, individual participants; shaded areas, within-participant SEM. Together, the results indicate that our analysis effectively removed any feature-specific evoked responses (i.e., discriminant patterns systematically time-locked to stimulus onsets or button presses), regardless of whether the classifiers were trained on the non-residualized or residualized data. This validates the residual activity time courses used for correlation analyses as estimates of ongoing fluctuations in activity (i.e., activity unrelated to the trial events).



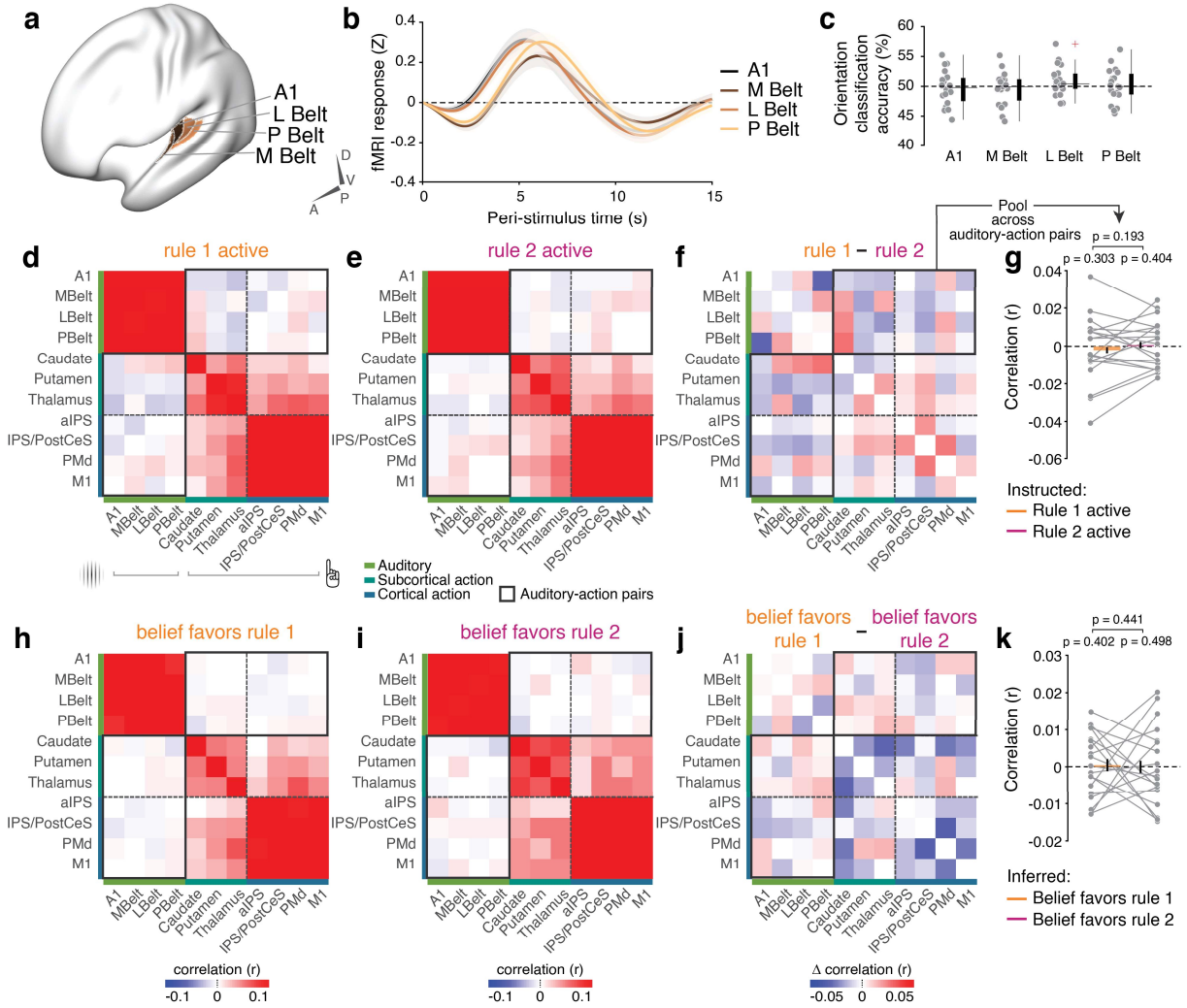
Supplementary Fig. 3. Analysis approach for assessing correlations between stimulus- and action-selective neural signals. First support vector machine decoders were trained on single-trial estimates of the multi-voxel response patterns evoked during the trials of the primary perceptual choice task. Different decoders were trained to discriminate the two stimulus categories, or the two actions, respectively. The resulting decision function (i.e., weight vector) was projected onto the full time-series fMRI data, from which evoked responses had been removed (Materials and Methods). This projection resulted in a scalar time-series quantifying the graded output of the decoder, indicating if the spontaneous activity patterns at a given point in time the tended towards one of two stimuli (or actions). Here, this is shown for the orientation decoder applied to V1 as an example. Finally, the outputs of the stimulus decoders were correlated with the outputs of the action decoders, separately for segments in time corresponding to the two rules, after shifting by hemodynamic delay (Materials and Methods).



