Behavioral/Cognitive

Dynamics of Hierarchical Task Representations

Dillan Cellier,[1,2,3] **Isaac T. Petersen,**[1,3] **and Kai Hwang**[1,2,3,4]

[1]Department of Psychological and Brain Sciences, University of Iowa, Iowa City, Iowa 52242, [2]Cognitive Control Collaborative, University of Iowa, Iowa City, Iowa 52242, [3]Iowa Neuroscience Institute, University of Iowa, Iowa City, Iowa 52242, and [4]Department of Psychiatry, University of Iowa, Iowa City, Iowa 52242

Task representations are critical for cognitive control and adaptive behavior. The hierarchical organization of task representations allows humans to maintain goals, integrate information across varying contexts, and select potential responses. In this study we characterized the structure and interactive dynamics of task representations that facilitate cognitive control. Human participants (both males and females) performed a hierarchical task that required them to select a response rule while considering the contingencies from different contextual inputs. By applying time- and frequency-resolved representational similarity analysis to human electroencephalography data, we characterized properties of task representations that are otherwise difficult to observe. We found that participants formed multiple representations of task-relevant contexts and features from the presented stimuli, beyond simple stimulus–response mappings. These disparate representations were hierarchically structured, with higher-order contextual representations dominantly influencing subordinate representations of task features and response rules. Furthermore, this cascade of top-down interactions facilitated faster responses. Our results describe key properties of task representations that support hierarchical cognitive control.

Key words: **cognitive flexibility; EEG; hierarchical cognitive control; task representations; working memory**

Significance Statement

Humans can adjust their actions in response to contingencies imposed by the environment. Although it has long been hypothesized that this ability depends on mental representations of tasks, the neural dynamics of task representations have been difficult to characterize. Our study used electroencephalography data from human participants to demonstrate the neural organization and interactive dynamics of task representations. Our results revealed a top-down, hierarchically organized representational structure that encoded multiple contexts and features from the environment. To support cognitive control, higher-level contextual representations influenced subordinate representations of task-relevant features and potential responses, facilitating response selection in a context-dependent manner. Our results provide direct evidence of organizational properties of task representations, which are cornerstones of cognitive control theories.

Introduction

Human behavior is adaptive to different contexts. Our actions are not governed by immutable rules but can be adjusted to meet changing conditions—or contextual contingencies—from the environment. Cognitive control enables such adaptive behavior by maintaining task representations of goals, context, and potential responses (Sakai, 2008; Schumacher and Hazeltine, 2016; Cohen, 2017). Theses maintained representations influence perceptual, motor, and cognitive processes to align actions with goals and context (D'Esposito, 2007), and thus task representations are an integral component of cognitive control.

To support cognitive control, task representations should encode multiple features and contexts from the environment to allow individuals to consider and integrate information from different sources and select the most appropriate action. For example, representations of abstract, contextual goals (e.g., drive cautiously near a playground) can influence subordinate representations of intermediate goals (e.g., speed of driving) and immediate actions (e.g., step on or off the gas pedal). Thus, task representations can be hierarchically organized (Botvinick, 2008; Badre and Nee, 2018). This hierarchical structure is advantageous when subjects must generate different responses to the same stimuli depending on different contexts.

Although it is thought that the human brain maintains task representations to facilitate cognitive control (Braver, 2012; Badre et al., 2021), there is still limited neural evidence for the dynamics and structure of hierarchical task representations.

Received Jan. 27, 2022; revised July 8, 2022; accepted Aug. 15, 2022.

Author contributions: D.C. and K.H. designed research; D.C. and K.H. performed research; D.C., I.T.P., and K.H. contributed unpublished reagents/analytic tools; D.C., I.T.P., and K.H. analyzed data; and D.C. and K.H. wrote the paper.

D.C. and K.H. were supported by the National Institute of Mental Health (Grant R01MH122613), and I.T.P. was supported by the Eunice Kennedy Shriver National Institute of Child Health and Human Development (Grant HD098235).

The authors declare no competing financial interests

Correspondence should be addressed to Kai Hwang at kai-hwang@uiowa.edu.

https://doi.org/10.1523/JNEUROSCI.0233-22.2022

In the laboratory, hierarchical cognitive control is often studied by administering tasks that involve multiple levels of contextual contingencies. For example, in a simple stimulus–response (S–R) task the correct response can be directly determined by the stimulus, whereas in a hierarchical task, subjects are required to determine the correct response by considering the contingencies among context, stimulus, and response. In tasks with varying levels of contingencies, functional magnetic resonance imaging (fMRI) studies have identified caudal frontal regions to be associated with executing simple S–R associations, whereas rostral midlateral prefrontal regions are associated with performing hierarchical tasks (Badre and D'Esposito, 2007). However, most existing studies have used univariate analyses and neglected the distributed and multivariate nature of neural representations (Diedrichsen and Kriegeskorte, 2017; Waskom and Wagner, 2017; Kikumoto and Mayr, 2020). Moreover, it is difficult to attribute differences in univariate response amplitudes to differences in task representations, given other components of task designs that differ between simple S–R tasks and hierarchical tasks (e.g., differences in stimuli complexity, difficulty, task instruction, and response requirement). Different levels of representations exist within a hierarchical task representation, yet the interactions between these different representational levels are not well understood.

The goal of the current study was to obtain neural evidence for the hierarchical organization and interactive dynamics of task representations. Specifically, we ask the following questions: First, can we decode the neural representations of task-relevant contexts and features? Second, can we determine how different components of task representations are structured? Third, can we characterize the interactive dynamics of these different components, leading to appropriate behavioral responses? To address these questions, we conducted an EEG study in which human subjects performed a hierarchical control task. Subjects selected the correct response for a match-nonmatch judgment while considering the contingencies from different contexts. We analyzed both behavioral and EEG data to probe task representations. For the behavioral data, we used a transitional reaction time (RT) analysis to determine how different models of task representations can influence switch costs in RT (Dykstra et al., 2022). To obtain neural evidence of task representations, we performed representational similarity analysis (RSA) on EEG data to decode task-relevant contexts, features, and response representations. To determine the structure of task representations, we performed Granger causality analysis (GCA) to test the dynamics of hierarchical interactions between representations. To determine the behavioral relevance of task representations, we conducted a structural equation modeling (SEM) analysis to determine how representational interactions influence trial-by-trial RT. Our results indicate that task representations are hierarchically structured, through which the highest level contextual information influences subordinate representations before a response, and the strength of the response-rule representation correlates with trial-by-trial variability in performance. These results reveal the structure and dynamics of task representations that are critical for hierarchical cognitive control.

Materials and Methods

Subjects. Forty healthy subjects (20 male, 18–35 years old, mean age 22.4) were recruited from the student body and surrounding areas of the University of Iowa. Subjects gave informed consent, and the study procedure was approved by the University of Iowa Institutional Review Board. Subjects were compensated with either money or class credit for their participation. Three subjects were excluded because of hardware issues rendering data unusable, and one subject was excluded because of excessive head motion. After exclusion, 36 subjects were included for the analyses.

