Anticipatory influences on simple sensory-motor behaviors are encoded by rapidly fluctuating neural dynamics across the human brain

Ashwin G. Ramayya[1]*, Vivek Buch[2], Andrew Richardson[1], Timothy Lucas[1], and Joshua I. Gold[3]

Affiliations:

[1]Department of Neurosurgery, Hospital of University of Pennsylvania, Philadelphia, PA 19104, USA

[2]Department of Neurosurgery, Stanford University, Stanford, CA 94305, USA

[3]Department of Neuroscience, University of Pennsylvania, Philadelphia, PA 19104, USA

*Corresponding author. Email: ashwinramayya@gmail.com

Abstract

A major challenge in understanding how expectations guide sensory-motor behavior has been to relate relatively simple algorithmic explanations to their potentially complex neural substrates. We studied this relationship using intracranial recordings from 23 patients with medically refractory epilepsy performing a simple visual-detection task with variable foreperiod delay. We identified two distinct anticipatory influences on behavior by fitting an abstract motor-preparatory model ("rise-to-bound") to response times (RTs). The first "predictive" process increased reaction speed and premature responses prior to stimulus arrival ("false alarms"), whereas the second "reactive" process increased reaction speed without affecting false alarms. We linked these anticipatory processes to rapidly fluctuating neural-activity patterns across the brain by leveraging the high spatio-temporal resolution and broad coverage of intracranial recordings. Predictive and reactive signals were highly distributed and interspersed throughout the brain but distinguished by their timing (e.g., occurring around response and stimulus onset, respectively), anatomical distribution (e.g., greater representations in MTL and perirolandic areas, respectively), and modulation patterns (e.g., decreased and increased activity during periods of higher expectation, respectively). These results provide new insights into the complex, but identifiable, mappings between specific, behaviorally derived forms of anticipatory information processing and the widespread neural dynamics that implement them.

Introduction

The ability to use expectations of future events to guide sensory-motor behavior is a fundamental building block of cognition (Friston, 2010; Nobre & Van Ede, 2018). For example, even simple stimulus detection is not a passive, purely stimulus-driven process but involves predictive inference that combines incoming sensory information with expectations learned from prior experience (Helmholtz, 1866; Rao & Ballard, 1999). However, the mechanisms that allow expectations to affect sensory-motor processing in the human brain (a set of processes that we refer to as "anticipation") remains unclear, reflecting at least three major challenges.

The first major challenge is conceptual: there is a disconnect between two broad views on anticipatory processing of sensory-driven behaviors that have dominated the field, particularly when expectations are related to the timing and probability of occurrence of an impending event. A "predictive" view emphasizes anticipatory mechanisms that prepare behavioral responses that are associated with expected sensory inputs (Friston, 2010; Los et al., 2014; Montague et al., 1996; Rescorla & Wagner, 1972; Salet et al., 2022). In contrast, a "reactive" view emphasizes anticipatory mechanisms that prioritize (e.g., via selective attention) how expected sensory inputs are processed once they arrive (Summerfield & Egner, 2009). The relative contributions of predictive and reactive processes to anticipatory sensory-driven behaviors, and how and where they are implemented in the brain, are not fully understood (Nobre & Van Ede, 2018).

A second major challenge is computational: we have a limited understanding of how the relatively simple algorithmic models often used to quantify anticipatory effects on behavior relate to the possibly complex neural dynamics that implement those algorithms (O'Connell et al., 2018). For example, anticipatory influences on response times (RTs) have been studied

extensively (Cattell, 1886; Luce, 1986; Niemi & Näätänen, 1981) and can be modeled as modulations of a “rise-to-bound” process that triggers a motor response once a latent signal rises to a certain value (the “bound”; Carpenter and Williams, 1995; Noorani and Carpenter, 2016). A compelling feature of these models is that, in addition to providing parsimonious accounts of behavior, they have algorithmic components that are thought to map directly onto the activity patterns of localized sets of neurons that contribute to motor preparation and execution (Hanes & Schall, 1996). However, this mapping remains contentious and not well understood (Latimer et al., 2015; Shadlen et al., 2016), reflecting in part a mismatch in complexity. Whereas rise-to-bound and other algorithmic models provide low-dimensional descriptions of cognitive processes that underlie behavior, the human brain is a high-dimensional system that generates emergent behavior from functionally heterogeneous, widely distributed and highly interconnected and interacting neural elements (Mesulam, 1998) whose relationships to behaviorally derived models are not well understood.

The third major challenge is practical: few studies have measured neural activity with appropriate resolution and scale to identify neural representations of anticipatory processing throughout the brain. Rise-to-bound models suggest that anticipatory modulation of RTs likely involves rapidly fluctuating neural dynamics between stimulus and response, a variable time interval that often ranges between ~200–700 ms (Luce, 1986). Neural correlates of RT variability have been identified in several motor-preparatory brain regions, including activity patterns that map directly (Hanes and Schall, 1996; O’Connell et al., 2012) or indirectly (Heitz and Schall, 2012; Hauser et al., 2018; Paraskevopoulou et al., 2021) onto model dynamics. However, these findings have provided only a limited understanding of the underlying brain

signals. Previous studies that identified RT-related representations have either had: 1) high resolution but local scope (e.g., invasive recordings in animals that typically target individual neurons in specific motor or sensory-motor brain regions thought to be relevant to a particular task but neglect potential contributions by the rest of the brain; Hanes & Schall, 1996; Hauser et al., 2018; Purcell et al., 2010), or 2) more global scope but much lower resolution (e.g., brain-wide human imaging studies that lack the temporal resolution to identify all but the coarsest RT modulations, or EEG that lacks the spatial resolution to distinguish local versus global representations; Callaway & Yeager, 1960; Dustman & Beck, 1965; Fox et al., 2007; O'Connell et al., 2012; Ratcliff et al., 2009).

To overcome these challenges, we obtained intracranial electroencephalography (iEEG recordings) from 23 patients with medically refractory epilepsy with indwelling intraparenchymal electrodes in widespread brain regions as they performed a stimulus-detection task with a variable foreperiod delay (Fig. 1A, Table S1). This design is known to modulate behavioral RTs based on differences in temporal expectations that depend on the foreperiod delay (Los et al., 2014; Los & Van Den Heuvel, 2001; Luce, 1986). We related these behavioral effects, which we modeled via commonly used rise-to-bound models (Stone, 1960; Ratcliff, 1978; Carpenter and Williams, 1995; Usher and McClelland, 2001; Brown and Heathcote, 2005; Noorani and Carpenter, 2016; Ratcliff et al., 2016), to high-frequency iEEG activity which reflects local spiking activity (70–200 Hz power; Manning et al., 2009; Dubey and Ray, 2019) sampled broadly across many parts of cortex and certain subcortical structures (Parvizi & Kastner, 2018). We show that relatively simple computations that account for predictive and reactive anticipatory influences on sensory-motor behavior have neural representations that are

distinct but distributed throughout the brain, in stark contrast to the more localized representations of task-relevant sensory and motor events alone. The results provide new views of the complex spatial and temporal patterns of activity in the brain that support predictive and reactive anticipatory influences on behavior.

Results

The stimulus-detection task allowed us to quantify anticipatory influences on sensory-motor behavior via modulation of RTs. Briefly, each trial began with the presentation of a visual target ("warning signal") on a computer screen that changed color after a randomly selected foreperiod delay of 500ms ("short") or 1500ms ("long"). Subjects were instructed to respond via button press as soon as they noticed the color change ("stimulus"). RT was measured as the elapsed time between stimulus and response. The different foreperiod delays provided categorically different levels of temporal expectation of stimulus arrival at the time of stimulus presentation (Luce, 1986; Nickerson, 1965; Niemi & Näätänen, 1981; Nobre & Van Ede, 2018; Ollman & Billington, 1972). On short-delay trials, the stimulus was presented when there was uncertainty about whether the trial was a short- or long-delay trial, resulting in relatively low temporal expectation of stimulus arrival. On long-delay trials, the stimulus was presented when the trial could be identified unequivocally as a long-delay trial, resulting in relatively high temporal expectation of stimulus arrival (Fig. 1A). On correct trials (those with responses that occurred within 1000 ms of stimulus onset), subjects were provided with visual feedback immediately after the response. On premature false-alarm trials (those with responses that occurred after target onset but before stimulus onset) and trials with no response within 1000 ms), subjects were shown a blank screen for ~2.5 s before the next trial began.

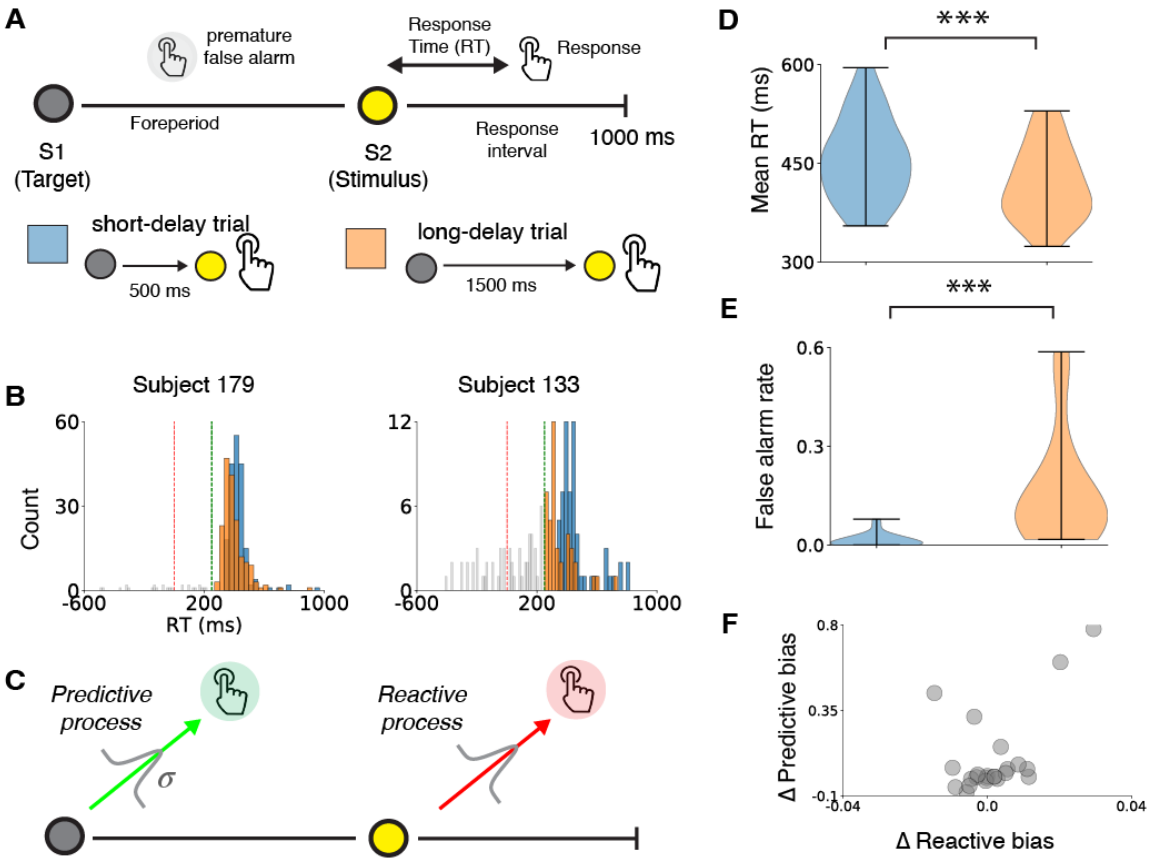

Figure 1. Behavior on a simple detection task reflected predictive and reactive forms of anticipatory processing. (**A**) Stimulus-detection task with variable foreperiod delay. (**B**) RT distributions from two example subjects. Red line indicates stimulus onset, green vertical line indicates 250 ms after stimulus onset (the fast-response threshold). Grey histogram indicates timing of premature responses; blue and orange histograms indicate timing of responses on short- and long-delay trials, respectively. (**C**) Model schematic illustrating abstracted preparatory motor processes contributing to sensory-motor behavior. Dashed boxes illustrate two distinct algorithmic components accounting for anticipatory influences on behavior: 1) a predictive process that tends to produce false alarms, and 2) a reactive process that included delay-related changes in the variance of the rate-of-rise of the latent decision variable. (**D,E**) Violin plots showing distributions of mean RTs (**D**) and anticipatory false-alarm rates (**E**) on short- (blue) and long- (orange) delay trials for all 23 subjects. (**F**) Best-fitting values of the two delay-sensitive model parameters: Δ predictive bias (ordinate, positive values imply more false alarms on long-versus short-delay trials) and Δ reactive bias (abscissa, positive values imply an increase in across-trial variance of the stimulus-driven rising process on long- versus short-delay trials). Each circle corresponds to data from a single subject.

Anticipatory influences on behavior

The subjects' RTs included stochastic variability for both delay conditions (median per-subject RT inter-quartile range=66.67 ms for short-delay and 66.75 for long-delay trials; examples are shown in Fig. 1B), with two primary effects of anticipation that were consistent with previous findings. First, subjects showed delay-related changes in RTs, with, on average, a tendency to respond with faster (paired *t*-test, $t\,(22)=5.57$, $p<0.001$) and less variable ($t(22)=2.4$, $p=0.03$) RTs on long- versus short-delay trials (Niemi & Näätänen, 1981), albeit with substantial individual variability (mean RT range across subjects=354–595 ms and 323–529 ms on short- and long-delay trials, respectively, Fig. 1D). Second, subjects had a wide range of false-alarm rates (range 1.7–58%) that were correlated with faster RTs ($r=0.47$, $p=0.02$) on long-delay trials (Luce, 1986; Nickerson, 1965; Ollman & Billington, 1972). Subjects rarely had false alarms on short-delay trials (range 0–7.8%, Fig. 1E).

Model-fitting results and alternatives

We modeled these behavioral effects using a pair of "rise-to-bound" processes that assumed that motor responses occur when a latent variable that increases in value linearly and deterministically over time reaches a threshold level or "bound" (Noorani & Carpenter, 2016; Fig. 1C and 1.S1A). The first process was triggered by the warning signal and accounted for premature false alarms. We interpret this process as a form of predictive motor plan that builds up in anticipation of the yet-to-occur onset of the sensory cue. The second process was triggered by the sensory cue and accounted for correct responses. We interpret this process as a form of reactive motor plan in response to the sensory cue. Stochastic, trial-by-trial RT variability was captured by stochastic variability in these two parallel rise-to-bound processes (R^2 mean=0.83,

range=0.55–0.99, Fig. 1.S1A,B). This model had two free parameters that could account for anticipation-dependent behaviors (i.e., a dependence of false-alarm rate and/or RT on delay condition) along separate dimensions of individual differences (Fig. 1), as follows.