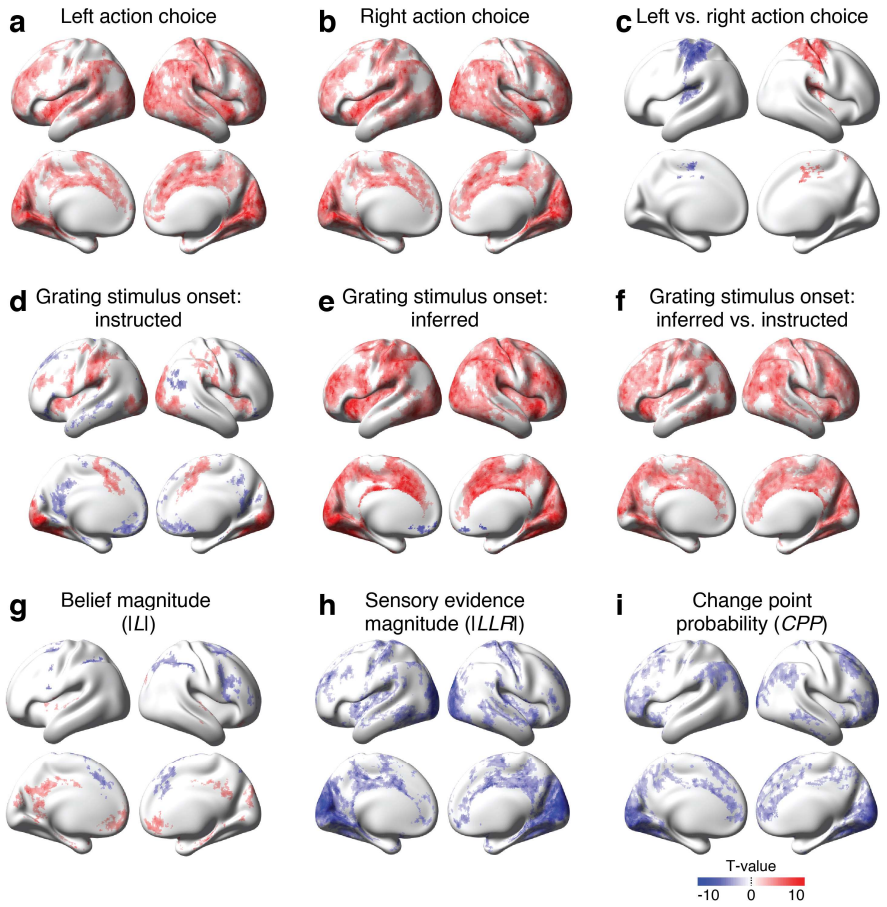
Supplementary Fig. 4. Rationale of control analysis of evoked responses. a. Signal correlation pattern (decoder outputs for evoked responses) in instructed rule task. **b.** Noise correlation (decoder outputs for intrinsic signal fluctuations) in instructed rule task (data from Fig. 3c; see Main Text), for comparison with panel a. **c.** The edges in panels a and b that link the stimulus and action regions were vectorized, and correlated with one another to quantify their spatial similarity. This yielded a correlation coefficient for each participant (Px), that was compared to zero. **d.** Same as panel a, but for the inferred rule task, and split by dominant belief rather than active rule. **e.** Data from Fig. 5c (see Main Text), for comparison with panel d. **f.** Same as panel c, but for the inferred rule task, and split by dominant belief rather than active rule.