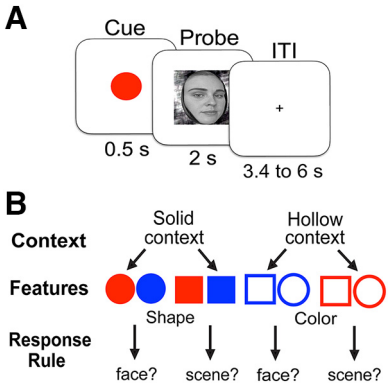

Figure 1. Behavioral paradigm and task representation. ***A***, Subjects were first presented with a cue, followed by a probe picture. Subjects were required to respond to the probe according to a response rule determined by the cue. ***B***, We hypothesized that to determine the correct response rule, subjects formed a hierarchical representation of attributes to determine the correct response rule. The highest level of contextual information informed which midlevel feature was task relevant, and this feature mapped to the correct response rule. We counterbalanced the mapping between context and shapes across subjects.

Task design and stimuli. The paradigm was designed to engage hierarchical cognitive control (Fig. 1*A*). Subjects were first presented with one of eight different colored geometric cues, immediately followed by a probe image of either a face or a scene. Subjects were required to make a button response to the probe according to one of two response rules (Face rule: Is the image a face? Scene rule: Is the image a scene?). For each trial, the cue was displayed for 0.5 s and varied across three attribute dimensions, shape, color, and texture fill. The probe images were pictures of a face or scene and remained on the screen for a maximum of 2 s or until the subject responded. The subjects can adopt a hierarchical representation of the task, in which the correct rule could be determined by evaluating different attributes of the cue, thus imposing response contingencies. The context, which is the fill texture of the cue (solid vs hollow), was first to be ascertained. This context determined which feature of the cue the subject should use to determine the response rule. A solid cue required subjects to next evaluate whether the cue was a circle or a square. Conversely, a hollow cue indicated subjects should evaluate the color of the cue, blue or red. Each feature was then mapped to either the face or scene response rule (Fig. 1*B*). After adopting the correct response rule, subjects executed a keyboard press with either the number one key or or the zero key, corresponding to yes and no, in accordance with the response rule and on viewing the probe image. The mappings between the texture of the cue and the shape/color feature were balanced and switched for half of the subjects as were the mappings of the one key or zero key for yes or no. Trials were separated by an intertrial interval ranging between 3.4 and 4.6 s long, during which a central fixation cross was displayed. To ensure subjects understood the task, we first administered a self-paced tutorial of the task and practice session. The hierarchical task rules were relayed to participants verbally (no task schematics were shown), and participants therefore knew which aspects of the stimuli to evaluate to determine their response. During the practice session, explicit feedback (correct or incorrect) was given after each trial, and total accuracy feedback was provided after the end of each practice block. During the practice session the experimenter monitored performance and answered questions from the subjects. The practice block was repeated until the subject achieved a practice accuracy of 80% or higher. In total each subject performed 415 trials over six blocks.

To infer the representational structure of the task from behavioral data, we performed a transitional RT analysis (Dykstra et al., 2022). Briefly, a transitional RT matrix was constructed for each subject by calculating the RT differences (switch cost) when switching between different cues (Fig. 2). Assuming switch cost reflects similarities in the

underlying task representations (Allport et al., 1994), the transitional RT matrix can reveal structures of task representations. If subjects formed a flat representation with no hierarchical relationship among attributes, we would expect no systematic patterns in the transitional RT matrix. Alternatively, if subjects formed hierarchical representations of tasks, we would expect the patterns of switch cost to reflect such representational structure. We therefore regressed the off-diagonal components of the observed transition RT matrices onto representational models that reflect our hypothesized task structure (Fig. 2). We evaluated a restricted and a full model. For both models, we included the representation of individual cue identities, predicting the general switch cost between different cues, and a response-rule representation, predicting switch cost of switching between face and scene response rules. For the full model, we included representations of context and feature attributes that we hypothesized to form hierarchical representations (Fig. 1*B*).

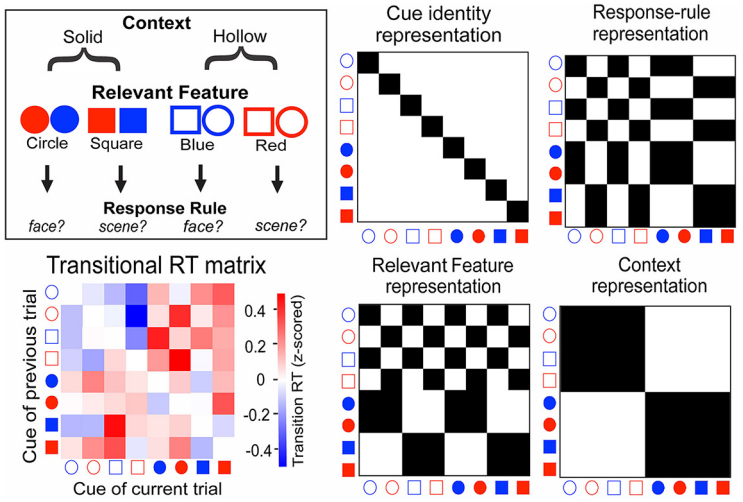

Figure 2. RT matrix and representational models. Top left, Task representation. Bottom left, The transitional RT matrix depicts the switch cost (difference in RT) for switching between trials with different cues. Trials that repeated the same cue were removed from this matrix. The color bar, on the right, indicates the magnitude of switch cost, normalized by mean and SD within each subject. Middle and right, The representation models predict patterns of switch cost based on the assumption that switching between trials with similar representations will show smaller switch costs. The observed transitional RT matrix was simultaneously regressed onto these models to evaluate their ability in explaining variances in switch cost patterns.

EEG data collection and preprocessing. EEG data were collected using a BioSemi ActiveTwo 64-channel system with a sampling rate of 512 Hz. Sixty-four Ag AgCl Pin-Type Active electrodes were applied to a fitted cap filled with electroconductive gel. To record muscle, eyeblink, and heartbeat-related activity, we also applied external electrodes on mastoids, around the eyes, and on the nose and left wrist.

Data were preprocessed using MNE-Python software (Gramfort et al., 2013). Raw data were first filtered using a finite impulse response lowpass filter of 50 Hz. Data from each channel were plotted, and noisy electrodes were rejected then interpolated (on average 1.03 channels rejected per subject, range 0–5). Data were re-referenced to the mastoid electrodes and epoched from −0.8 to 3 s time windows time locked to the probe onset. Epoched data were visually inspected, and noisy epochs were rejected (on average 3.5 of the 415 epochs were rejected per subject). Data were then subjected to an independent component analysis (ICA) using the fastICA algorithm (Bell and Sejnowski, 1995); noisy components that were characteristics of eyeblink or muscle artifacts were rejected. On average 3.3 of 64 independent components were rejected per subject.