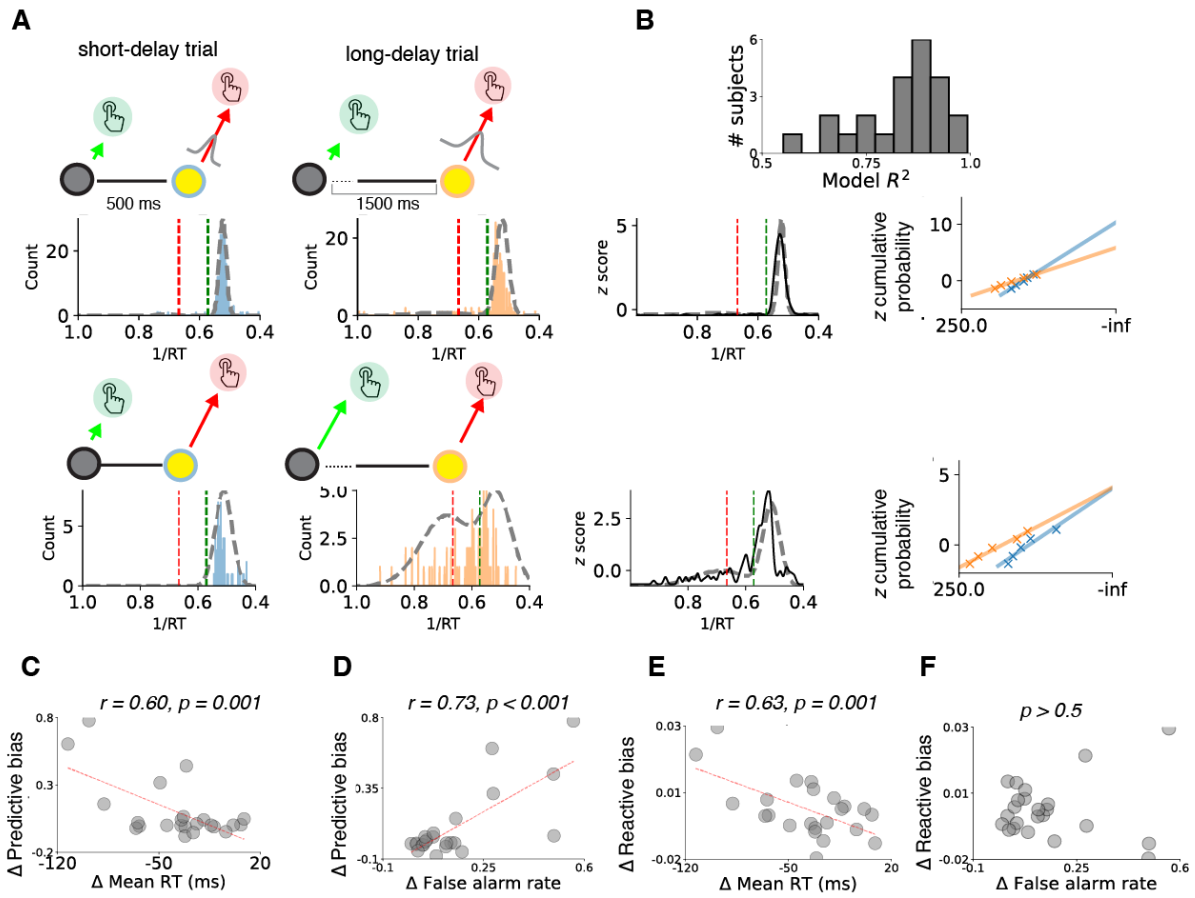

Figure 1 - Figure Supplement 1: Rise-to-bound model of RT variability. (**A**) Model schematic, RT distributions, and model fits (dashed lines) for two example subjects. Panels from left-to-right: 1) reciprocal RTs (blue) and model-fits on short-delay trials; 2) reciprocal RTs (orange) and model-fits on long-delay trials; 3) full reciprocal RT distribution and model fits (*z*-scored probability distribution); and 4) RTs from each delay condition (points are binned data, lines are model fits) shown on a reciprobit axis (where RT distributions predicted by the model are expected to fall along a straight line whose slope and intercept are governed by the model parameters; Noorani & Carpenter, 2016). (**B**) Distribution of R^2 values indicating goodness-of-fit across subjects. (**C–F**) Scatterplots showing relationships between model parameters (ordinates) and delay-related changes in model-free measures of RT (abscissas). Lines show linear fits when there was a significant correlation, as indicated.

The first parameter (Δ predictive bias) governed the relative contributions of the target- versus stimulus-driven rise-to-bound processes for long- versus short-delay trials (positive values imply a stronger target-driven contribution and more false alarms on long-delay trials). This parameter was intended to account for not just the increased frequency of false alarms on long- versus short-delay trials but also the associated increase in reaction speed with more false-alarms on long-delay trials (r=0.47, p=0.023). Accordingly, variability in best-fitting values of this model parameter across subjects captured delay-dependent decreases in RT (r = 0.63, p < 0.001, Fig. 1.S1C) that were associated with increases in false-alarm rate (r = 0.76, p < 0.001, Fig. 1.S1D).

The second parameter (Δ reactive bias) captured delay-dependent changes in RT in terms of changes in the variance of the stimulus-triggered rise-to-bound process on long- versus short-delay trials. Variability in best-fitting values of this model parameter captured decreases in RTs that, unlike those captured by Δ predictive bias, were not associated with increased false-alarms (r=0.60, p=0.002, Fig. 1.S1E,F; also see partial correlation analysis below). This parameter also captured individual differences in delay-related changes in RT variability (increased rise-to-bound variance correlated with faster RTs on long delay trials; r = 0.71, p < 0.001). Thus, delay-related changes in behavior could be explained by two anticipatory influences on RTs: 1) a predictive bias, which captured correlated changes in faster RTs and more false alarms on long delay trials (Fig. 1.S1C,D); and 2) a reactive bias, which captured changes in RTs that were not related to false alarms (Fig. 1.S1E,F).

We also performed partial-correlation analyses to test for independent relationships between each model parameter and delay-related changes in behavior (i.e., between Δ predictive

bias and behavior after accounting for variability in *Δ* reactive bias and vice versa). *Δ* predictive bias was independently correlated with faster mean RTs (partial r=0.46, p=0.03), decreased RT variance (partial r=0.53, p=0.01), and an increased rate of false alarms (partial r=0.84, p<0.001). The relationship between *Δ* reactive bias and mean RTs was weaker and not statistically significant (partial r=0.39, p=0.07). However, in contrast to *Δ* predictive bias, *Δ* reactive bias was independently correlated with increased RT variance (partial r=0.71, p<0.001) and a decreased rate of false alarms (partial r=0.57, p=0.005). These results further support the argument that the two parameters of our model, *Δ* predictive bias and *Δ* reactive bias captured distinct forms of anticipatory processing that were guiding behavior.

We tested several alternatives to *Δ* reactive bias to explain anticipatory modulations in RTs that were not captured by the *Δ* predictive bias parameter, each of which provided a similar goodness-of-fit (mean R^2 across subjects was between 0.82–0.84 for each of these alternative models, compared to 0.84 for our main model). Within the LATER framework, a change in a rise-to-bound process can be modeled using any two of the following three parameters (for each pair, a change in the other parameter can always be matched by a certain coupled change in the value of each parameter in the pair): 1) a change in the mean of the rate-of-rise, 2) a change in the variance of the rate-of-rise, or 3) a change in the distance traveled (Noorani & Carpenter, 2016). We rejected all models that accounted for delay-related changes using three free parameters chosen from any two of the above plus *Δ* predictive bias because none resulted in a noticeable improvement in goodness-of-fit, despite the extra parameter. Neither alternative two-parameter models (a delay-related change in either the mean of the stimulus-triggered rise-to-bound process or the distance traveled, plus the *Δ* predictive bias) captured delay-related changes

in RT variability, and both had much stronger correlations in the values of the best-fitting model parameters across subjects than for our main model. We therefore conclude that the model parameterization we used, with *Δ* predictive bias and *Δ* reactive bias, provided the most parsimonious description of the behavioral data in terms of two distinct computational factors responsible for delay-related modulations of RTs.

Distributed neural representations of anticipatory influences on behavior

To identify neural signals related to the anticipation-dependent effects on RT behavior, we analyzed intracranial high-frequency neural activity (HFA, 70–200 Hz power) recorded from bipolar pairs of intraparenchymal depth electrodes implanted in our subjects with medically refractory epilepsy for clinical purposes (Fig. 2A). In total, we studied neural activity from 2,609 bipolar pairs of intraparenchymal depth electrodes distributed widely throughout the brain in 23 patients (mean=113.4 electrodes/subject), which we localized to various brain regions and intrinsic brain networks (Figs. 2B, 2.S1). These HFA signals, which are a reliable surrogate of local neural population spiking activity (Dubey & Ray, 2019; Manning et al., 2009; Ray & Maunsell, 2011), exhibited diverse forms of task-related sensory-, delay-, and motor-related processing at many electrodes (Figs. 3 and 3.S1).

We identified task-related neural representations from electrodes that showed activity modulations by delay condition and/or stochastic RT variability. Specifically, we measured the task-driven "activation function" of the neural population near each electrode by averaging z-scored HFA time-locked to target onset and motor response. We then identified all electrodes that had activation functions that were modulated by either the delay condition or stochastic

(delay-independent, trial-by-trial) RT variability, after accounting for transient sensory- and motor-driven responses (Fig. 3.S2).

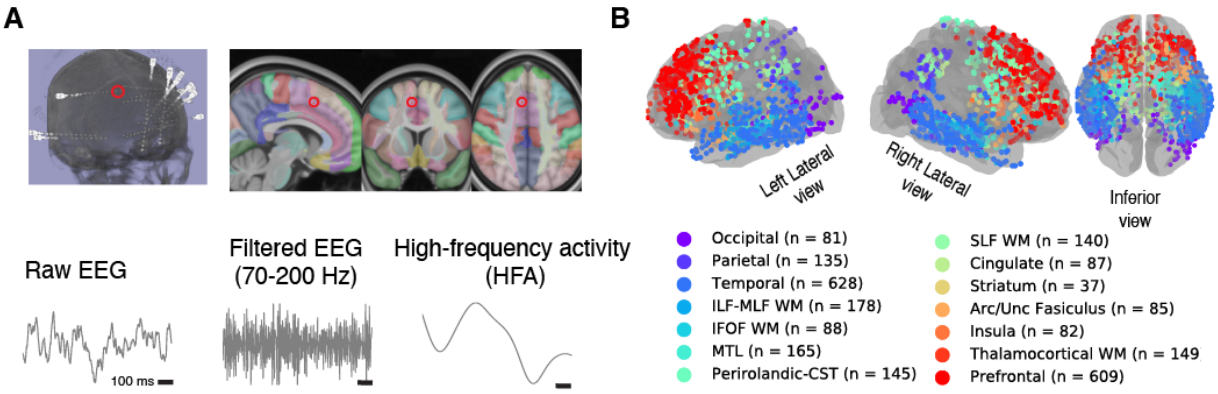

Figure 2: Measuring neural activity using intracranial electrodes (**A**) Extracting high-frequency activity (HFA, 70–200 Hz power) from intraparenchymal depth electrodes as an estimate of local spiking (red circle indicates example prefrontal electrode described in Fig. 3). (**B**) Brain plot showing electrode locations from all subjects in standard MNI coordinates. Colors indicate Regions of Interest (Table 4).

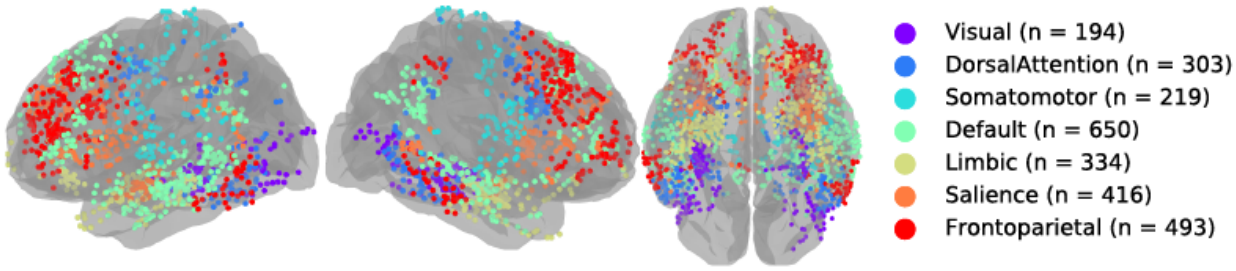

Figure 2 - Figure Supplement 1: Relating electrode locations to intrinsic brain networks. Brain plot showing electrode locations from all subjects in standard MNI coordinates as in Fig 2. Colors indicate intrinsic brain networks based on a normative atlas (Yeo et al., 2011)

In some cases, localized task-related neural representations could be mapped directly onto certain aspects of rise-to-bound model dynamics. For example, Fig. 3A shows activity recorded from an example electrode with pre-stimulus RT and delay modulations (the electrode location in prefrontal cortex is indicated in Fig. 2A). These modulations included increased pre-stimulus baseline activity that was relatively increased during faster RTs and during long-delay trials, followed by a largely RT-independent rate of rise before the motor response, as predicted by the model (Fig. 3B).

However, in general, correlates of both RT variability and delay condition were not restricted to such localized representations. Instead, these correlates were: 1) widespread across the brain, 2) heterogeneous (e.g., most brain regions contained electrodes that showed opposing direction of modulation, including increases and decreases in activity with faster RTs), and 3) dynamic (distinct electrodes showed modulations over the course of the trial; Figs. 3C and 3.S1-3).

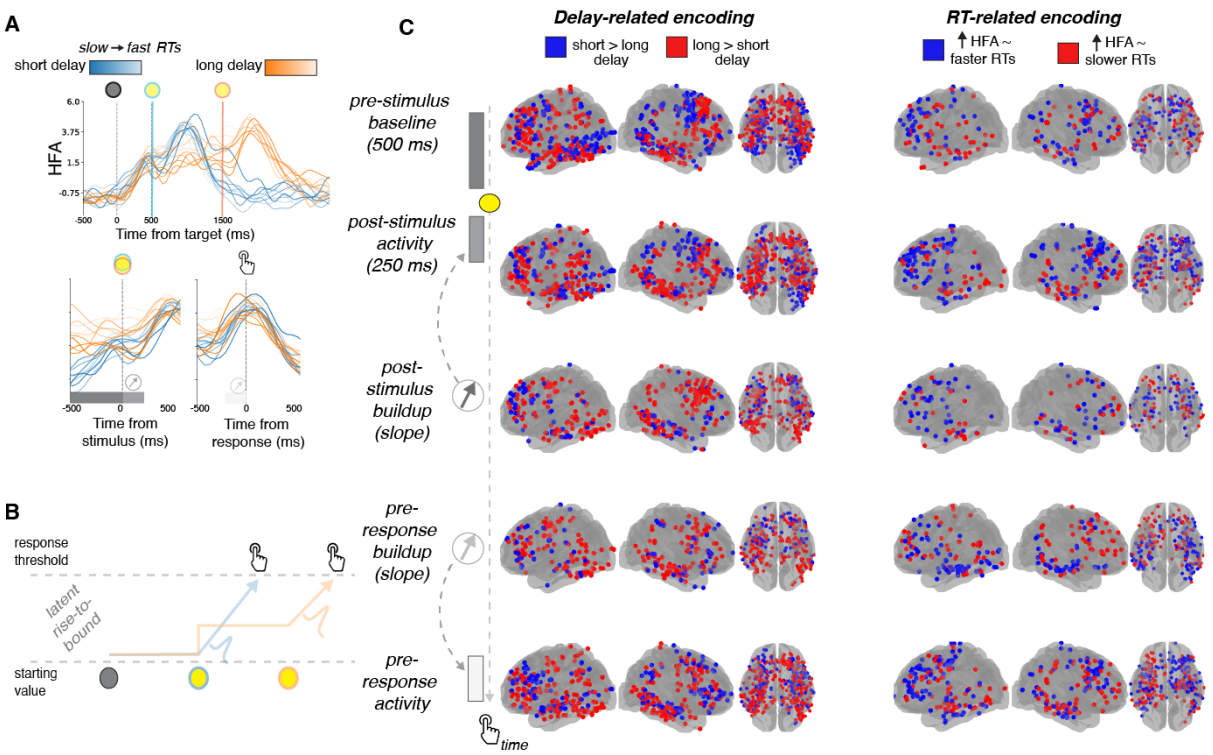

Figure 3: Task-related neural representations were widespread, heterogeneous, and dynamic. (**A**) Task-driven responses of local neural activity (average z-scored HFA) measured at the prefrontal electrode indicated in Fig. 2**A**, plotted separately for short- (blue) and long- (orange) delay trials and binned by stochastic RT percentile (10 bins; lighter shading indicates faster RTs; Fig. 3.S2). Top row shows target-locked activity (vertical lines indicate time of target onset (grey), short-delay color change (blue), and long-delay color change (orange), respectively). Bottom row shows stimulus- and response-locked activity (left and right, respectively). (**B**) Schematic illustrating how neural activity from this electrode can be mapped on directly to algorithmic "rise-to-bound" model dynamics. (**C**) Brain plots showing the anatomical distribution of delay- and RT-related neural representations (left and right columns, respectively, uncorrected $p<0.05$, non-parametric tests). Blue electrodes indicate negative effects (relatively increased activity during short delay/fast RTs); red electrodes indicate positive effects (relatively increased activity during long delay/slow RTs). Each row represents task-related activity patterns over time for different within-trial temporal dynamics (indicated on left, corresponding to task events illustrated in **B**).