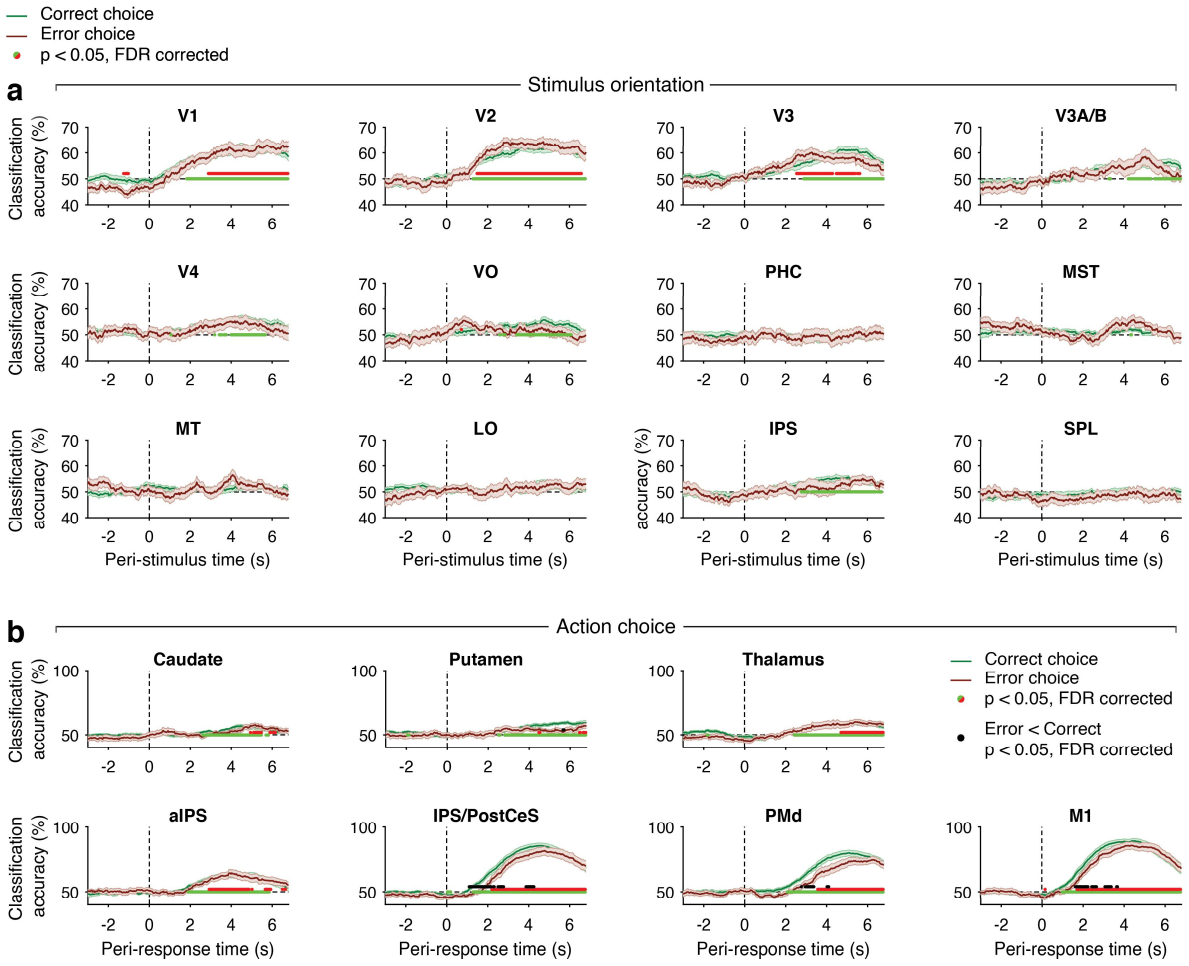
Supplementary Fig. 5. Control analysis excluding data segments following stimulus onset. a. Schematic of analysis approach. The three consecutive volumes (5.7 s) following the onset of each stimulus were discarded from the graded decoder output time series, preserving only the data segments in the inter-trial intervals, prior to correlating stimulus and action decoder outputs. **b.** Correlation of the graded outputs of a stimulus orientation decoders and action decoders, collapsed across all visual-action ROI pairs, in the instructed rule task. The rule-specific pattern of correlations of decoder outputs remains the same as for the main analysis shown in Fig. 3e, indicating that the effect is expressed in the inter-trial intervals. Thin lines indicate individual participants. Bars show the group mean. Error bars show the within-participant SEM. **c.** Same as panel b, but for the inferred rule task.