Time-frequency and decoding analysis. Because past studies suggest that neural signals from different frequency bands carry task-relevant information for hierarchical cognitive control (Buschman et al., 2012; Voytek et al., 2015; Riddle et al., 2021, 2020), we performed time-frequency transformation on our EEG data. Single-trial EEG epochs were convolved with a family of complex wavelets to obtain the following time-frequency spectrum:

$$G(t,f) = \frac{1}{\sqrt{2\pi f}} \exp\left(\frac{-t^2}{2\sigma^2}\right) \exp(i2\pi f t),$$

where t is the time point within the trial epoch, f is the frequency that increased from 1 to 40 Hz in 30 logarithmically spaced steps, and σ is defined as follows:

$$n/2\pi f,$$

where n increased from 3 to 10 also in logarithmical steps. This spacing was used to balance the time and frequency resolutions. For each time point, power was calculated using the squared amplitude of the resulting complex spectrum.

To determine whether EEG data contained task-specific information, we performed decoding analysis using the linear discriminatory analysis (LDA) implemented via the scikit-learn package (Pedregosa et al., 2011). Specifically, we used the observed EEG data as training features and trials as observations to decode individual cue types separately for each time point. For time-resolved decoding, whole-scalp potentials from 64 channels were used as features, whereas for time-frequency decoding, power values from the wavelet analysis were used as features. To prevent overfitting, we repeated 10 runs of k-fold cross-validations by randomly partitioning trials into four independent folds. Three folds were used as the training set, and the remaining one was used as the test set. Resulting classification probabilities were averaged across the 10 cross-validated runs. For each time point, frequency, and trial, this decoding analysis returned a confusion profile of classification probabilities for each cue. We repeated the same time-frequency decoding analysis to decode context (texture fill of the cue, i.e., filled, or hollow), task-relevant feature (the relevant cue feature, i.e., color, red or blue; or shape, circle or square), and response rule (face or scene).

Representational similarity analysis. To determine the representational structure of our task, we applied RSA to the confusion profiles from LDA predicting individual cues. Specifically, for each subject we regressed the trial-by-trial confusion profiles onto the following representational models: context, task-relevant feature, individual cue identity, and response rule (Fig. 2). These representational models were orthogonal, with variance inflation factors (VIFs) all smaller than three (VIF values, context = 1.54 feature = 2.46, cue identity = 1.54, task = 2.46, mean r among models $\leq$ 0.23). Each model describes which cues share an attribute and further predict that decoding probability will increase among cues that shared the same attribute. For example, cues that share the same solid or hollow texture fill will exhibit similarly increased decoding probabilities in the confusion profile. In the regression model, decoding probabilities were log transformed and entered as the dependent variable, and representational models were simultaneously entered as regressors. Trial-by-trial RTs were entered as a covariate, and error trials and trials with overly long RT (5× the SD) were excluded from this analysis. The resulting regression coefficients were treated as representational strengths of each representation, and the time-frequency varying patterns of coefficients reflected the representational dynamics. The statistical significance of representational strength was determined using a randomized permutation test (see below, Cluster-based permutation test). Note that we opted to evaluate representational strength with this approach instead of comparing raw prediction accuracy from LDA because the baseline accuracy between predicting individual cues (chance level, 12.5%) and context or features (chance level, 50%) were not comparable. Therefore RSA is a more suitable method as it can simultaneously evaluate the relative representational strengths of multiple models.

Granger causality analysis on representational dynamics. After obtaining the representation strength of context, task-relevant feature, and

response rules, we then tested patterns of hierarchical interactions among these representations. To this end, we tested the strengths of directional interactions among representations by performing GCA (Barnett and Seth, 2014; Seth et al., 2015). For each subject, we first epoched the representational time courses that covered the cue and the probe period. We then subtracted the mean from each representational time series and divided by the SD. The normalized time series were entered into an autoregressive model, and the model order was determined using the Bayesian information criterion, explored between 10 and 100 ms. Granger causality values were estimated in a bivariate fashion, fully crossing between representations (context, feature, cue identity, response rule) and frequencies (1–40 Hz). This procedure resulted in matrices describing the strength of cross-frequency directional interactions among representations. To determine the statistical significance, we calculated a dominant influencing pattern measure between two representations (for example X and Y) by subtracting the Granger causality values of X-to-Y from Y-to-X. We then used cluster-based permutation tests to determine statistical significance. A significant result suggests that one representation was dominantly influencing another representation stronger in one direction than in the reverse direction.

Cluster-based permutation test. To test for statistical significance of representational strength and Granger causality patterns, we performed nonparametric cluster-based permutation tests to control for multiple comparisons (Maris and Oostenveld, 2007). For representational strength, we first performed one-sample t tests by testing the observed regression coefficients across subjects against a null value of zero for each time-frequency sample. For GCA, we contrasted the difference of influencing directions (X-to-Y vs Y-to-X) for every frequency pair. We then determined the uncorrected statistical threshold of $t = 2.04$, which corresponds to an uncorrected significance threshold of $p < 0.05$, and clustered continuous time-frequency samples that exceeded this uncorrected threshold. We calculated the cluster statistical mass by summing the t statistics within each cluster. We then randomly permuted the sign of the observed values 1000 times, recalculated the cluster mass for each randomized permutation, and pooled the results to derive empirical null distributions of cluster masses. Significant effects could only occur by chance in these null distributions, thus satisfying the null hypothesis of no effect. The corrected significance value was determined by the proportion of values in the null distribution that were greater than in the original observed unpermuted cluster mass. This approach controls for multiple comparisons by testing the significance of a single cluster mass rather than performing a separate significance test for each time-frequency sample.

Structural equation modeling. To determine the behavioral relevance of task representations, we tested a path model that describes a hierarchical relationship among representations and trial-by-trial RT. This analysis required us to obtain trial-by-trial estimates of representational strengths. To obtain such trial-by-trial estimates, for each trial, we pooled the confusion profiles from a broad frequency range (1–40 Hz) during the cue (the full 500 ms) and the response period (the first 500 ms for response-rule representations). This pooling procedure ensures sufficient observations to conduct the RSA regression procedure separately for each trial. We then repeated the RSA regression to obtain trial-by-trial estimates of context, feature, and cue identity. We entered these trial-by-trial representational estimates and RT data into an SEM analysis. To account for nonindependence in the data, we adopted cluster-robust SEs by including the participant as the clustering variable.

For the path model, we hypothesize that the strength of the response-rule representation influences RT, whereas the response-rule representation receives input from the feature representation. The feature representation in turn receives input from the context representation. We predict that the cue identity representation will have a weaker effect on the response-rule representation when compared with the feature representation, indicating that subjects adopted a hierarchical representation to infer the appropriate response rule rather than memorizing S–R mappings for each individual cue. The path coefficients of these relationships were evaluated with structural equational modeling using the lavaan package (version 0.6) in R software (Rosseel, 2012). Acceptable model fit was evaluated using the following established cutoffs: comparative fit index (CFI) ≥ 0.95, root mean square error of approximation (RMSEA) < 0.08, and standardized root mean square residual (SRMR) ≤ 0.08 (Schreiber et al., 2006).