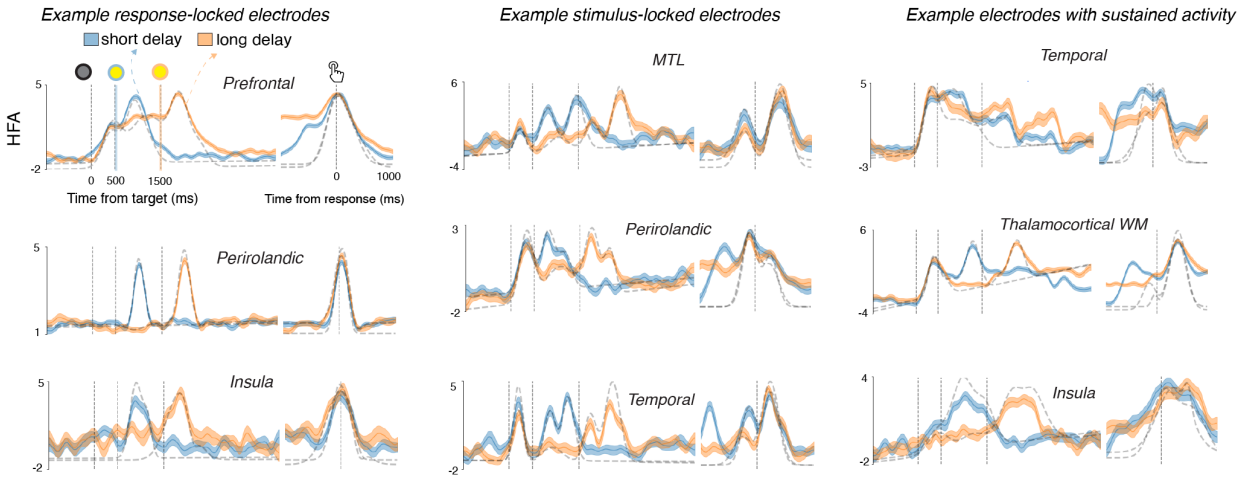

Figure 3 - Figure Supplement 1: Example task-driven activation functions. Each panel shows an activation function for a single electrode (average z-scored HFA locked to target and response) separately for short- (blue) and long- (orange) delay trials. Lines and ribbons are mean±sem across trials. Vertical lines indicate various task events, as indicated in the upper-left panel (which depicts activity recorded from the electrode from Figs. 2A and 3A). Dashed grey lines indicate model fits from the Gaussian curve-fitting algorithm used to remove stereotypical sensory and motor responses. Columns (left-to-right) show example electrodes with response-locked, stimulus-locked, and sustained activity patterns.

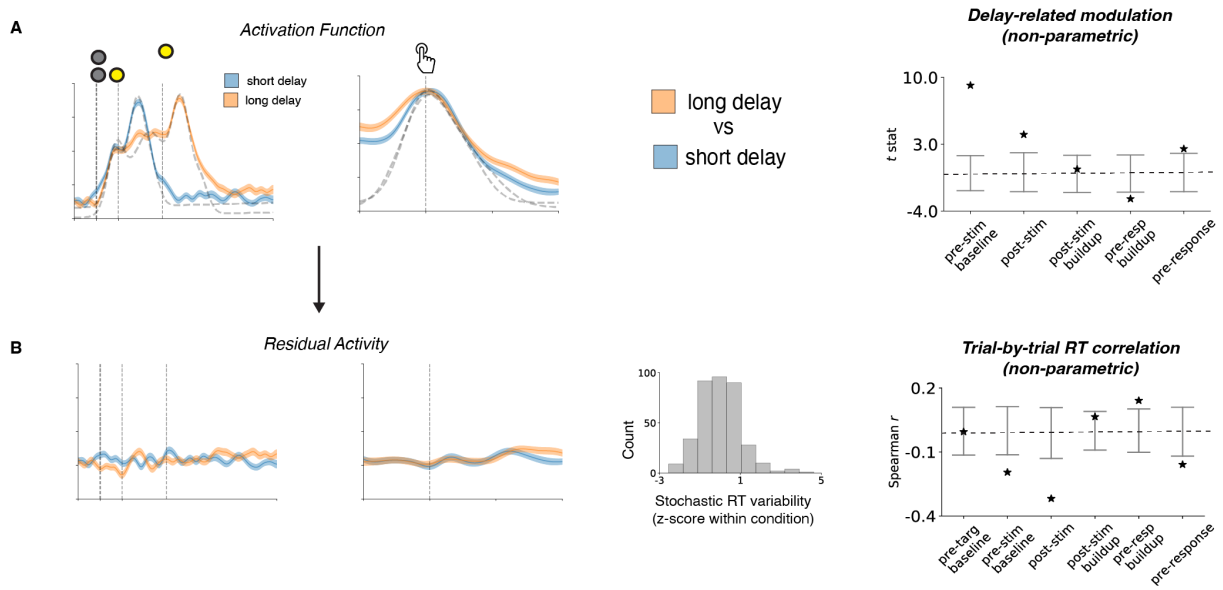

Figure 3 - Figure Supplement 2: Overview of methods used to relate each electrode's activation function to the delay condition and stochastic RT variability. (**A**) Example target-, stimulus-, and response-locked activation functions from a single electrode (same electrode as in Figs. 2A and 3A). (**B**) Same as in A, but with stereotypical sensory and motor responses (determined via Gaussian fits to data across trials) removed. Right panels show how activity measured at different times within a trial related to delay condition (top, using raw activation functions) and stochastic RT variability (bottom, using residual activation functions). Top: *s indicate two-tailed *t* statistic measured from the data comparing long- versus short-delay trials; error bars indicate 95% confidence intervals from non-parametric shuffle procedure of the *t* statistic indicating the difference on short- versus long-delay trials. Bottom: *s indicate Spearman *rhos* computed from the data between trial-by-trial z-scored RT and neural activity; error bars indicate 95% confidence intervals.

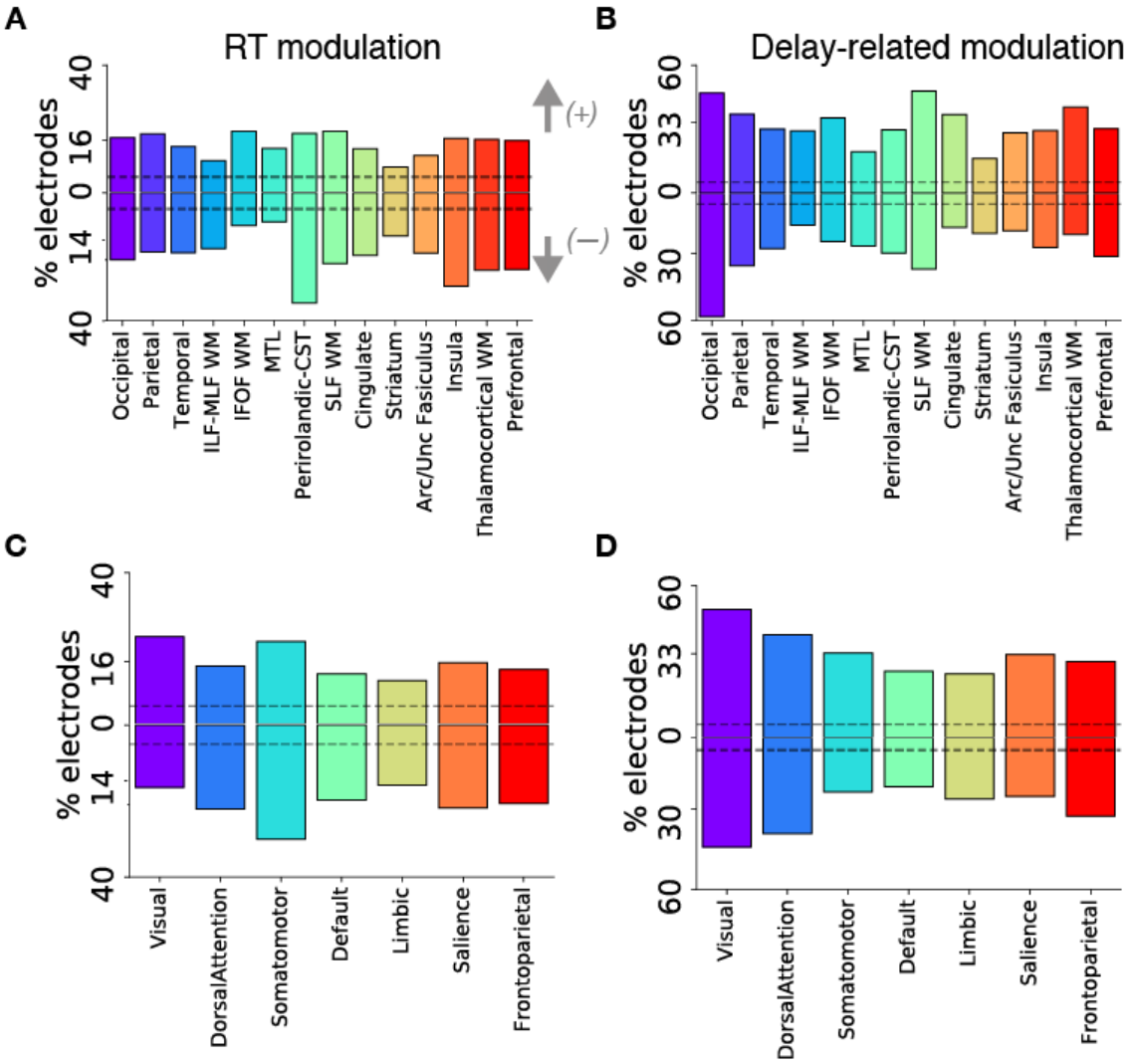

Figure 3 - Figure Supplement 3: Anatomical distribution of neural activity modulated by stochastic RT variability and delay condition. (**A**) Percentages of electrodes in each region with activity that was correlated with stochastic, trial-to-trial RT variability during any time interval. Negative values indicate increased activity with faster RTs; positive values indicate increased activity with slower RTs. (**B**) Percentages of electrodes in each region with activity that was modulated by delay condition during any time interval. Negative values indicate weaker activity on long-versus short-delay trials; positive values indicate increased activity on long-versus short-delay trials. (**C,D**) Same as **A** and **B**, respectively, but for the indicated intrinsic brain networks (Fig. 2.S1). Colors are matched to Fig. 2. and 2.S1.

Specifically, neural correlates of stochastic, trial-to-trial RT variability were distributed widely across the brain (*n*=638/2609, or 24.3% of electrodes; false-positive rate per omnibus test was 5% Fig. 2.S3A). Although these correlates of stochastic RT variability were most common in sensory (occipital, *n*=28/81, or 34.5% of electrodes) and motor (perirolandic, *n*=56/145, or 38.6% of electrodes) regions, they were also observed frequently in almost all sampled regions including deep structures such as the insula and MTL (χ^2test, FDR-corrected *p*s<0.05 for each region of interest, except striatum). Moreover, the neural activity patterns that were modulated by stochastic RT were heterogeneous. For example, most brain regions included separate neural populations that showed either increases or decreases in activity with faster RTs (Fig. 3.S3A). We observed similar patterns when studying the prevalence of these RT-modulated signals across various intrinsic brain networks (Figs. 3.S3C). The correlates of stochastic RT variability that we identified were also dynamic and varied throughout the course of a trial (Fig. 3C).

Likewise, neural modulations by foreperiod delay condition were widespread throughout the brain (*n* = 1271/2609, or 48.7% of electrodes; Fig. 3.S3B). Some regions had activity that tended to be higher for long versus short-delay trials (ILF-MLF WM, Cingulate, and Thalamocortical WM, all FDR-corrected *ps*<0.05), but otherwise we did not identify regional preferences for a particular direction of delay-related modulation (FDR-corrected *ps*>0.2). We observed similar patterns when studying the prevalence of these delay condition-modulated signals across various intrinsic brain networks (Figs. 3.S3D). We also found that electrodes showed distinct patterns of delay-related modulation during distinct time intervals throughout the task, suggesting the delay-related neural representations were dynamic (Fig. 3C). These delay-modulated signals also had diverse temporal dynamics, including differences in both transient

and sustained activity patterns between cue onset and movement onset for short- versus long-delay trials (Fig 3.S1).

Relating network dynamics to anticipatory behavior

Across these distributed electrodes with activity sensitive to RT and delay conditions, we identified two distinct sets of neural representations that, across subjects, were correlated with the two behaviorally derived components of our rise-to-bound models that best captured the effects of anticipation on behavior. We identified these neural representations from moment-to-moment fluctuations in activity measured between stimulus and response across all electrodes, which we quantified as trajectories in a high-dimensional coordinate space (Ebitz & Hayden, 2021; Humphries, 2021; Vyas et al., 2020; Figs. 4A and 4.S1).

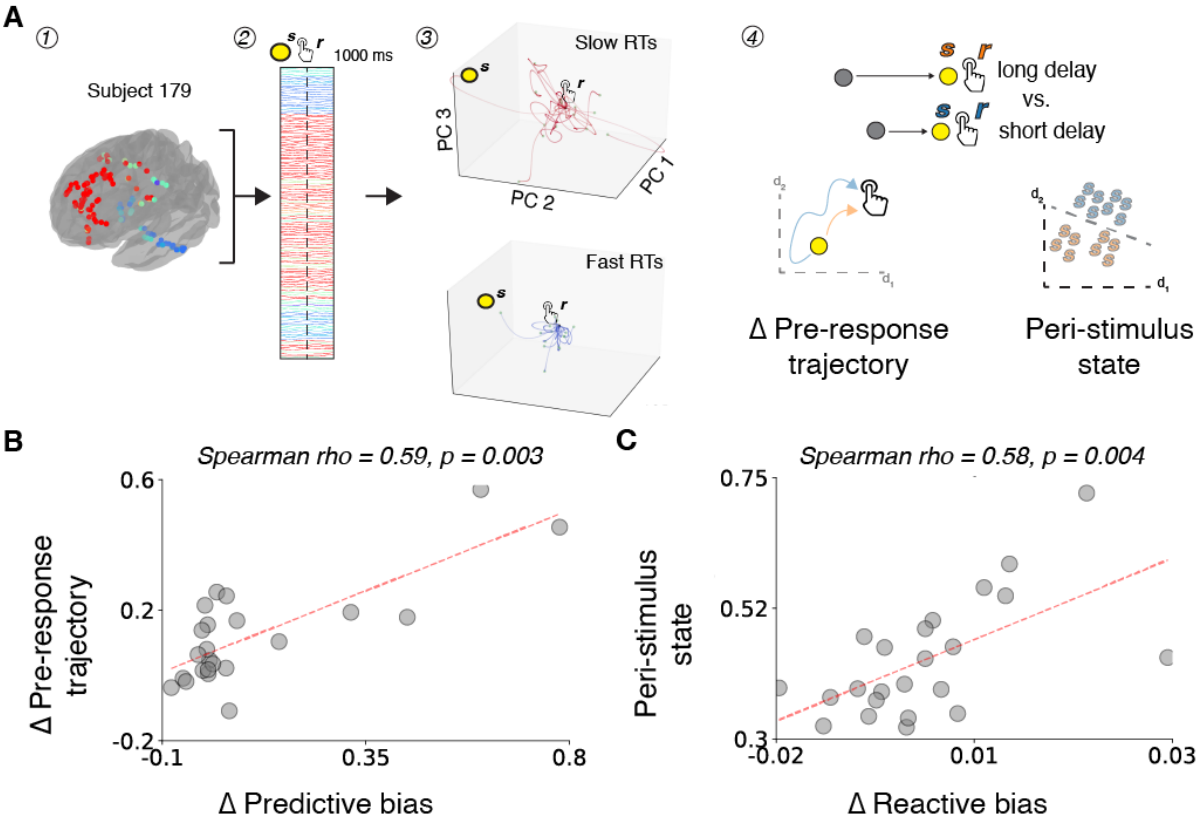

Figure 4: Predictive and reactive anticipatory biases are reflected in distinct latent, brain-wide neural dynamics. (**A**) Illustration of methods used to study relationships between anticipatory biases on behavior and network dynamics in a single example subject (#179). Left-to-right: 1) electrode locations in standardized MNI coordinates; 2) example segments of HFA from each electrode for a single trial (1000 ms of stimulus-locked activity; each row indicates an electrode colored by region of interest as shown in Fig. 2); 3) reduced three-dimensional representations of widespread HFA measured from stimulus onset to response, shown for the 15 slowest RTs (red, left panel) and 15 fastest RTs (blue, right panel), where each "neural trajectory" is drawn centered on the neural state at the time of the motor response (finger press, black dot; yellow circle represents neural state at the start of one trial); and 4) two features of these network dynamics that showed delay-related differences that correlated with anticipatory influences on behavior. (**B**) Across-subject relationship between behaviorally derived Δ predictive bias (abscissa) and neural Δ pre-response trajectory (ordinate; see **A**). (**C**) Across-subject relationship between behaviorally derived Δ reactive bias (abscissa) and neural peri-stimulus state (ordinate; see **A**). In **B** and **C**, each circle corresponds to data from a single subject; lines are linear fits.