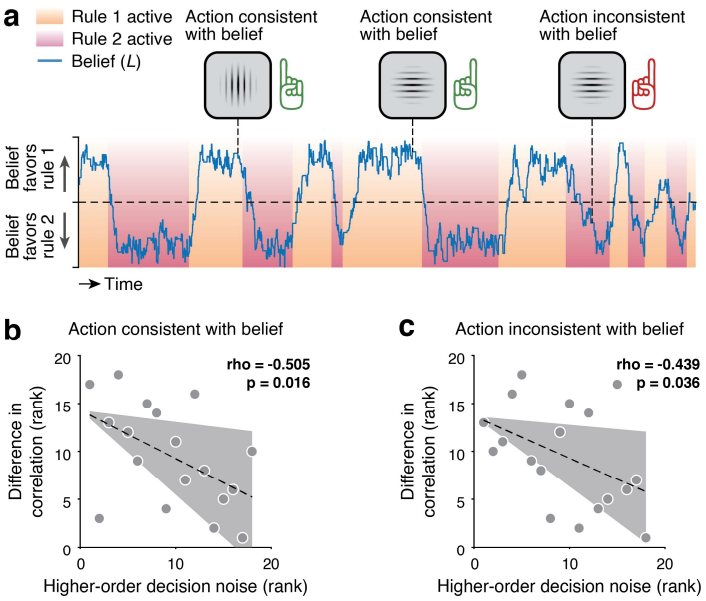
Supplementary Fig. 6. Control analysis using auditory cortex instead of visual cortex. a. Regions of interest in auditory cortex that were used to replace visual cortical regions, and entered into stimulus-classification analysis. **b.** Stimulus-locked responses in auditory regions obtained with deconvolution. Auditory regions showed robust responses following trial onset. Error bars, SEM. **c.** Cross-validated stimulus orientation decoder performance. The responses in auditory cortices to stimulus onset were uninformative about stimulus orientation (as expected). Gray dots show individual participants. The whiskers extend to the most extreme values non-outlier values of each distribution. Outliers are shown in red. **d.** Correlation matrix of graded decoder output for stimulus and action in the instructed rule task when rule 1 was active. **e.** Same as panel b, but for rule 2. **f.** The difference in coupling between rule 1 and rule 2. The gray outline delineates the edges of the matrix that link stimulus and action ROIs. **g.** Correlation of the graded outputs of a stimulus orientation decoder (trained on auditory cortex) and action decoder. Thin lines indicate individual participants. Bars show the group mean. Error bars show the within-participant SEM. **h.** Correlation matrix of graded decoder output for stimulus and action in the inferred rule task when model-derived belief favored rule 1. **i.** Same as panel b, but for rule 2. **j.** The difference in coupling between rule 1 and rule 2. The gray outline delineates the edges of the matrix that link stimulus and action ROIs. **k.** Correlation of the graded outputs of a stimulus orientation decoder (trained on auditory cortex) and action decoder. Thin lines indicate individual participants. Bars show the group mean. Error bars show the within-participant SEM.



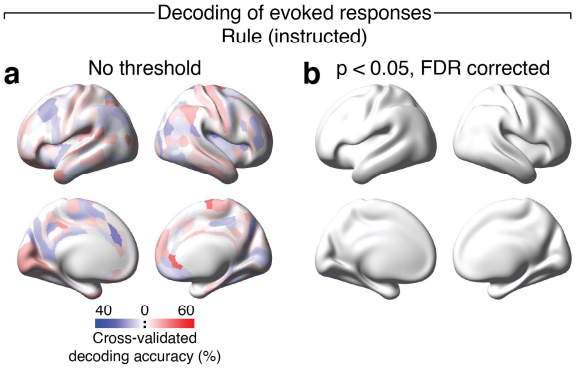
Supplementary Fig. 7. Voxel-level regression analysis of fMRI responses evoked by action, stimulus onset, and computational variables. a. Left hand button presses. **b.** Right hand button presses. **c.** The difference between left and right hand button presses, showing the expected hand movement-selective lateralization of activity within sensory-motor cortical areas including the action-selective regions of interest used here (PMd, IPS/PostCes, aIPS.. **d.** Grating stimulus onset in instructed rule task. **e.** Grating stimulus onset in inferred rule task. **f.** Difference between stimulus-locked responses from instructed and inferred rule tasks. This widespread differential activity likely reflects the increased attentional demand in the inferred rule task. **g.** Magnitude of model-derived belief, showing positive effects in posterior cingulate and ventromedial prefrontal cortices and negative effects in posterior parietal and lateral prefrontal cortices. This cortical distribution resembles the sign-flipped map of effects previously shown for relative uncertainty in a similar learning task (McGuire et al, 2014). This can be understood by noting that belief magnitude in our task roughly corresponds to the complement of relative uncertainty (Murphy et al., 2021). **h.** The magnitude of the sensory evidence in the form of a log-likelihood ratio between rules. **i.** Change point probability, the posterior probability that a change in rule has just occurred given the sensory evidence and belief. Panels a-f were computed in one regression model. Panels g-i were computed in another regression model. All colored regions indicate significant co-variation with local fMRI responses, corrected for multiple comparisons using threshold-free cluster enhancement.