Data availability. Code and data are available at https://github.com/HwangLabNeuroCogDynamics/TaskRep.

Results

Behavioral results

Subjects performed a hierarchical control task (Fig. 1*A*) in which the correct response rule can be determined by evaluating different attributes of the cue if subjects engaged hierarchical cognitive control (Badre and Nee, 2018). We hypothesized that subjects formed a hierarchical representation of attributes to determine the correct rule (Fig. 1*B*). Specifically, subjects first considered the highest level contextual attribute (whether the cue was filled with color or hollowed in texture). This higher-level contextual attribute further informs which lower-level feature is task relevant (color or shape); the feature was then mapped to the correct response rule.

Subjects performed the task accurately (mean accuracy = 0.91, SD = 0.07, mean reaction time = 825 ms, SD = 169 ms). We then preformed the transitional RT analysis (see above, Materials and Methods) and found that the regression coefficient for context representation was statistically significant ($b = -0.22$, $t = 4.048$, $p < 0.001$, 95% confidence interval, -0.112 to -0.322), suggesting that switching between cues within the same contexts showed a smaller switch cost. Regression coefficients for feature, cue identity, and response-rule representations were not statistically significant. We further evaluated two models, a full model incorporating context, feature, cue identity, and response-rule representations versus a restricted model including only cue identity and response-rule representations. Model comparison showed that the full model outperformed the restricted model [$F_{(2,2012)} = 9.4$, $p < 0.001$; full model Akaike information criterion (AIC) = 5919, restricted model AIC = 5934; full model Bayesian information criterion (BIC) = 5942, restricted model BIC = 5949]. This indicates that switching between contexts can explain significantly more variances in switch cost patterns than the restricted model assuming a flat representation. We also tested an alternative model in which all squares mapped to the scene task (except the hollow red square), and all circles mapped to the face task (except the hollow blue circle), altogether forming four groups of representations. We reasoned that if subjects adopted such a strategy, switch cost would increase in trials that switched between these four groups of representations. We entered this alternative model and the response rule model into our transitional RT regression analysis and found that this model did not explain a significant amount of variance in the observed transitional RT patterns ($b = -0.0295$, SE = 0.05, $t = -0.592$, $p = 0.55$, 95% confidence interval, -0.127 to 0.068), indicating no statistically significant switch costs when switching among these four representations. In summary, our behavioral results suggest that subjects formed task representations beyond individual cues by extracting contextual information from the presented stimuli. We performed additional analyses on the simultaneously collected EEG data to further explore the structure and dynamics of neural representations.

Decoding of EEG data

To determine whether EEG data contain task-specific information, we conducted a time-resolved decoding analysis to predict individual cues from EEG trial epochs. For each time sample,

whole-scalp potentials were used to train LDA classifiers using a four-fold cross-validation procedure (see above, Materials and Methods). Decoding results were then compared with an empirical null distribution of predictions trained with 1000 randomly shuffled cue labels. We found above-chance decoding performance; immediately after the cue onset, elevated decoding accuracy lasted throughout the cue period, showed another transient increase after the probe onset, and lasted through the first second of the probe period before falling to chance level (Fig. 3*A*). This temporal pattern is consistent with a hypothetical mental chronometry of evaluating the presented cue to determine the correct response rule. In addition to the temporal decoding analysis, we further performed a time-frequency decoding analysis using band-specific frequency power (ranging from 1 to 40 Hz) at each time point as training features. After both cue and probe onsets, we found above chance decoding in a broad frequency range (1–30 Hz; Fig. 3*B*).

We then repeated the time-frequency decoding analysis for decoding context (filled vs unfilled texture), shape (circle vs square), color (red vs blue), and response rule (face vs scene). After both cue and probe onsets, we found above chance decoding in a broad frequency range (Fig. 4). Context (texture) had the strongest decoding accuracy when compared with color, shape, and task (context vs color, $t_{(36)} = 9{,}74$, $p < 0.001$; context vs shape, $t_{(36)} = 7{,}62$, $p < 0.001$; context vs task: $t_{(36)} = 9{,}82$, $p < 0.001$). These results suggest that task information can be decoded from the multivariate time-frequency patterns of EEG activity.

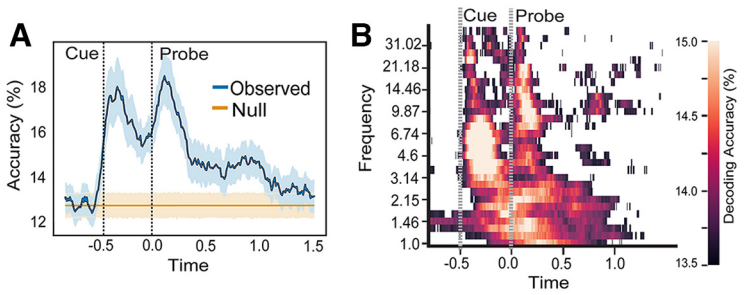

Figure 3. Decoding accuracy for individual cues. ***A***, Time-resolved decoding accuracy. Shaded errors represent 95% confidence interval across subjects. ***B***, Time-frequency decoding accuracy, thresholded by the 95th percentile from the empirical null distribution. For all graphs, *x*-axis is time in seconds.

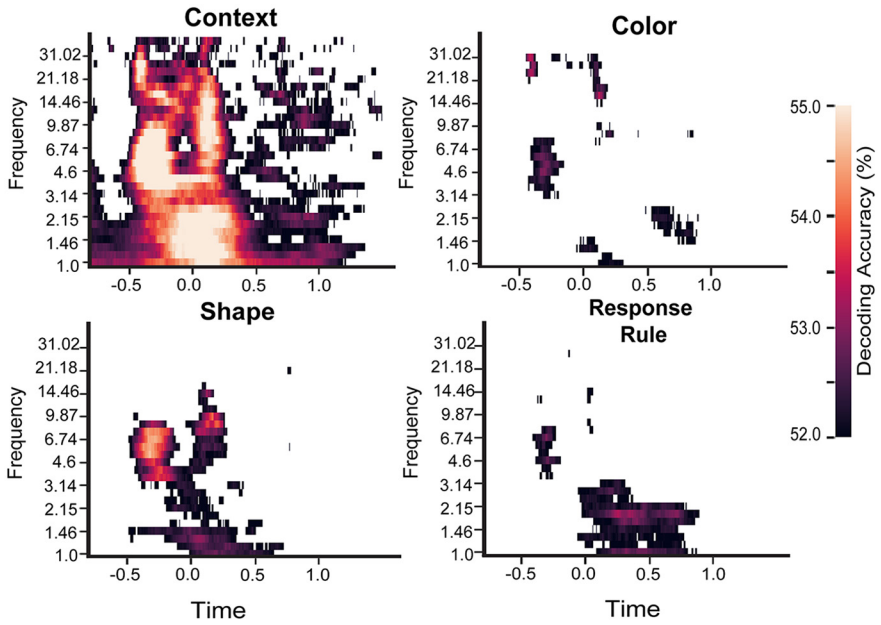

Figure 4. Decoding accuracy for context, color, shape, and response rule. Decoding accuracies thresholded by the 95th percentile from the empirical null distribution. For all graphs, *x*-axis is time in seconds.