First, subjects with more direct network trajectories between stimulus and response onset on long- versus short-delay trials ("pre-response trajectory", which is conceptually related to the rate of rise in the rise-to-bound behavioral model) tended to have larger Δ predictive biases (i.e., a greater tendency to respond more quickly on long-delay trials, with an associated increase in false-alarms; *rho*=0.59, FDR-corrected *p*=0.021; Figs. 4B and 4.S2). That is, this measure of pre-response network dynamics was related to the "predictive" form of processing in the model that accounted for false alarms. Second, subjects with greater differences between network state just before the short- versus long-delay cues ("peri-stimulus state," which is conceptually related to the starting point in the rise-to-bound model) tended to have higher Δ reactive biases (i.e., they tended to respond more quickly and with more variability on long-delay trials, but without an increase in false alarms; *rho*=0.58; FDR-corrected *p*=0.021; Fig. 4C and 4.S2). That is, this measure of peri-stimulus network state was related to the "reactive" form of processing in the model that accounted for delay-related changes in RT that were not coupled to false alarms.

Several features of these relationships support the idea that anticipatory modulations of RT involve widespread neural signals that can be related to specific behaviorally derived computations. First, these relationships depended on activity patterns that were highly distributed across many electrodes, as quantified by Principal Components Analysis (Fig. 4.S1). Second, these relationships did not simply reflect differences in mean RT for the two delay conditions but rather more nuanced features of the model-derived parameters: pre-response trajectory was correlated with Δ predictive bias (partial *rho*=0.58, *p*=0.004) and peri-stimulus state was correlated with Δ reactive bias (partial *rho*=0.55, *p*=0.008) even after accounting for delay-related differences in mean RT. Third, the pre-response trajectory was also positively correlated

with Δ reactive bias (Spearman's *rho*=0.55, FDR-corrected *p*=0.021), but this correlation was not significant after accounting for delay-related differences in mean RT (partial *rho*=0.36, *p*=0.09). We did not observe any other significant correlations between best fitting rise-to-bound model parameters and other measures of network dynamics on short- versus long-delay trials (FDR-corrected *p*s>0.4; Fig. 4.S2). Thus, below we focus on better understanding the spatial localization of signals that contributed to the representation of predictive biases in activity just before the response and the representation of reactive biases in activity around the time of the cue.

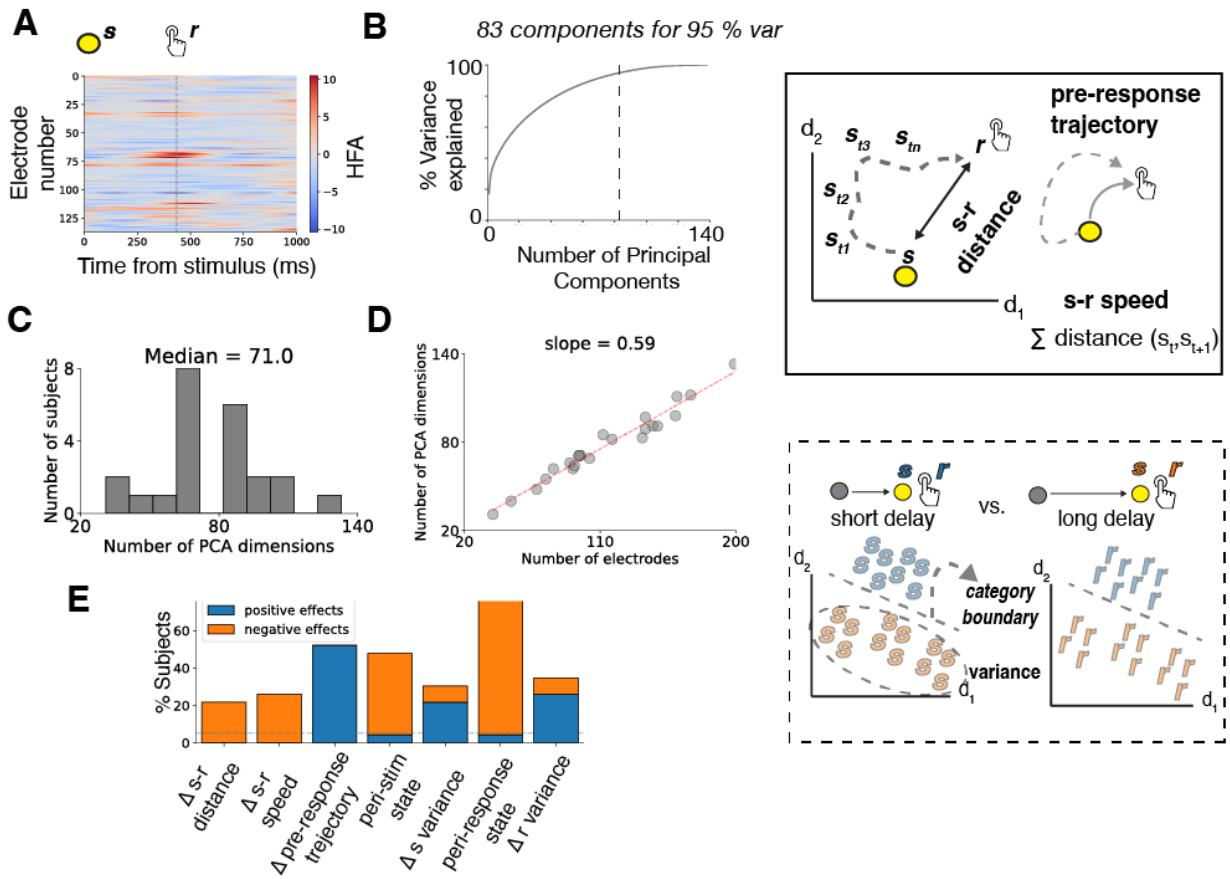

Figure 4 - Figure Supplement 1: Measuring widespread network dynamics from simultaneous multi-site recordings. (**A**) Color plot representing simultaneous recordings from 138 electrodes (ordinate) as a function of time within a single trial relative to stimulus onset (abscissa). Same data as Fig. 4A. Red/blue indicate positive/negative z-scored HFA values. Grey vertical dashed line indicates time of response on this trial (finger press). Network dynamics are computed between time of stimulus and response on each trial. (**B**) Network-state variability across electrodes over time was high dimensional: 83 principal components were required to explain 95% of variability (dashed line). (**C**) Network dynamics show high dimensional variability at the group level. Histogram across subjects of the number of components needed to explain 95% of the variance per subject. (**C**) Correlation between the number of components needed to explain 95% of the variance per subject and the number of implanted electrodes in each subject. (**E**) Number of subjects that showed significant delay-related modulations in neural-state dynamics (non-parametric test, $p < 0.05$). The boxes to the right show schematics illustrating methods used to quantify network state dynamics. *Solid box:* Schematic illustrating measures used to quantify trial-to-trial variability neural-state dynamics using a schematic 2-dimensional representation of neural-state space. *Dashed box:* Schematics illustrating measures of changes in network-state dynamics related to the delay condition. Left and Right panels illustrate discriminability (determined by the category boundary) and variance (illustrated as the spread of activity) for stimulus and response neural states, respectively.

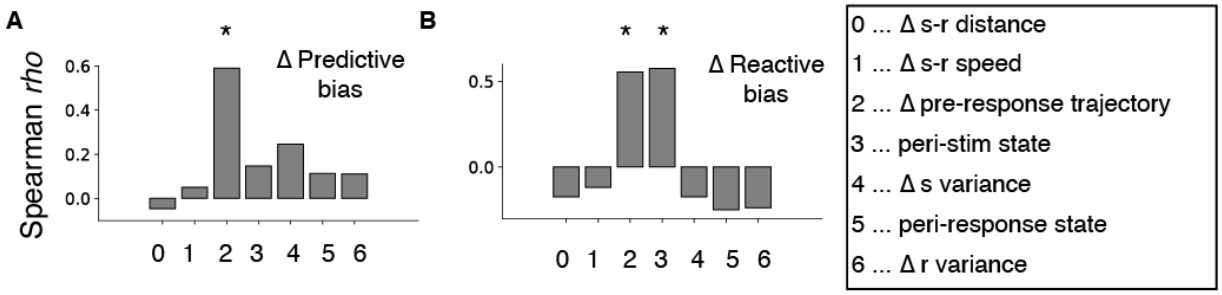

Figure 4 - Figure Supplement 2: Relating network dynamics to anticipatory modulation of RT variability. Spearman correlations between individual differences in delay-related changes in Δ predictive bias (**A**) and Δ reactive bias (**B**) and various neural dynamics (Fig. 4, 4.S1). * indicates FDR-corrected $p<0.05$

Localized neural representations of anticipatory influences on behavior

Because task-related neural representations were heterogeneous and distributed across the brain, we grouped electrodes based on similar activity modulations related to stochastic RT variability and delay condition using data-driven unsupervised clustering (Figs. 5.S1–3). This approach allowed us to study specific neural representations of task-relevant events that were distributed across many regions and avoid arbitrarily aggregating activity from nearby neural populations that might have heterogeneous response properties and functions. We identified 20 clusters with similar activity profiles using an objective function that maximized the identification of distinct activity patterns that were reproducible across subjects in our dataset (Fig. 5A).

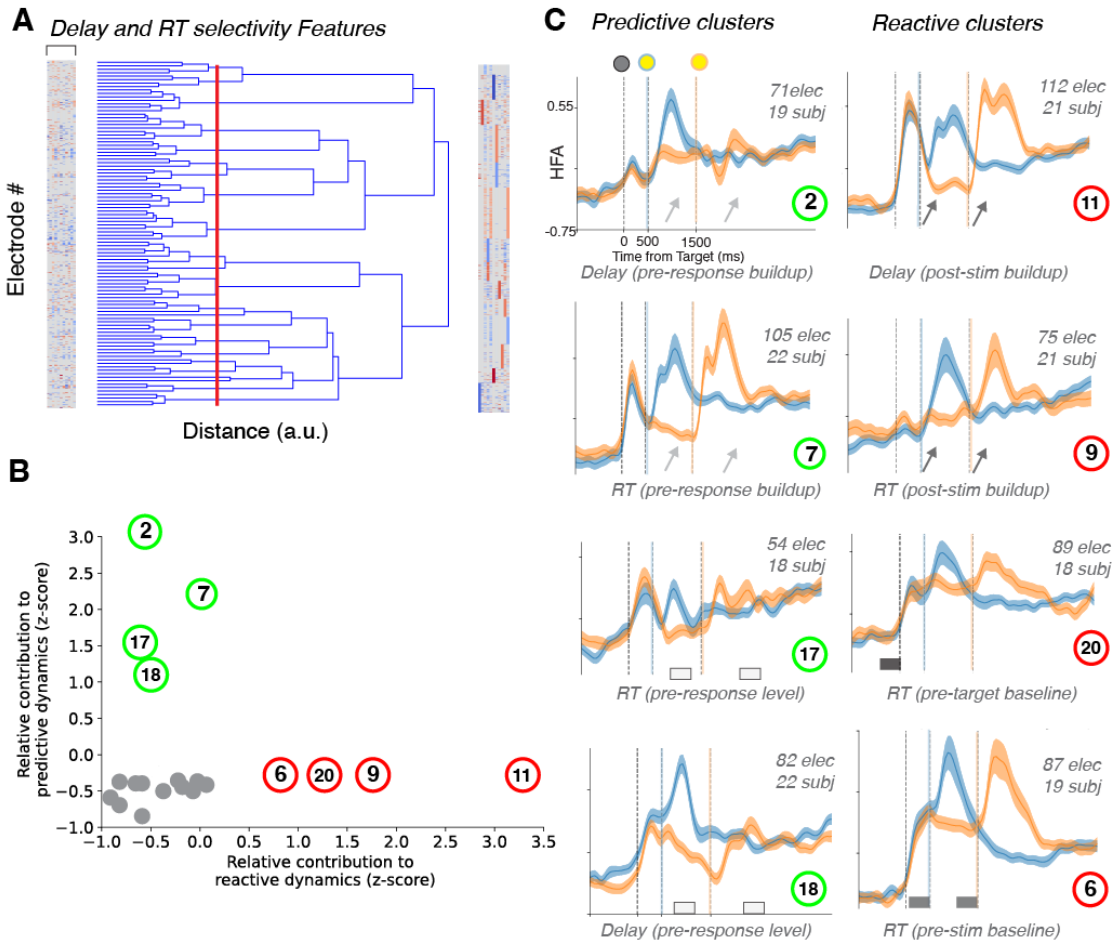

Figure 5: Predictive and reactive anticipatory biases were reflected in distinct neural populations. (A) Illustration of our data-driven clustering of electrodes based on similar task-related activity modulations (left: colormap representing a feature matrix across all electrodes, where each row represents an electrode, and each column indicates the correlation with RT or delay modulation across various time intervals as in Fig. 3A; middle: dendrogram representing similarities between electrodes; right: feature matrix re-organized based on similar RT and delay modulations at the clustering level indicated by the red line in middle panel, which we used in these analyses). (**B**) Scatterplot showing the relative contributions of each functionally defined cluster to "predictive dynamics" (i.e., the across-subject correlation between Δ predictive bias and Δ pre-response trajectory shown in Fig. 4B; ordinate) and "reactive dynamics" (i.e., the across-subject correlation between Δ reactive bias and peri-stimulus state shown in Fig. 4C; abscissa). (**C**) Functional profiles of clusters that showed strong influences on predictive dynamics (left column, "predictive clusters") or reactive dynamics (right column, "reactive clusters"). For each cluster, we show average target-locked activation functions to illustrate functional differences across the clusters (blue and orange indicate activity on short- and long-delay trials, respectively; lines and ribbons are mean±sem across electrodes). Each cluster is labeled below based on the predominant feature of task-related activity modulation that is shared among electrodes in the group.