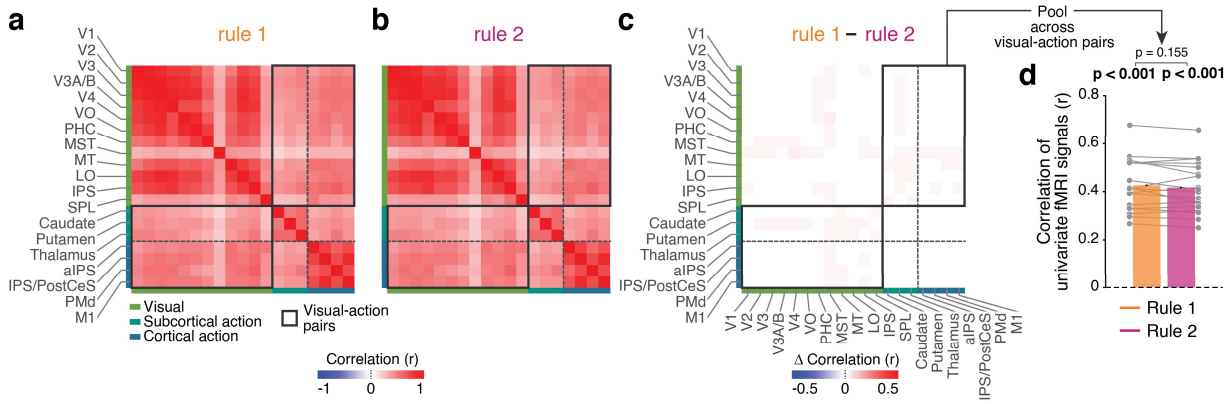
Supplementary Fig. 8. Cross-validated classification accuracy on correct and error trials in inferred rule. a. Cross-validated classification accuracy for stimulus orientation in visual cortical areas. Stimulus decoding was significant on, and did not differ between, correct and error trials for early visual cortex (V1-V3). In higher-tier visual cortical area, stimulus decoding was generally weaker, but nowhere was there a difference between correct and error trials. **b.** Cross-validated classification accuracy for action in action-related brain regions. Again, action decoding was significant for both trial categories in all regions. Yet, action decoding was significantly weaker on error than correct trials within IPS/PostCeS, PMd, and M1. Overall, this pattern of decoding results indicates that noise corrupting the sensory representation of stimulus orientation judgement was a negligible factor limiting behavioral performance of the inferred rule task, in line with the near-perfect performance based on the same stimuli in instructed rule. At the same time, erroneously chosen actions were encoded less precisely in action-related cortical regions, similar to effects observed for high vs. low decision confidence in motor cortex (Wilming et al, 2021). Error bars show the within-participant SEM. For stimulus decoding in the visual ROIs, no significant differences were found between correct and error trials.



Supplementary Fig. 9. Relationship between correlated variability patterns and behavioral performance. **a.** Choice trials were divided up based on if the action of the participant was consistent with- or inconsistent with the sign of the model-derived belief variable. **b.** Across-participant correlation between the model-derived noise estimates and the difference in the co-fluctuations of stimulus and action codes between the two belief states, for choices that were consistent with the belief parameter. Shaded area, 95% CI (bootstrapped) around the least squares regression line. **c.** Same as panel b, but for choices that were inconsistent with belief.



Supplementary Fig. 10. Local encoding of rule from the evoked responses to the SR-mapping cue. Maps show cross-validated classification accuracies. **a.** Active rule (instructed rule task), decoded by means of support vector machine). **b.** Areas showing significant decoding. The mapping rule could not be decoded from the cue-evoked responses in any area, indicating that evoked responses did not drive decoding of continuous signals (Fig. 7 in main text).



Supplementary Fig. 11. Results of a conventional correlation analysis based on region-average signals. a. Inter-regional activity correlation when rule 1 was active / belief favored rule 1. **b.** Same as panel a, but for rule 2. **c.** The difference between rule 1 and rule 2. **d.** Inter-regional correlation strength for the edges in panel c that link the visual and motor ROIs. In contrast to the correlation of graded outputs of stimulus and action decoders, this analysis yielded no sign flip between mapping rules. Gray lines show the individual participants. Bars show the group mean. Error bars show the within-participant SEM.