Representation similarity analysis

We performed RSA (Kriegeskorte et al., 2008; Kikumoto and Mayr, 2020) to evaluate the representational structures that can explain the observed multivariate EEG decoding results presented in Figure 3*B*. First, for each subject, trial, and time-frequency measurement, we performed LDA and obtained classification probabilities for each of the eight possible task cues. The observed posterior probabilities were converted into confusion profiles and regressed onto representational models of context, feature, cue identity, and response rule (Fig. 5*A*). All model regressors were simultaneously evaluated via linear regression, and the resulting coefficients and their associated *t* statistics indicated the unique variances in EEG decoding patterns explained by each representational model. Henceforth, the resulting regression coefficients can be interpreted as the representational strength of different task representations. Only correct trials were included in the analysis, and reaction times were entered as a covariate. Statistical significance was determined via a cluster-based randomized permutation procedure.

We found statistically significant coding of context, feature, and cue identity representations (Fig. 5). For context, relevant feature, and cue identity representations, clusters in the observed data were extended between the delta to theta frequency range (1–8 Hz) after the cue and probe onsets (Fig. 5*B*). For context and cue identity representations, clusters were also observed in alpha and beta frequency ranges (8–30 Hz) and for context representation, the slower delta frequency range (1–4 Hz). For relevant feature and response-rule representations, clusters in the observed data overlapped in the delta frequency range and after the probe onset (Fig. 5*B*). These results suggest distinct representations of different attributes were constructed for the task.

Interaction between representations

We further explored patterns of interactions between the representations shown in Figure 5*B* to determine the hierarchical structure of these representations and identify which representations exhibited directional interaction with other representations. For each subject, we obtained the time series of context, feature, cue identity, and response-rule representations by extracting regression coefficients from the full-time-frequency RSA analysis (Fig. 5*B*). We then performed GCA on the representational time series to examine the directional interactions between representations. GCA was performed separately for each frequency pair. To determine the direction of interactions between

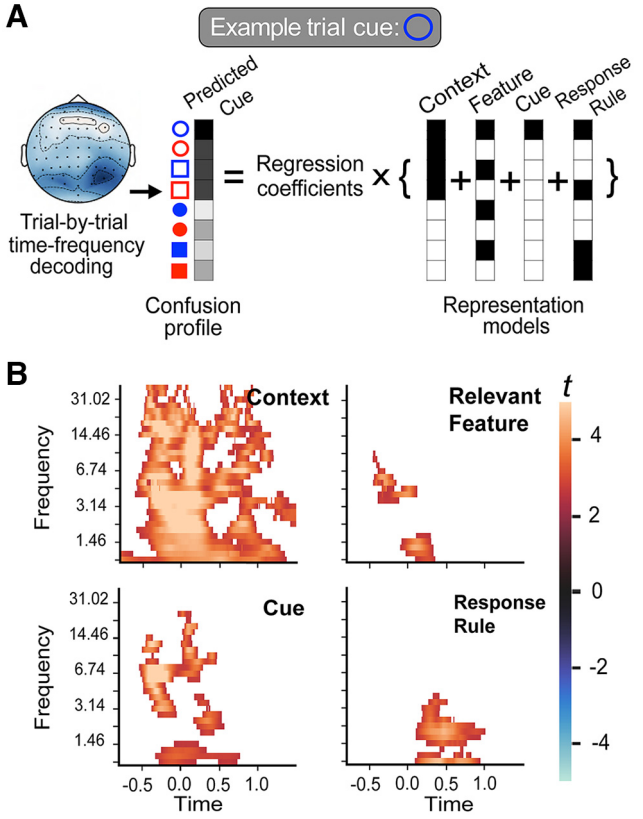

Figure 5. Time-frequency RSA. ***A***, Overview of analytical approach. Trial-by-trial decoding results for each time-frequency data point were converted into confusion matrices, where each cell represents the posterior probability from decoding. The vector of confusion profiles across trials was then regressed onto representational models of different task attributes. ***B***, Statistic (*t*) values of regression coefficients for each representation model. Significant clusters determined using cluster-based permutation test ($p < 0.05$ corrected). For all graphs, *x*-axis is time in seconds, time point −0.5 s depicts cue onset, and time point 0 depicts probe onset.

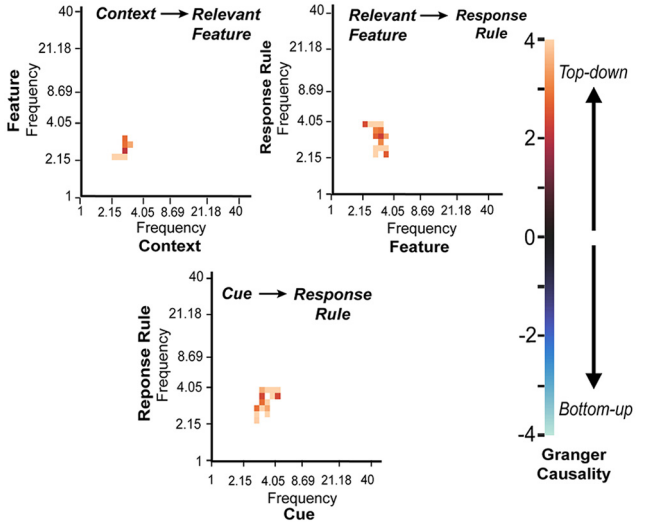

Figure 6. GCA on representational time series. Significant clusters determined by cluster-based randomized permutation test ($p < 0.05$). Far left, The color bar depicts the difference in Granger causality values; positive values represent top-down influencing direction, negative values represent bottom-up influencing direction. For all graphs, the *x*-axis and *y*-axis indicate the frequency of the representational time series. A significant cluster indicates strong influencing interaction from one representation to another in a given frequency pair.

two representations—for example, from context to feature—we performed cluster-based permutation tests to contrast the strength of Granger causality values from context to feature versus those from feature to context. A statistically significant result indicates that one representation was dominantly influencing another representation.