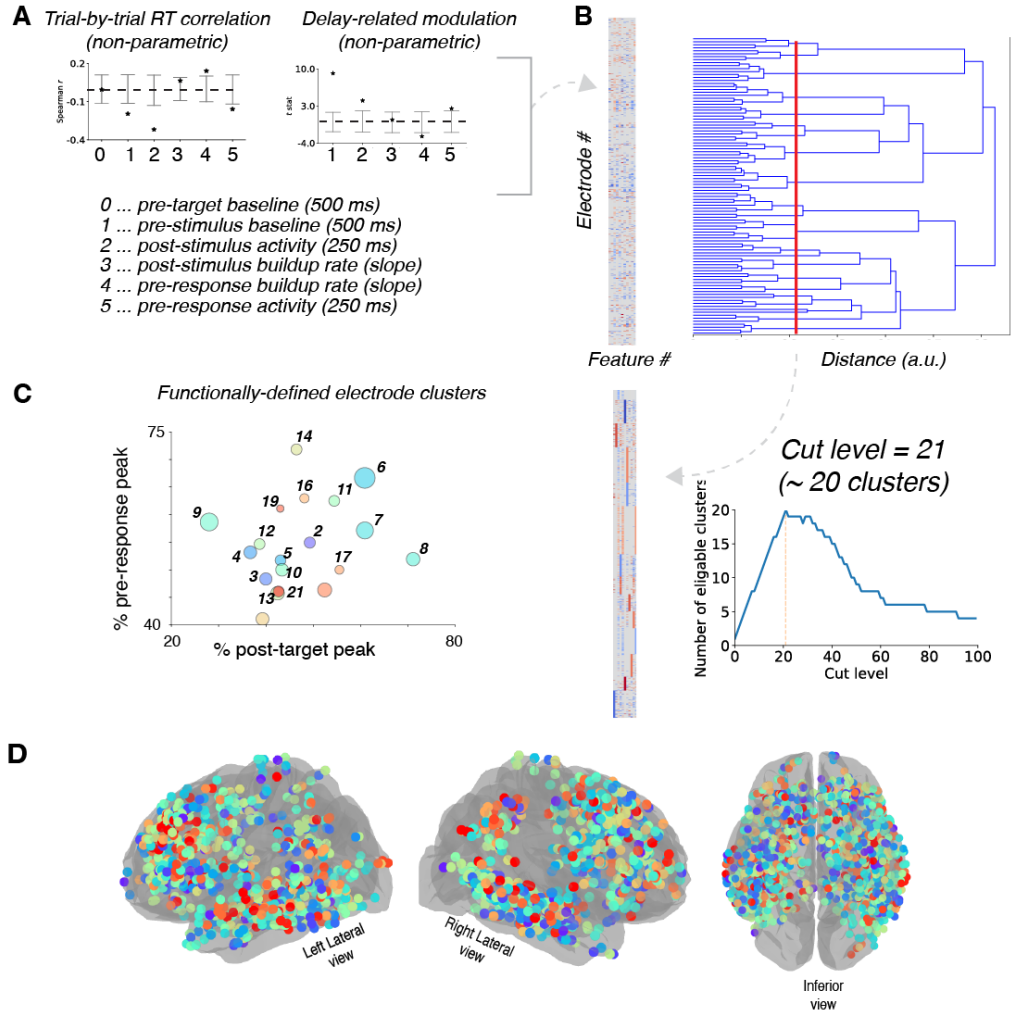

Figure 5 - Figure Supplement 1: Overview of the hierarchical clustering algorithm. (**A**) Example feature vector for a single electrode, representing the presence and direction of a significant modulation by stochastic RT and delay condition ($p < 0.05$, non-parametric tests, 0 for non-significant relation, -1 for significant negative relation, and +1 for significant positive relation; same data as Fig. S3). (**B**) Overview of the clustering algorithm shown in Fig. 5A. Each electrode from a given patient was described by a pair of feature vectors as in **A** (blue/red marks in the left panels show positive and negative *z*-statistics relative to null distributions, respectively; color range -4.5 to 4.6). These feature vectors were used as input to a clustering algorithm. We identified a level of clustering (red vertical line in dendrogram) based on an objective function that maximized the number of clusters that were well sampled in our dataset (i.e., each cluster consisted of at least 50 electrodes from at least 5 subjects). The main analyses used a clustering level of 21, resulting in 20 eligible clusters, as depicted in the top row). (**C**) Scatterplot illustrating each cluster (circles) based on two functional properties of the responses measured from electrodes in that cluster: 1) percentage of electrodes showing post-target peaks in activity (abscissa), and 2) percentage of electrodes showing pre-response peaks in activity

(ordinate). Size of circles indicate degree of anatomical localization (larger circles are more localized, based on χ^2 statistics comparing anatomical distribution to uniform).

In general, these clustering analyses identified: 1) a much greater diversity of activity patterns than the kinds of joint modulations by delay condition and stochastic RT variability expected from simple versions of the rise-to-bound model, and 2) a much more widespread anatomical distribution of task-related responses than just premotor areas that have been the targets of many previous studies (Figs. 5 and 5.S1–3). For example, the clusters identified activity patterns that differed in terms of not just the magnitude of post-target and pre-response neural responses (Fig. 5.S1C) and modulations by error feedback (Fig. 5.S2, left column), false alarms (Fig. 5.S2, middle column), and spatial selectivity (Fig. 5.S2, right column), but also in terms of the timing and magnitude of ramp-like responses that depended on delay condition (e.g., clusters 2, 3, 8, 11–13, and 19) and of other patterns that covaried with RT variability (e.g., clusters 4–7, 9, 10, 14–17, 20). Generally, these clusters each included electrodes throughout the brain, with no obvious bias towards sensory, motor, or sensory-motor areas (except cluster 6, 7 and 9 that showed increased prevalence in perirolandic, occipital and insular cortex, respectively).

Additionally, we identified clusters with activity patterns that appeared to be relevant for anticipatory information processing, including several clusters with activity modulations related to the short- versus long-delay conditions (e.g., clusters 2, 3, 8, 11–13, and 19). Some of these modulations affect ramp-like signals that are roughly consistent with certain theoretical models of how a neural population in aggregate could represent represent elapsed time from a sensory event (Salet et al., 2022; Shankar & Howard, 2012). Thus, these clustering analyses showed a much more widespread anatomical distribution than just premotor areas that have been the

targets of many previous studies and a much greater diversity of activity patterns than the kinds of joint modulations by delay condition and stochastic RT variability expected from simple versions of the rise-to-bound model (Fig. 3).

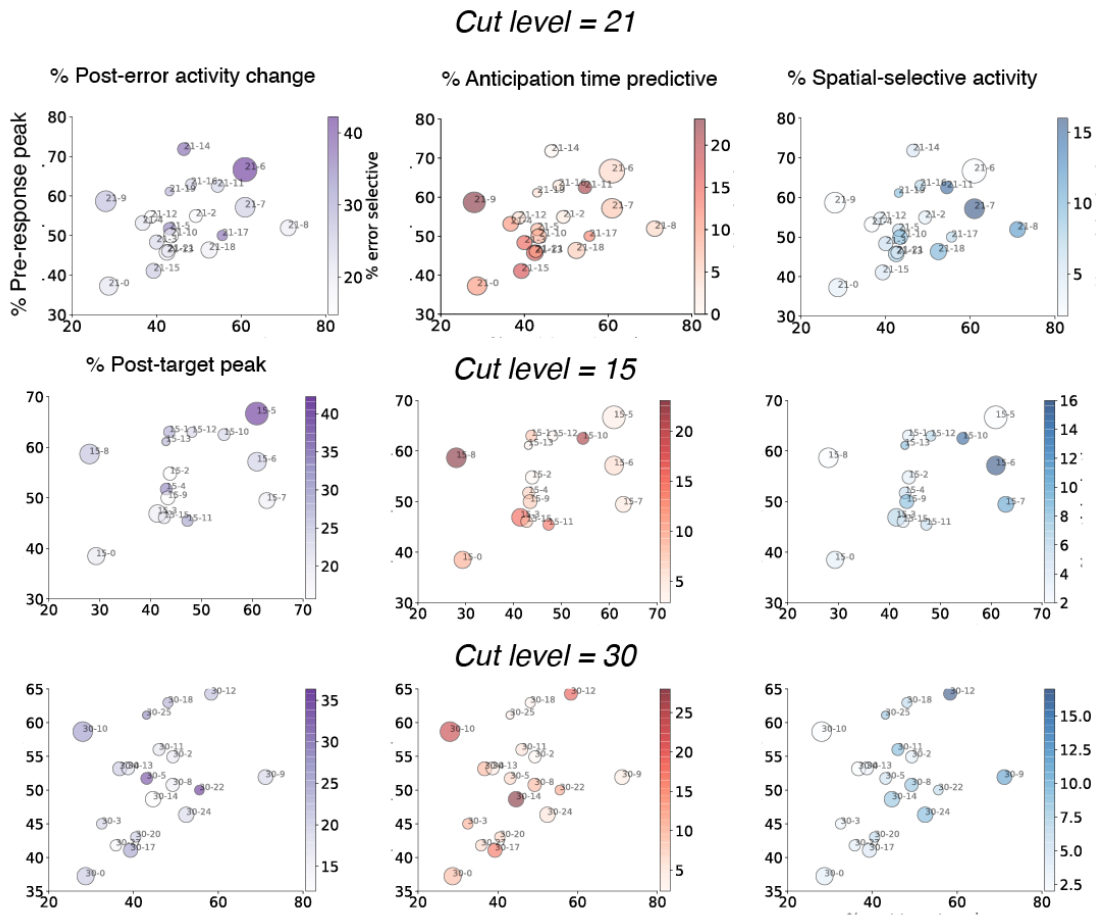

Figure 5-Figure Supplement 2: Functional properties of clusters across different clustering levels. Each panel shows a scatterplot as in Fig. 5.S1C, but color coded to indicate additional task-related functional properties (based on parametric statistical tests), as follows. Left column (purple) shows percentage of electrodes that showed altered activity following errors (as compared to control trials, *t*-test, $p<0.05$). Middle column (red) shows percentage of electrodes that showed correlations between post-target activity and timing of anticipatory false alarms (Spearman correlation, $p<0.05$). Right column (blue) shows percentage of electrodes that showed differential post-target activity based on the location of the target on the screen (one way ANOVA, $p<0.05$). Top row shows clustering level used in main analyses (20 clusters). Bottom two rows show functional properties of clusters for two different clustering levels, illustrating qualitatively similar distributions of functional properties.

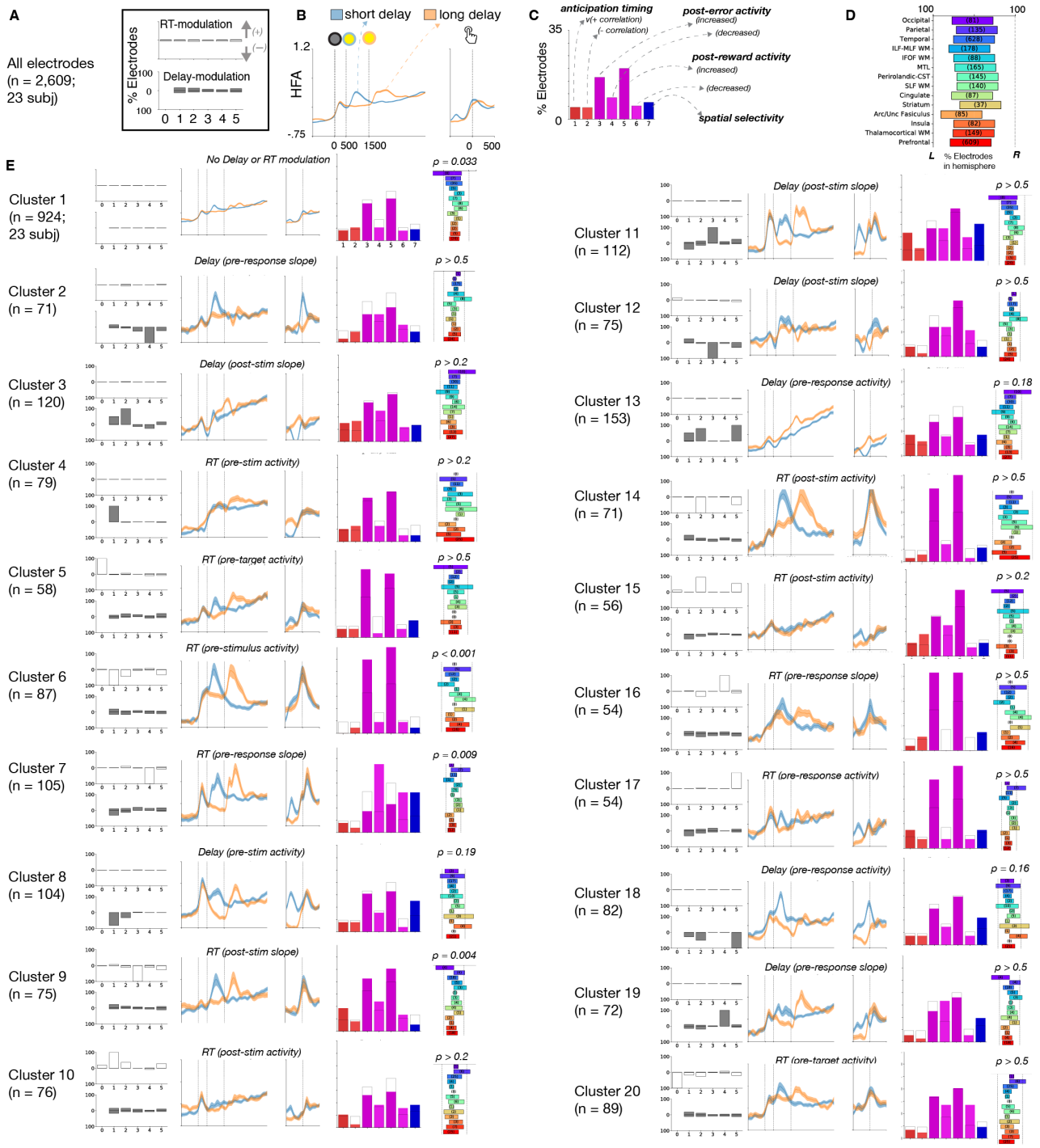
A
All electrodes
(n = 2,609;
23 subj)
% Electrodes
RT-modulation
Delay-modulation
B
short delay
long delay
HFA
C
anticipation timing
(+ correlation)
(- correlation)
post-error activity
(increased)
(decreased)
post-reward activity
(increased)
(decreased)
spatial selectivity
% Electrodes
D
Occipital
Parietal
Temporal
ILF-MLF WM
IFOF WM
MTL
Perirolandic-CST
SLF WM
Cingulate
Striatum
Arc/Unc Fasiculus
Insula
Thalamocortical WM
Prefrontal
% Electrodes
in hemisphere
L
R
E
Cluster 1
(n = 924;
23 subj)
No Delay or RT modulation
p = 0.033
Cluster 2
(n = 71)
Delay (pre-response slope)
p > 0.5
Cluster 3
(n = 120)
Delay (post-stim slope)
p > 0.2
Cluster 4
(n = 79)
RT (pre-stim activity)
p > 0.2
Cluster 5
(n = 58)
RT (pre-target activity)
p > 0.5
Cluster 6
(n = 87)
RT (pre-stimulus activity)
p < 0.001
Cluster 7
(n = 105)
RT (pre-response slope)
p = 0.009
Cluster 8
(n = 104)
Delay (pre-stim activity)
p = 0.19
Cluster 9
(n = 75)
RT (post-stim slope)
p = 0.004
Cluster 10
(n = 76)
RT (post-stim activity)
p > 0.2
Cluster 11
(n = 112)
Delay (post-stim slope)
p > 0.5
Cluster 12
(n = 75)
Delay (post-stim slope)
p > 0.5
Cluster 13
(n = 153)
Delay (pre-response activity)
p = 0.18
Cluster 14
(n = 71)
RT (post-stim activity)
p > 0.5
Cluster 15
(n = 56)
RT (post-stim activity)
p > 0.2
Cluster 16
(n = 54)
RT (pre-response slope)
p > 0.5
Cluster 17
(n = 54)
RT (pre-response activity)
p > 0.5
Cluster 18
(n = 82)
Delay (pre-response activity)
p = 0.16
Cluster 19
(n = 72)
Delay (pre-response slope)
p > 0.5
Cluster 20
(n = 89)
RT (pre-target activity)
p > 0.5

Figure 5 Figure Supplement 3: Cluster descriptions. (**A–D**) Anatomical and functional properties used to describe each cluster, here illustrated using data from all (unclustered) electrodes. (**A**) Bar plot showing frequency of RT- and delay-modulation at various time intervals. Positive and negative values indicate frequency of electrodes showing positive effects and negative effects, respectively (as described in Fig. 3.S3). (**B**) Average activation function (same format as Fig. 3.S1). (**C**) Bar plot summarizing the frequency of activity modulations ($p < 0.05$) related to: false-alarm response times (1, 2), post-error changes in activity (3,4), post-reward changes in activity (5,6), and spatial location of the stimulus (7); (**D**) Anatomical distribution of electrodes. Horizontal bars indicate proportions of electrodes in each region (bar width) and per hemisphere (L/R, indicated by offset relative to center). Vertical dashed lines indicate uniformly distributed proportions. (**E**) Properties shown in **A–D** for each cluster. Each cluster is labeled by the dominant RT or delay-related effect. *P*-values indicate deviation of anatomical distribution of the cluster from expected distribution (χ^2 test, as shown in **D**).

To relate these diverse clusters to the behaviorally relevant network dynamics shown in Fig. 4, we quantified the relative contribution of electrodes within each cluster to the across-subject correlations related to predictive processing (Δ predictive bias and pre-response trajectory) and reactive processing (Δ reactive bias and peri-stimulus state). Specifically, we measured a relative contribution score for each cluster that quantified the change in the effect size of given across-subject correlation after excluding electrodes within that cluster and correcting for the number of electrodes in the cluster. We used *z*-scored relative contribution scores to identify clusters with the strongest relative influence on particular anticipatory network dynamics. We found two distinct sets of clusters that each showed relatively strong influences on just one of the two behaviorally relevant anticipatory network dynamics described above, as follows (Fig. 5A; see Figs. 5.S1–3 for a more detailed description of these clusters).

The first set of "predictive" clusters made the strongest contributions to the correlation between Δ predictive bias and pre-response trajectory but minimal contributions to the correlation between Δ reactive bias and peri-stimulus state (clusters 2, 7, 17, 18, each of which had z-scores >0.5 for pre-response trajectory but <0 for peri-stimulus state; colored green in Fig.

5). These predictive clusters generally showed delay and RT-modulations near the time of response (e.g., modulations of pre-response build up in Clusters 2 and Cluster 7 and of pre-response level in Clusters 17 and 18), along with attenuated activity on long-delay trials, when expectation was relatively high (Cluster 2 and Cluster 18; see Fig. 6A for the aggregate activation function across these 4 clusters).

The second set of "reactive" clusters made the strongest contribution to the correlation between Δ reactive bias and peri-stimulus state but minimal contributions to the correlation between Δ predictive bias and pre-response trajectory (clusters 11, 9, 21, 6, each of which had z-scores >0.5 for peri-stimulus state, but <0 for pre-response trajectory, colored red in Fig. 5). These reactive clusters generally showed modulations near the time of stimulus processing (e.g., modulation of post-stimulus buildup in Clusters 11 and 9, of pre-target baseline in Cluster 20, and of pre-stimulus baseline in Cluster 6), along with increased activity on long-delay trials, when expectation was relatively high (Cluster 11; see Fig 6A for the aggregate activation function across these 4 clusters).

The electrodes in these predictive and reactive clusters were widely distributed and interspersed throughout the brain, but showed clear differences in their anatomical distributions (Fig. 6B, electrodes contributing to any predictive cluster are colored green and to any reactive cluster are colored red). These differences were evident when comparing their distributions across brain regions (χ^2 statistic = 36.5, $p < 0.001$, $df = 13$) and intrinsic brain networks (χ^2 statistic = 14.4, $p = 0.025$, $df = 6$, Fig. 6.S1). For example, the medial temporal lobe (MTL) had a relatively high prevalence of predictive clusters (more predictive clusters and fewer reactive clusters than expected by chance given their overall prevalence, $n = 26/165$ predictive versus

11/165 reactive; similar trends for Limbic networks). In contrast, perirolandic cortex had a relatively high prevalence of reactive clusters (more reactive clusters and fewer predictive clusters than expected, n = 37/145 reactive versus 10/145 predictive; similar trends for somatomotor networks; Fig 6C, D). However, in general multiple brain regions and multiple networks contributed to both forms of anticipatory processing.

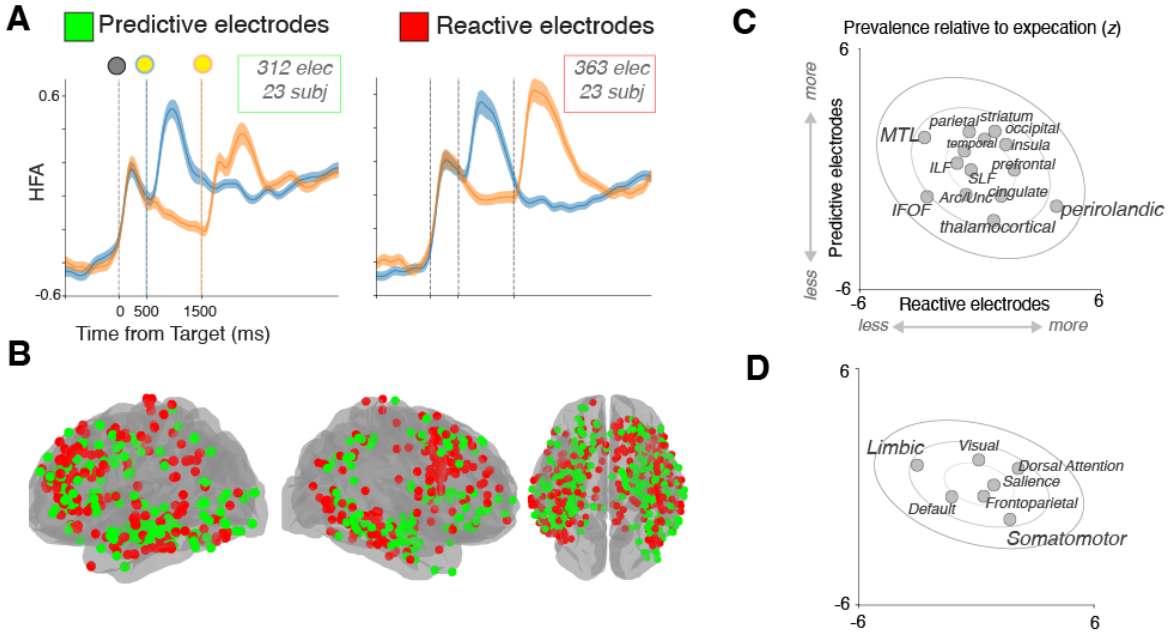

Figure 6: Comparing neural populations contributing to predictive and reactive anticipatory biases. (A) Aggregate activation functions for electrodes in predictive and reactive clusters (left and right panel, respectively; same format as Fig. 3.S1). (**B**) Brain plot showing anatomical distribution of predictive (green) and reactive electrodes (red). (**C**) Scatterplot showing the relative frequency of predictive and reactive electrodes in each brain region of interest relative to their overall (expected) frequency across the brain (z-scores; positive values indicate increased relative frequency; negative values indicate decreased relative frequency). Regions towards the top left show a bias towards predictive clusters, whereas regions towards the bottom right show a bias towards reactive clusters. Ellipses indicate confidence intervals derived from the joint distribution (from largest, middle and smallest indicates 3, 2 and 1 σ, respectively). (**D**) Same as **C**, but for intrinsic brain networks.

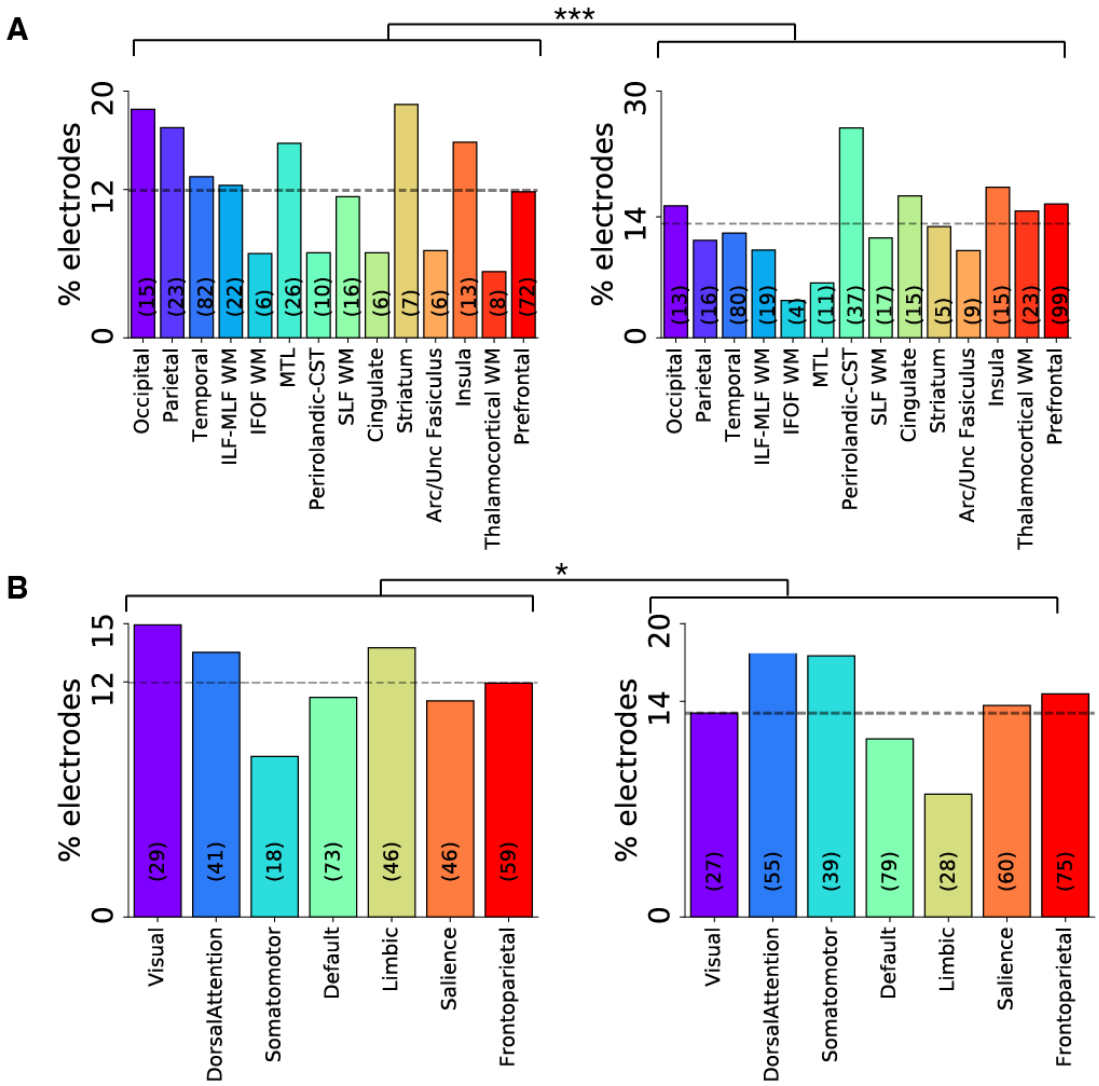

Figure 6 - Figure Supplement 1: Comparing anatomical distributions of predictive and reactive electrodes. (**A**) Percentages of electrodes in each region that contributed to predictive and reactive clusters (left and right panel, respectively). Dashed line indicates percentage expected based on overall prevalence across the brain. *** indicates $p < 0.001$; * indicates $p < 0.05$. (**B**) Same as **A** but for the indicated intrinsic brain networks (Fig. 2.S1). Colors are matched to Fig. 2. and 2.S1.

Comparing visuospatial and movement-related electrodes

To better appreciate the distributed nature of these representations of predictive and reactive signals, we performed similar analyses on lower-level sensory and motor signals that presumably arise from more localized regions of the brain. Specifically, we identified "visuospatial" electrodes as those that showed activity modulations based on the spatial position of the stimulus on the screen during the 500 ms after initial presentation (n = 173 electrodes across 22 subjects; one-way ANOVA, $p < 0.05$). In addition, we identified "movement-related" electrodes as those that showed pre-response peaks within 100 ms of the button press based on Gaussian modeling of activation functions (n = 119 electrodes across 20 subjects; see Fig. 7A for aggregate activation functions). Visuospatial and movement-related electrodes showed distinct anatomical distributions across brain regions (χ^2 statistic = 36.5, $p < 0.001$, $df = 13$) and intrinsic brain networks (χ^2 statistic = 14.4, $p < 0.001$, $df = 6$, Fig. 7.S1) that were less uniformly distributed than for the predictive and reactive representations (compare Figs. 6 and 6.S1 with Figs. 7 and 7.S1). Visuospatial electrodes had a very high prevalence in occipital cortex and the Visual network, and a relatively low prevalence elsewhere. Likewise, movement-related electrodes had a very high prevalence in perirolandic cortex (particularly in the left hemisphere, which is consistent with the fact that participants pressed the response button with their right hand) and the Somatomotor network, and a relatively low prevalence elsewhere. These results support our interpretation that anticipatory influences on simple sensory-motor behaviors involve brain networks that were more distributed than for simpler sensory or motor processing alone and whose distributed nature was not simply artifacts of our data-collection or analysis approaches.

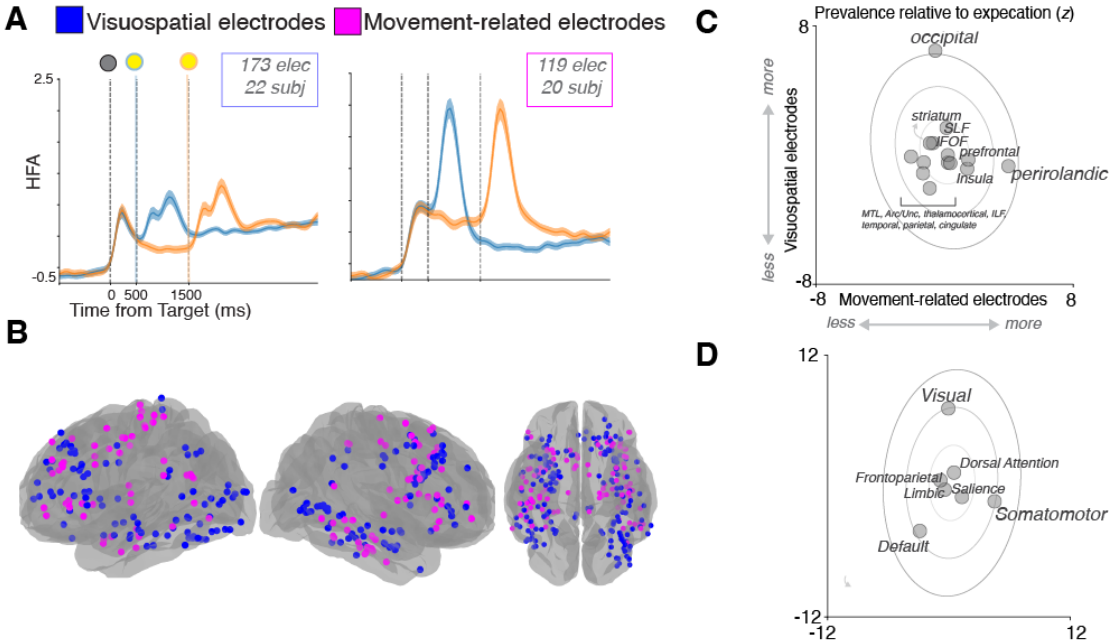

Figure 7: Comparing visuospatial and movement-related electrodes. (A) Aggregate activation functions for visuospatial and movement-related electrodes (left and right panel, respectively; same format as Fig. 3.S1). (**B**) Brain plot showing anatomical distribution of visuospatial (blue) and movement-related electrodes (magenta). (**C,D**) Same as Fig. 6 C,D, but for visuospatial (abscissa) and movement-related electrodes (ordinate).