We found significant influencing interactions from context (texture of the cue identity) to feature (color or shape of the cue), from task-relevant feature (color or shape of the cue identity) to response rule (face or scene rule), and from cue identity to response rule (Fig. 6). Clusters in the observed data overlapped with the delta frequency band (1–4 Hz). We did not find any significant effects between other representations, nor did we find any significant influencing interactions from a lower-level representation to a supraordinate representation. This suggests a top-down hierarchical structure of task representations. This pattern of interaction is consistent with the notion of hierarchical information flow between representations, where the higher-level context influences feature and rule representations to facilitate cognitive control.

Structure of task representations

To test the role of task representations facilitating behavioral performance, we tested a path model (see above Structural equation modeling) specifying the structure of representations and examined its relationship with RT using SEM (Fig. 7). Fit indices suggest the hypothesized hierarchical relationship among representations explained the observed patterns in the data (CFI = 0.999, RMSEA < 0.001, SRMR = 0.014; Model Test User Model, $\chi^2 = 5.72$, $p = 0.24$; Model Test Baseline Model, $\chi^2 = 3964.41$, $p < 0.001$). We found that context representation was significantly associated with the feature representation ($\beta = 0.149$, SE = 0.009, $z = 16.89$, $p < 0.001$), and further found significant associations between feature and rule representations ($\beta = 0.121$, SE = 0.008, $z = 15.41$, $p < 0.001$), and between rule representation and RT ($\beta = -0.02$, SE = 0.004, $z = -4.59$, $p < 0.001$). Note that the path coefficient was negative between rule representation and RT, indicating that robust rule representation correlated with faster RT. We further found significant paths between feature and cue identity representations ($\beta = 0.738$, SE = 0.01, $z = 75.23$, $p < 0.001$) and between cue identity and context representations ($\beta = 0.242$, SE = 0.01, $z = 23.47$, $p < 0.001$). Critically, we found that the path coefficient between response rule and cue identity representations was significantly weaker than the path coefficient between feature and rule (difference in $\beta = 0.089$, SE = 0.012, $z = 7.49$, $p < 0.001$), suggesting information from feature representation was used to determine the response rule. We further compared this hierarchical model with a flat model, which hypothesizes that subjects relied on individual cue identity representations to determine the appropriate response rule and thus did not construct a hierarchical relationship among cue identity, feature, and context representations (Fig. 6*B*). Comparing these two nested models indicate that the hierarchical model fit significantly better than the flat representation model (hierarchical model, AIC = 129451, BIC = 131221, $\chi^2 = 5.72$; flat model, AIC = 135534, BIC = 137629, $\chi^2 = 6096.28$; χ^2 difference = 6090.6; degrees of freedom difference = 4, $p < 0.001$).

To summarize, results from the SEM model provide further support of a hierarchical structure of task representations. Specifically, we observe a cascade of information flow from context to feature. Feature-to-response-rule representations influence trial-by-trial

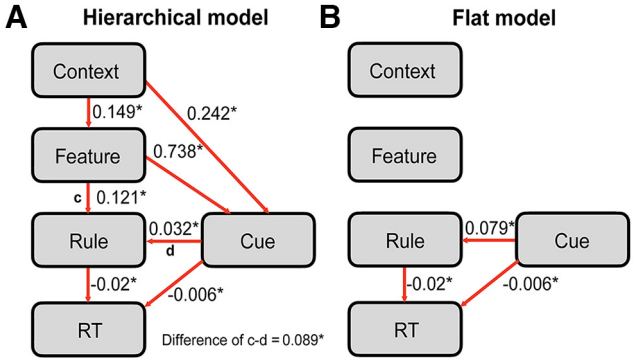

Figure 7. Path model testing the hierarchical structure of task representations. ***A***, Hierarchical model hypothesizing that higher-order contextual and feature representations influence response-rule representation for optimal task performance. ***B***, Flat representation model hypothesizing that the subject relies on individual cue identities to determine the correct response rule. Model comparison indicates that the hierarchical model outperforms the flat model. The asterisk (*) denotes statistically significant path coefficients ($p < 0.05$). All path coefficients were standardized coefficients.

RT, supporting the notion that hierarchical task representations facilitate goal-directed behaviors. Critically, this was not observed for cue identity.

Discussion

Task representations are an integral component of cognitive control (Cohen et al., 1990; Rougier et al., 2005; Braver, 2012; Badre et al., 2021). For task representations to be useful, relevant information from disparate contextual sources should be maintained and allowed to interact with one another, generating actions that satisfy contextual contingencies. However, the content and dynamics of task representations have been difficult to study. Behavioral measures, for example RT switch cost and dual-task cost, can infer boundaries and similarities between representations (Arrington et al., 2003; Hazeltine and Schumacher, 2016) but have no direct access to neural representations. Univariate response amplitudes measured with neuroimaging techniques, such as blood-oxygenated-level-dependent signal, can functionally localize changes in activity corresponding to experimental manipulations (Badre and D'Esposito, 2007; Koechlin et al., 2003) but cannot be used to identify multivariate representational content (Freund et al., 2021). Furthermore, the relatively slow temporal resolution of fMRI makes it difficult to detect interactions between representations, which likely occur at the time scale of milliseconds. Recent methodological advances in combining EEG with RSA directly address these limitations. Specifically, the structure of task representations can be revealed by testing models of information representation with multivariate brain activity patterns (Kriegeskorte et al., 2008; Nili et al., 2014). Applying RSA to time-resolved EEG data can track the representational dynamics of cognitive chronometry (Hubbard et al., 2019; Kikumoto and Mayr, 2020). The present study adopted these methodological advances to study how task representations facilitate hierarchical cognitive control.

Consistent with the notion that task representations encompass information beyond simple S–R mappings (Hazeltine and Schumacher, 2016; Schumacher and Hazeltine, 2016), we identified representations of contingent contextual information from multiple relevant sources. Neural evidence for representations of context, cue features, cue identity, and response-rule representations were observed. To further characterize how these representations interact to facilitate hierarchical control, we subsequently performed GCA and SEM analyses. We found that higher-level contextual representation influenced subordinate representations of task-relevant features and response rules, bridging the cascade of information transfer between stimuli input to response selection in a context-dependent manner. Below, we discuss how our findings reveal properties of hierarchical task representations.

Useful task representations should be behaviorally relevant. Results from our SEM analysis showed that response-rule representation exhibited the strongest correlation with RT. Moreover, our GCA results showed that response-rule representation was in turn under the influence of higher-order representations including context and cue features. We found that the interactions between representations were asymmetric; context representation dominantly influenced feature representation, and feature representation dominantly influenced response-rule representation. We did not find significant influencing interactions from response-rule representation to higher-order representations. This asymmetric interaction suggests a top-down, hierarchical structure of constituent representations. This observation is further consistent with theories of cognitive control where response selection is constrained by hierarchically organized contextual information (Botvinick, 2008; Badre and Nee, 2018). The identified hierarchical cascade of representational interaction directly supports the top-down control processes in previously proposed models of cognitive control (Norman and Shallice, 1986; Cohen et al., 1990; Posner et al., 2004).