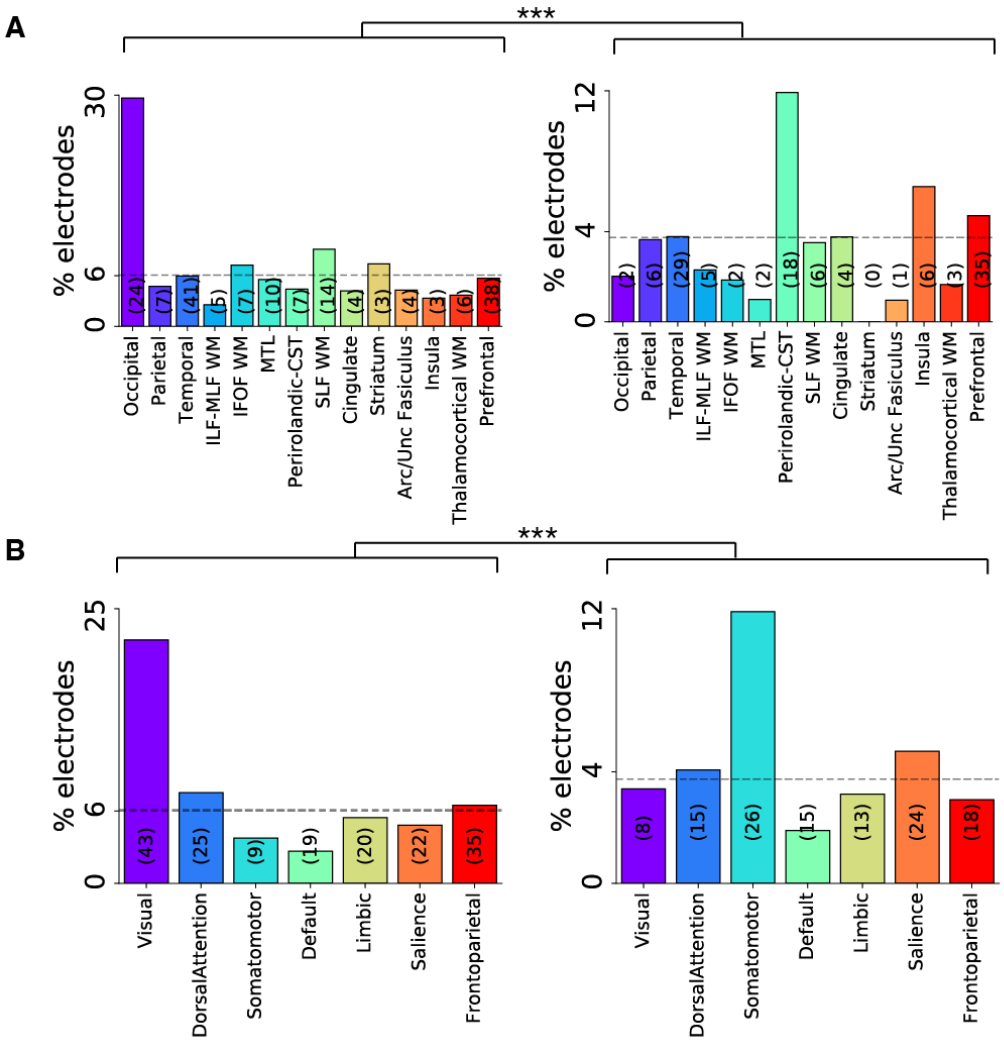

Figure 7- Figure Supplement 1: Comparing anatomical distributions of visuospatial and movement-related electrodes. (A,B) Same as Fig. 6.S1, but for visuospatial and movement-related electrodes (left and right panels, respectively).

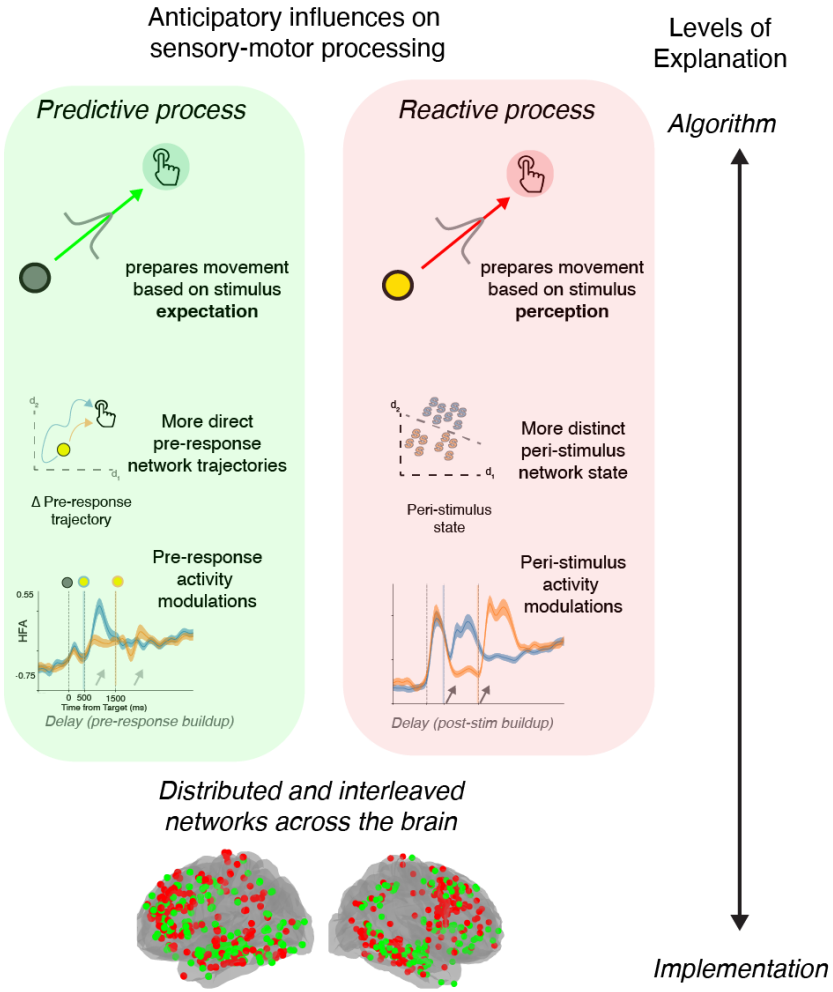

Figure 8: Summary of neural processes related to predictive and anticipatory influences on sensory motor behavior. From top to bottom, we illustrate algorithm- to implementation-level explanations of predictive (green) and reactive (red) anticipatory influences on sensory-motor behavior. In the colored boxes, we highlight key results showing distinguishable and identifiable representations of predictive and reactive processing based on algorithmic behavioral modeling (top row), latent network dynamics (middle row) and localized neural representations (bottom row). Bottom panel shows the distributed and interleaved distribution of these processes across the brain.

Discussion

We studied neural correlates of anticipatory influences on sensory-motor behavior in the human brain using a combination of computational behavioral modeling of RTs and high-resolution intracranial neural recordings from patients with medically refractory epilepsy. We found that this relatively simple cognitive function had a complex mapping onto neural activity in the human brain. This mapping included widespread, heterogeneous, and dynamic representations of behaviorally derived model components that described distinct predictive and reactive modulations of sensory-motor behavior. Below, we further discuss these results in the context of previous findings and argue that this combination of approaches, involving high-resolution, large-scale neural measurements along with computational models of behavior, can bridge algorithm- and implementation-level explanations of cognitive behavior.

Identifying predictive and reactive influences on behavior

We modeled these heterogeneous anticipatory influences on behavior using a rise-to-bound framework (Stone, 1960; Ratcliff, 1978; Carpenter and Williams, 1995; Usher and McClelland, 2001; Brown and Heathcote, 2005; Noorani and Carpenter, 2016; Ratcliff et al., 2016). This framework has close ties to statistical decision theory (Wald, 1945; Good, 1979) and can provide parsimonious descriptions of cognitive influences on RT variability (Noorani and Carpenter, 2016). We included two rise-to-bound processes to capture predictive and reactive components of RT variability (including premature false alarms) within the same model and help bridge prior findings. The pre-stimulus process, which accounted for relatively high rates of false alarms in some participants, is a simplified mechanism to capture certain aspects of predictive information processing that allow individuals to generate responses based on expectations formed based on

prior experience (Los et al., 2014; Los & Van Den Heuvel, 2001; Salet et al., 2022), possibly via associative learning mechanisms (Rescorla and Wagner, 1972; Montague et al., 1996; Friston, 2010). We also wanted to account for the shape of RT distributions, which are often not accounted for by such models (Salet et al., 2022). The post-stimulus process, which resulted in faster RTs without a corresponding increase in false alarms, is akin to other reactive models that that emphasize mechanisms that allow rapid fluctuations in temporal expectation (within trial) to alter behavior, for example, via selective attention that prioritizes incoming sensory information) (Summerfield & Egner, 2009). Our behavioral analyses suggest that both components are needed to fully account for within- and across-individual variability in anticipatory behavior.

Using these model fits to guide our neural analyses, we found certain local neural representations that could be mapped directly onto key components of the rise-to-bound model. For example, we identified RT-dependent activity that appeared to rise to a particular level just before each motor response from some electrodes found in premotor regions. This finding, is consistent with other reports of similar representations of rise-to-bound like processes in the primate brain (Gold & Shadlen, 2007; Hanes & Schall, 1996; Heitz & Schall, 2012; O'Connell et al., 2012). However, these signals in our data comprised only a subset of neural activity that was relevant for anticipatory modulation of behavioral RTs. This result suggests that the mapping between model and brain dynamics is not as straightforward as suggested by certain local and aggregate neural signals (Gold & Shadlen, 2007; Hauser et al., 2018; O'Connell et al., 2012).

More generally, we found that task-related activity modulations were encoded by complex, rapidly fluctuating activity patterns across widespread brain regions. These results are consistent previous findings in intracranial EEG studies showing widespread and heterogeneous

task-related representations across the brain (Mesgarani et al., 2014; Paraskevopoulou et al., 2021; Ramayya et al., 2015), illustrating the benefits of the unique combination of high spatio-temporal resolution and broad anatomical sampling of these neural measurements (Parvizi & Kastner, 2018). Additionally, our results expand on the large body of literature that has described task-related neural representations during anticipatory sensory-motor behavior (Nobre & Van Ede, 2018). Specifically, our results suggest that such representations are: 1) not limited to sensory and motor neural populations as might be inferred from certain animal electrophysiology studies (Ghose & Maunsell, 2002; Janssen & Shadlen, 2005); 2) more heterogeneous than might be inferred from scalp EEG studies that aggregate neural activity across large cortical areas (Miniussi et al., 1999; Rohenkohl & Nobre, 2011; Walter et al., 1964); and 3) more dynamic than might be inferred from functional neuroimaging studies with relatively low temporal resolution (Coull & Nobre, 1998; A Vallesi, 2010; Antonino Vallesi et al., 2009).

These results highlight a major challenge in linking anticipatory influences on sensory-motor behavior with neural activity: a mismatch in complexity between low-dimensional algorithmic explanations of sensory-motor behavior and high-dimensional networks of the human brain. Whereas algorithmic models often represent sensory-triggered motor processes using a single latent variable (e.g., the rising quantity in rise-to-bound models), sensory-motor processing in the human brain involves far more complex patterns of neural activity. For example, functional neuroimaging studies have shown that the cortex displays widespread activity in response to simple sensory stimuli (Gonzalez-Castillo et al., 2012), which may reflect the many recurrent feedback connections from higher order association regions to lower order sensory and motor regions (Felleman and Van Essen, 1991). Such results are consistent with the

emerging view that although the cortex can be segmented into distinct regions based on structural features (Desikan et al., 2006; Fischl et al., 2004), information processing is largely distributed across parallel, interleaved processing streams ("intrinsic brain networks"), with localized information processing limited to certain primary sensory and motor regions (Glasser et al., 2016; Van Essen et al., 1992; Yeo et al., 2011). Consistent with this view, we found that neural representations that showed activity resembling low-level sensory-motor processing were relatively localized to visual and motor regions. These organization principles suggest that network-level processes are not reserved for higher-order cognitive functions that are carried out in isolation from low-level sensory and motor functions. Instead, brain networks function to integrate cognitive computations, such as generating expectations of future events, into ongoing sensory and motor processing (Friston, 2010; Goel & Buonomano, 2014; Nobre & Van Ede, 2018).

Interpreting network dynamics that reflect anticipatory influences on behavior

The measures of network dynamics we used (Ebitz & Hayden, 2021; Vyas et al., 2020) to identify correlates of anticipatory influences on behavior are related to techniques used to study dynamic systems that have gained popularity in neuroscience because of their rich representational potential (Barack & Krakauer, 2021). Such measures of latent neural dynamics quantify how widespread neural activity patterns change over time and are influenced by correlated neural variability across measured neural elements (Churchland et al., 2010; Umakantha et al., 2021) and may contain more information than certain population-level analyses that only consider activity levels across measured neural elements (Kriegeskorte et al., 2008). Whereas latent neural dynamics are often studied across neural populations within a

single region, we measured these dynamics across a much broader spatial scale (i.e., all sampled neural populations for a given subject) and thus refer to them as network dynamics (alternatively described as “mesoscale” when applied to human iEEG recordings; Natraj et al., 2022).

One of our key sets of findings was that, despite the lack of a straightforward mapping from the relatively simple, behaviorally derived descriptions of anticipatory effects on behavior to the complex network dynamics that encoded those effects, the behaviorally relevant network dynamics nevertheless had distinguishable anatomical and functional characteristics (Fig. 8 for a summary of key results). To identify these characteristics, we grouped electrodes based on similar task-related activity modulations using data-driven clustering. We found that two distinct anticipatory network dynamics, one related to predictive processing and the other to reactive processing, were based on distinct subpopulations of neurons that had different task-related firing patterns (as distinguished by the clustering algorithm) and different patterns of anatomical localization across the brain.

Certain properties of these predictive and reactive network dynamics were loosely consistent with predictions from algorithmic models and prior theoretical and empirical studies. Predictive network dynamics were influenced strongly by neural populations with attenuated responses during settings of relatively increased temporal expectation, consistent with a representation related to prediction errors (Montague et al., 1996; Nobre & Van Ede, 2018; Rescorla & Wagner, 1972). These signals tended to have relatively high prevalence in the medial temporal lobe, where memory-related functions and representations of time (Eichenbaum, 2014; Howard et al., 2015; MacDonald et al., 2011; Shankar & Howard, 2012) have been suggested to contribute to predictive information processing (Salet et al., 2022; Stachenfeld et al., 2017). In

contrast, reactive network dynamics were influenced strongly by neural populations that showed task-related activity modulations near the time of stimulus presentation, consistent with a role in selective attention (Summerfield & Egner, 2009). These signals tended to have a relatively high prevalence in preparatory motor regions, which have been implicated in reactive information processing (Gold & Shadlen, 2007; Hanes & Schall, 1996; Hauser et al., 2018).