An alternative explanation is that subjects form a flat representation associating each individual cue identity with a response rule. Behaviorally, if subjects had adopted such a flat representation with no hierarchical relationship, then we would not have observed structural switch costs when subjects switched between trials presented with different cues. Our behavioral results did not support this prediction, and instead a regression analysis revealed that a model with hierarchical context representation can explain significantly more variances in switch cost patterns beyond a flat representation. Results from our SEM analysis on representational dynamics also favored the hierarchical interpretation. Specifically, the path coefficient between the cue identity representation and the response-rule representation was significantly weaker than that between feature and context representations. This result suggests that subjects likely did not rely on memorizing eight individual S–R mappings to determine the correct response rule but instead used a hierarchical context-to-rule mapping. More broadly, our results are consistent with previous observations that human subjects can learn and form hierarchical representations of tasks (Collins and Frank, 2013; Eichenbaum et al., 2020; Vaidya et al., 2021), which has been argued to have the advantage of reducing memory load and increasing efficiency for satisfying complex response contingencies (Badre et al., 2010; Frank and Badre, 2012).

Although our results suggest a hierarchical structure of task representations, they do not suggest a strict step-like, sequential order of transition from context to rule representations. Instead, given that multiple representations were simultaneously activated, interactions among constituent representations likely occurred simultaneously. These parallel neural dynamics of hierarchical interaction are consistent with behavioral results reported from a previous study using a response-deadline manipulation (Ranti et al., 2015). Ranti et al. (2015) observed similar error rates across

different contingency levels, regardless of the amount of time given to subjects to make a response. This suggests that subjects do not traverse multiple hierarchical levels in a serial fashion. Instead, after extracting information from relevant sources, information across hierarchical levels is processed in parallel to converge on the most appropriate response rule. Such a parallel processing scheme may be more efficient for hierarchical cognitive control when compared with a strictly serial step-like decision process.

Our EEG data cannot pinpoint the precise anatomic localization or circuit mechanisms that instantiate task representations across hierarchy levels. Past studies suggest that the frontal and parietal cortices are likely important for task representations. For example, prior fMRI studies that used multivoxel pattern analysis were able to successfully decode rule representations from frontal and parietal cortices (Woolgar et al., 2011; Waskom et al., 2014; Cole et al., 2016; Etzel et al., 2016; Pischedda et al., 2017; Qiao et al., 2017), and tasks that required contextual control over actions more selectively involved the middle lateral prefrontal cortex (Badre and D'Esposito, 2007; Nee and D'Esposito, 2016; Vaidya et al., 2021). It is conceivable that task representations decoded from our EEG data were primarily driven by these frontal (and potentially parietal) regions. Whether the representations we observed across different contingency levels were encoded by common, overlapping, or segregated regions will need to be addressed by applying RSA to data with greater spatial specificity.

We performed LDA and RSA on both time- and frequency-resolved EEG data because past research on nonhuman primate models suggests that electrophysiology signals from different frequency bands may encode rule representations (Buschman et al., 2012; Siegel et al., 2015). Henceforth, including frequency information may increase our sensitivity in identifying task representations. We observed that clusters from the randomized permutations test for context and feature representations overlapped with the delta to beta frequency range and for response-rule representation in the delta frequency range. These observations are consistent with prior studies showing that delta and beta neural oscillations may be associated with different aspects of hierarchical cognitive control; specifically, delta signal may be involved in hierarchical abstraction of tasks for resolving multiple contingencies, whereas beta band oscillations may inhibit competing task sets (Riddle et al., 2021, 2020). Other studies suggests that alpha band oscillatory power decrease during tasks that switch among contextual contingencies (Cooper et al., 2016), which may reflecting task updating processes.

It is important to point out that prior studies focused on contrasting the oscillatory magnitudes between task conditions and did not focus on determining the representational content nor the interactive dynamics among different levels of representations. Our results suggest that different levels of hierarchical representations are encoded by our subjects. We found statistically significant top-down interactions between context, feature, and response-rule representations. Clusters in observed interaction data overlapped with the delta frequency range, and we hypothesized that delta signal may be a carrier frequency integrating distinct representations across different levels of hierarchy. We cannot determine whether this effect was driven by spatially overlapping or segregated anatomic regions encoding representations with similar or distinct oscillatory signals. Another important limitation of our study is that although we counterbalanced which stimuli dimension (texture, shape, color) was assigned as the task-relevant feature, we did not counterbalance which was assigned as the context in the task hierarchy. This limits the generalizability of our findings on whether the context of our task can be generalized to other dimensions of the cue. This critical issue will need to be addressed by future studies.

In summary, task representations have long been thought to be critical for context-dependent, adaptive cognitive control. However, the way in which task representations can integrate information from multiple task-relevant features to satisfy contextual contingency has been difficult to study. The principal contribution of our study is the characterization of the hierarchical structure and interactive dynamics of task representations. Our results relate top-down cognitive chronometry with multivariate EEG activity and uncover latent dynamics of human information processing for hierarchical cognitive control.

References

Allport DA, Styles EA, Hsieh S (1994) Shifting intentional set: exploring the dynamic control of tasks. In: Attention and performance 15: Conscious and nonconscious information processing (Umiltà C, Moscovitch M, eds), pp 421–452). Cambridge, MA: MIT.

Arrington CM, Altmann EM, Carr TH (2003) Tasks of a feather flock together: similarity effects in task switching. Mem Cognit 31:781–789.

Badre D, D'Esposito M (2007) Functional magnetic resonance imaging evidence for a hierarchical organization of the prefrontal cortex. J Cogn Neurosci 19:2082–2099.

Badre D, Nee DE (2018) Frontal cortex and the hierarchical control of behavior. Trends Cogn Sci 22:170–188.

Badre D, Kayser AS, D'Esposito M (2010) Frontal cortex and the discovery of abstract action rules. Neuron 66:315–326.

Badre D, Bhandari A, Keglovits H, Kikumoto A (2021) The dimensionality of neural representations for control. Curr Opin Behav Sci 38:20–28.

Barnett L, Seth AK (2014) The MVGC multivariate Granger causality toolbox: a new approach to Granger-causal inference. J Neurosci Methods 223:50–68.

Bell AJ, Sejnowski TJ (1995) An information-maximization approach to blind separation and blind deconvolution. Neural Comput 7:1129–1159.

Botvinick MM (2008) Hierarchical models of behavior and prefrontal function. Trends Cogn Sci 12:201–208.

Braver TS (2012) The variable nature of cognitive control: a dual mechanisms framework. Trends Cogn Sci 16:106–113.

Buschman TJ, Denovellis EL, Diogo C, Bullock D, Miller EK (2012) Synchronous oscillatory neural ensembles for rules in the prefrontal cortex. Neuron 76:838–846.

Cohen JD (2017) Cognitive control: core constructs and current considerations In: The Wiley handbook of cognitive control (Egner T, ed), pp 1–28. Chichester, UK: Wiley.

Cohen JD, Dunbar K, McClelland JL (1990) On the control of automatic processes: a parallel distributed processing account of the Stroop effect. Psychol Rev 97:332–361.

Cole MW, Ito T, Braver TS (2016) The behavioral relevance of task information in human prefrontal cortex. Cereb Cortex 26:2497–2505.

Collins AGE, Frank MJ (2013) Cognitive control over learning: creating, clustering, and generalizing task-set structure. Psychol Rev 120:190–229.

Cooper PS, Darriba Á, Karayanidis F, Barceló F (2016) Contextually sensitive power changes across multiple frequency bands underpin cognitive control. Neuroimage 132:499–511.

D'Esposito M (2007) From cognitive to neural models of working memory. Philos Trans R Soc Lond B Biol Sci 362:761–772.

Diedrichsen J, Kriegeskorte N (2017) Representational models: a common framework for understanding encoding, pattern-component, and representational-similarity analysis. PLoS Comput Biol 13:e1005508.

Dykstra T, Smith DM, Schumacher EH, Hazeltine E (2022) Measuring task structure with transitional response times: task representations are more than task sets. Psychon Bull Rev. Advance online publication. Retrieved April 8, 2022. doi:10.3758/s13423-021-02035-3.

Eichenbaum A, Scimeca JM, D'Esposito M (2020) Dissociable neural systems support the learning and transfer of hierarchical control structure. J Neurosci 40:6624–6637.

Etzel JA, Cole MW, Zacks JM, Kay KN, Braver TS (2016) Reward motivation enhances task coding in frontoparietal cortex. Cereb Cortex 26:1647–1659.

Frank MJ, Badre D (2012) Mechanisms of hierarchical reinforcement learning in corticostriatal circuits 1: computational analysis. Cereb Cortex 22:509–526.

Freund MC, Etzel JA, Braver TS (2021) Neural coding of cognitive control: the representational similarity analysis approach. Trends Cogn Sci 25:622–638.

Gramfort A, Luessi M, Larson E, Engemann DA, Strohmeier D, Brodbeck C, Goj R, Jas M, Brooks T, Parkkonen L, Hämäläinen M (2013) MEG and EEG data analysis with MNE-Python. Front Neurosci 7:267.

Hazeltine E, Schumacher EH (2016). Understanding central processes: the case against simple stimulus-response associations and for complex task representation In: Psychology of learning and motivation (Ross BH, ed), pp 195–245. Amsterdam: Academic Press.

Hubbard J, Kikumoto A, Mayr U (2019) EEG decoding reveals the strength and temporal dynamics of goal-relevant representations. Sci Rep 9:9051.

Kikumoto A, Mayr U (2020) Conjunctive representations that integrate stimuli, responses, and rules are critical for action selection. Proc Natl Acad Sci U S A 117:10603–10608.

Koechlin E, Ody C, Kouneiher F (2003) The architecture of cognitive control in the human prefrontal cortex. Science 302:1181–1185.

Kriegeskorte N, Mur M, Bandettini P (2008) Representational similarity analysis—connecting the branches of systems neuroscience. Front Syst Neurosci 2:4.

Maris E, Oostenveld R (2007) Nonparametric statistical testing of EEG- and MEG-data. J Neurosci Methods 164:177–190.

Nee DE, D'Esposito M (2016) The hierarchical organization of the lateral prefrontal cortex. Elife 5:e12112.

Nili H, Wingfield C, Walther A, Su L, Marslen-Wilson W, Kriegeskorte N (2014) A toolbox for representational similarity analysis. PLoS Comput Biol 10:e1003553.

Norman DA, Shallice T (1986) Attention to action. In: Consciousness and self-regulation: advances in research and theory (Davidson RJ, Schwartz GE, Shapiro D, eds), pp 41–18. Boston: Springer.

Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V (2011) Scikit-learn: machine learning in Python. J Mach Learn Res 12:2825–2830.

Pischedda D, Görgen K, Haynes J-D, Reverberi C (2017) Neural representations of hierarchical rule sets: the human control system represents rules irrespective of the hierarchical level to which they belong. J Neurosci 37:12281–12296.

Posner MI, Snyder CR, Solso R (2004) Attention and cognitive control. Cognitive psychology: Key readings 205:55–85.

Qiao L, Zhang L, Chen A, Egner T (2017) Dynamic trial-by-trial recoding of task-set representations in the frontoparietal cortex mediates behavioral flexibility. J Neurosci 37:11037–11050.

Ranti C, Chatham CH, Badre D (2015) Parallel temporal dynamics in hierarchical cognitive control. Cognition 142:205–229.

Riddle J, Vogelsang DA, Hwang K, Cellier D, D'Esposito M (2020) Distinct oscillatory dynamics underlie different components of hierarchical cognitive control. J Neurosci 40:4945–4953.

Riddle J, McFerren A, Frohlich F (2021) Causal role of cross-frequency coupling in distinct components of cognitive control. Prog Neurobiol 202:102033.

Rosseel Y (2012) Lavaan: an R package for structural equation modeling. J Stat Softw 48:1–36.

Rougier NP, Noelle DC, Braver TS, Cohen JD, O'Reilly RC (2005) Prefrontal cortex and flexible cognitive control: rules without symbols. Proc Natl Acad Sci U S A 102:7338–7343.

Sakai K (2008) Task set and prefrontal cortex. Annu Rev Neurosci 31:219–245.

Schreiber JB, Nora A, Stage FK, Barlow EA, King J (2006) Reporting structural equation modeling and confirmatory factor analysis results: a review. J Educ Res 99:323–338.

Schumacher EH, Hazeltine E (2016) Hierarchical task representation: task files and response selection. Curr Dir Psychol Sci 25:449–454.

Seth AK, Barrett AB, Barnett L (2015) Granger causality analysis in neuroscience and neuroimaging. J Neurosci 35:3293–3297.

Siegel M, Buschman TJ, Miller EK (2015) Cortical information flow during flexible sensorimotor decisions. Science 348:1352–1355.

Vaidya AR, Jones HM, Castillo J, Badre D (2021) Neural representation of abstract task structure during generalization. Elife 10:e63226.

Voytek B, Kayser AS, Badre D, Fegen D, Chang EF, Crone NE, Parvizi J, Knight RT, D'Esposito M (2015) Oscillatory dynamics coordinating human frontal networks in support of goal maintenance. Nat Neurosci 18:1318–1324.

Waskom ML, Wagner AD (2017) Distributed representation of context by intrinsic subnetworks in prefrontal cortex. Proc Natl Acad Sci U S A 114:2030–2035.

Waskom ML, Kumaran D, Gordon AM, Rissman J, Wagner AD (2014) Frontoparietal representations of task context support the flexible control of goal-directed cognition. J Neurosci 34:10743–10755.

Woolgar A, Hampshire A, Thompson R, Duncan J (2011) Adaptive coding of task-relevant information in human frontoparietal cortex. J Neurosci 31:14592–14599.