Finally, we found that individual differences in these predictive and reactive network dynamics could account for differences in behavior. It is possible that some of these results might be particular to our patient population. In general, epilepsy patients can show additional forms of inter-individual variability in their brain networks related to their pathology (Bettus et al., 2008) that might also manifest in certain categorical behavioral differences compared to healthy controls (Bruhn & Parsons, 1977). However, we sought to mitigate such concerns by focusing on neural signals associated with specific, highly controlled behaviors. Under these kinds of conditions, it has been shown that neural findings from intracranial EEG studies in patients with epilepsy can generalize to healthy controls populations (Long et al., 2014). Nevertheless, more work is needed to fully understand the relationship between behavioral and neural variability across individuals (Genon et al., 2022), which can have broad evolutionary (Bechara et al., 2000), developmental (Tenenbaum et al., 2011), and functional (Yang & Wang, 2020) causes.

Conclusions

Taken together, our results suggest that anticipatory influences on sensory-motor behavior, which can be described using relatively simple algorithmic models, emerge from a multitude of heterogeneous, distributed, and elementary computations represented throughout the

brain (Fig. 8). These results reflect the fundamental role of anticipation in higher brain function and can help explain why simple sensory-motor processing engages widespread brain networks (Gonzalez-Castillo et al., 2012; Mesulam, 1998). Moreover, these results motivate the need for more high-resolution and large-scale neural recordings combined with sophisticated computational models of brain activity and behavior to help bridge gaps in our understanding of the computational underpinnings of behavior and their complex neural substrates (Bogacz & Gurney, 2007; Wong & Wang, 2006; Yang & Wang, 2020).

Materials and Methods

Subjects

We studied 23 patients with medically refractory epilepsy who underwent surgical implantation of intracranial electrodes for seizure localization (Table 1). Patients provided informed consent to perform cognitive testing as part of our research study while they were admitted to the hospital. Our study was approved by the University of Pennsylvania Institutional Review Board. Clinical circumstances alone determined the number and placement of implanted electrodes.

Stimulus detection task

Subjects performed a stimulus-detection task with a variable foreperiod delay (Luce, 1986; Nickerson, 1965; Niemi & Näätänen, 1981; Ollman & Billington, 1972). Subjects viewed visual stimuli on a laptop computer screen and responded by pressing a button on a game controller with their right thumb. Each trial began with the presentation of a small white box at a randomized location on the screen as a fixation target (one of nine locations on a 3 x 3 grid). The stimulus changed color to yellow after one of two randomly selected foreperiod delays: 1) short delay=500 ms, or 2) long delay=1500 ms. We measured response time (RT) as the time between color change (stimulus) and button press (response). If a response was provided within a predefined response interval (1000 ms after color change), visual feedback was provided as follows: 1) for RTs ≤ 150 ms, salient positive feedback was shown, consisting of a smiley face with text below (“wow!”); 2) for RTs > 150 ms, the RT was shown, color-coded by binned values (green for 150–300 ms, yellow for 300–600 ms, red for 600–1000 ms). If a response

occurred before the color change of the stimulus ("premature false alarm") or if no response was provided within 1000 ms after color change, the subject was shown a blank screen for 2500–2750 ms (random inter-trial jitter). During each session, the subject performed blocks of 9 trials each. After six blocks, the subject was provided with an option of performing additional blocks of trials. On average, subjects performed a mean of 149 trials (i.e., 1.34 sessions). Two subjects performed additional trials with a 1000 ms foreperiod delay, which we excluded from our analyses.

Extended LATER model of RT

All model-fitting procedures described below were implemented by finding the minimum of the appropriate objective function using the SciPy Optimization package (*Python*). We modeled RT distributions using an extended version the LATER model (Carpenter & Williams, 1995; Noorani & Carpenter, 2016). This rise-to-bound model describes RTs as measuring the time it takes for a latent variable to rise linearly from a starting value (S_0) at stimulus presentation to a threshold value (S_t) at response. The model accounts for trial-to-trial RT variability by assuming that the rate-of-rise of this process varies stochastically from trial-to-trial per a Gaussian process. The basic form of the LATER model describes an RT distribution using two free parameters, as follows:

$$RT \sim \frac{1}{N\left(\mu, \sigma^2\right)}$$

Where N represents a Gaussian distribution with mean μ and standard deviation σ. Note that this version assumes that the threshold value S_t=1; an equivalent formulation assumes that σ=1, and S_t is a free parameter. These parameters are estimated by fitting a Gaussian probability density

function to reciprocal RTs, which typically resembles a Gaussian distribution (Noorani & Carpenter, 2016).

We extended this basic model to account for premature false alarms, which we modeled as arising from a parallel LATER unit that was triggered at target onset (Figs. 1C, and 1.S1). For these and other RT analyses, we assumed that false alarms included responses made either before stimulus onset or within 250 ms after stimulus onset ("fast responses"), because these fast responses are thought to arise from processes distinct from those driving slower RTs (we chose 250 ms as a conservative cutoff to ensure that our neural analyses do not conflate these two processes; Luce, 1986; Noorani & Carpenter, 2016). Specifically, each subject's RT distribution for a given foreperiod delay condition was modeled as a weighted average of: 1) the predictive LATER unit fit to premature false alarms that occurred between 500 ms prior to and 250 ms following stimulus onset, and 2) the reactive LATER unit fit to RTs between 250–1000 ms. Because premature false alarms were associated with negative RTs when measured relative to stimulus onset, we introduced a 500 ms offset such that all RTs were positive values. Thus, this extended LATER model included two LATER units (each with a mean μ and standard deviation σ) and a weighting factor ("premature response bias") that determined the overall probability of triggering an anticipatory false alarm.

We first fit this extended LATER model to behavior on the short-delay trials. We then fit a modified model to behavior on long-delay trials that included two additional free parameters representing delay-related changes in: 1) the weighting factor determining the probability of triggering predictive responses ("Δ predictive bias"), and 2) the variance (σ^2) associated with reactive LATER unit ("Δ reactive bias"). In Table 3, we report the best-fitting parameters for

short- and long-delay trials for each subject (Δ predictive bias and Δ reactive bias were derived by comparing best-fitting premature response bias and rise-to-bound σ parameters on short- and long-delay trials, respectively). We evaluated the overall model fits by computing an R^2 value comparing the (*z*-scored) model-predicted probability distributions and an empirically estimated probability distribution based on a Gaussian-smoothed reciprocal RT histograms (including all responses from 500 ms prior to stimulus onset to 1000 ms following stimulus onset, with an offset such that all RTs on anticipatory false alarms were positive values and RTs on short and long-delay trials were aligned relative to stimulus onset; Fig. 1.S1; Tables 2,3).

Intracranial neural recordings via intraparenchymal electrodes

Patients were implanted exclusively with intraparencymal depth electrodes ("stereo EEG," Ad-tech, 1.1 diameter, 4 contacts spaced 5 mm apart), except in one patient who also had subdural grid electrodes (Subject #142, Ad-tech, 4 mm contacts, spaced 10 mm apart). iEEG was recorded using a Natus recording system. Based on the amplifier and the discretion of the clinical team, signals were sampled at either 512 or 1024 Hz. Signals were converted to a bipolar montage by taking the difference of signals between each pair of immediately adjacent contacts on the same electrode. The resulting bipolar signals were treated as new virtual electrodes (henceforth, "electrodes"), originating from the midpoint between each electrode pair (Burke et al., 2013; Ramayya et al., 2015, 2021). Digital pulses synchronized the electrophysiological recordings with task events. We excluded electrodes that recorded prominent 60 Hz electrical line noise, defined as electrodes that showed greater spectral power in the 58–62 Hz range compared to the 18–22 Hz range, or electrodes that were disconnected (standard deviation=0).

We excluded trials with prominent noise artifacts (e.g., if voltage data were not recorded due to saturation, or if the mean or standard deviation of voltage was >10 standard deviations of all trials). We did not specifically exclude electrodes based on epileptic activity because our analyses focused on behaviorally linked neural activity, which should not be influenced systematically by epileptic networks (Liu & Parvizi, 2019). We analyzed data from 2,609 electrodes.

Anatomical localization of electrodes

Intracranial electrodes were identified manually on each post-operative CT scan and contact coordinates recorded using custom software based on the center of density of the radiodense contacts on thresholded images. To obtain contact locations in each patient's native anatomic space as well as a common reference space (MNI coordinates), we used Advanced Normalization Tools (Avants et al., 2011) to register post-operative CT to pre-operative MRI and each MRI to the Montreal Neurological Institute (MNI) average brain and the Neuromorphometrics whole brain atlas (Landman & Warfield, 2012). Electrodes were assigned anatomical labels in a hierarchical manner. First, electrodes were labeled using the Neuromorphometrics atlas. Second, for all electrodes labelled as “Cerebral White Matter,” we assigned a specific white-matter tract label by relating MNI coordinates of each electrode to the XTRACT Human Connectome Project Probabilistic atlas (Warrington et al., 2020). Third, for white-matter electrodes that remained without a specific tract label, we assigned a lobe-based label using the MNI Structural Atlas (e.g., “parietal white matter” or “frontal white matter”) (Mazziotta et al., 2001). For electrodes that were not assigned to an anatomical structure based on the above procedures, we assigned the label associated with the nearest labeled electrode in

MNI coordinate space (using Euclidean distance). We grouped electrodes into regions of interest (ROI) based on manual grouping of anatomical labels (Table 4). We illustrate our method in Fig. 2a using an example electrode in MNI coordinate space with overlying gray and white matter segmentations (Harvard-Oxford Cortical Structural Atlas is shown in place of the Neuromorphometrics, which is not in standard MNI152 space). We also assigned each electrode to various canonical intrinsic brain networks ("7 network model") using a volumetric atlas in MNI coordinates based on (Yeo et al., 2011; Fig. 2.S1). We refer to "ventral attention" network as the "salience" network due to resemblance to behaviorally-defined networks important for emotion (Seeley et al., 2007), as noted by (Yeo et al., 2011) , but otherwise use terminology as reported in the original study.

Extracting high-frequency activity (HFA)

We extracted 5000 ms segments of iEEG data surrounding the following task events: 1) 2000 ms prior to target onset, 2) 5000 ms after target onset, 3) the stimulus color change, and 4) onset of the motor response. For each segment, we extracted spectral power using 5 complex-valued Morlet wavelets (wave number 3, to increase temporal resolution) with logarithmically spaced center frequencies from 70 to 200 Hz. We squared and log-transformed the wavelet convolutions to obtain power estimates at each time sample. We removed 1000 ms buffers at the beginning and end of each segment to avoid contamination from edge artifacts. We averaged these power estimates across the 5 wavelets, resulting in a single power value for each time sample. We convolved each power-time series with a Gaussian kernel (half-width of 75 ms), resulting in a continuous representation of high-frequency activity (HFA) surrounding each task event. We z-scored HFA by the mean and standard deviation of the distribution of HFA values

obtained from randomly selected segments of iEEG data recorded from that session clips (matched to the number of total task events) to not bias values towards any particular task-related event (Burke et al., 2013; Ramayya et al., 2015, 2021). We henceforth refer to *z*-scored HFA as "HFA."

Measuring and preprocessing electrode-specific activation functions

We quantified the activity pattern recorded by each electrode by measuring the average HFA around target onset during, measured separately on short-delay and long-delay trials (subsequently referred to as "activation functions"). We measured target-locked activation functions from 1000 ms prior to and 4000 ms following target onset.

To ensure that all subsequent analyses focused on modulations of the activation functions that were separate from the immediate sensory and motor responses, we preprocessed the data as follows (all curve fits were performed using the *SciPy* Optimization package in *Python* ("cuve_fit"):

First, we estimated the overall trend in HFA that occurred over the course of the trial independent of changes that occurred immediately following stimulus presentation using the target-locked activation function (averaged across both trial types). Specifically, we fit a line to estimate how HFA changed over time using during two short segments of this activation function: 1) 1000 ms to -500 before target onset, and 2) 1000 ms after the long-delay color change occurred (2500 ms to 3000 ms after target onset). We subtracted this trend line from each target-locked activation function to obtain detrended functions.

Second, we estimated the changes in activity that occurred immediately after target onset across both trial types. Specifically, if the detrended target-locked activation function (averaged

across both trial types, from target onset to 500 ms afterwards) contained any activity peaks or troughs that were outside of the electrode's baseline activity range, we fit a Gaussian distribution to model these activity excursions. If both a peak and trough were detected in a given electrode's activation function, we modeled whichever one was larger. We then removed the post-target Gaussian from each target-locked activation function. For long-delay trials, we also fit a Gaussian to the 500–1500 ms interval after target onset and subtracted it from the activation function to remove any expectation-related changes in activity that occurred in response to the target not changing color.

Third, we modeled activity locked to the color change separately for short- and long-delay trials. Specifically, we first fit a Gaussian to a short time segment after the color change (0–500 ms post-stimulus) to capture any changes in activity that occurred immediately after color change. We also fit any residual peaks or troughs over a longer time segment (0–1000 ms post-stimulus). We then subtracted these fit Gaussians from the activation function from each trial.

Fourth, we modeled activity locked to the motor response separately for short- and long-delay trials. Specifically, we fit separate Gaussians to the 500 ms preceding the response, to capture any pre-response ramping activity or any smeared stimulus-locked activity from the preceding color change, and to any residual peaks or troughs during the 500 ms post-response interval, to capture further changes in activity related to either the response or feedback presentation. We then subtracted these fit Gaussians from the activation function from each trial.

In summary, these procedures resulted in residual trial-by-trial HFA measures that were time-locked to four task events: 1) target-locked, short delay; 2) target-locked, long delay; 3) response-locked, short delay; and 4) response-locked, long delay.

Relating neural activity to foreperiod delay

For each electrode, we assessed whether there were categorical differences in the activation function observed during short- and long-delay trials (Fig. 3.S2). We compared the following features of the activation function for long- versus short-delay trials via an independent-sample *t*-test (assuming unequal variance): 1) baseline interval prior to stimulus (500 ms), 2) post-stimulus activity (250 ms), 3) post-stimulus buildup rate (slope of a line fit to HFA trend 250 ms following stimulus), 4) pre-response buildup rate (slope of a line fit to HFA 250 ms prior to response), and 5) pre-response activity (250 ms). We assigned non-parametric *z*-statistics and *p* values (two-tailed) to each of these tests by comparing each *t*-statistic to their respective null distributions generated by randomly re-assigning category labels for each trial (permutation procedure, 1000 iterations).

Relating residual neural activity to stochastic RT variability

We measured stochastic RT variability for each subject as follows. First, we transformed RTs during correct responses (excluding “fast response” RTs < 250 ms) by taking the negative reciprocal ($-1/RT$), which transforms right-tailed RTs into an approximately Gaussian distribution (the negative sign is applied for convenience such that long RTs are still associated with larger values) (Noorani & Carpenter, 2016). Then, we removed delay-related RT variability by z-scoring reciprocal RTs within each delay condition.

We assessed whether stochastic RT variability was related to neural activity at a given electrode after removing stimulus- and response-locked components of the activation function, including ramping components (Fig. 3.S2). We extracted several trial-by-trial neural features locked to trial events: 1) baseline interval prior to target onset (500 ms), 2) baseline interval prior to stimulus (500 ms), 3) post-stimulus activity (250 ms), 4) post-stimulus buildup rate (slope of a line fit to HFA trend 250 ms following stimulus), 5) pre-response buildup rate (slope of a line fit to HFA 250 ms prior to response), and 6) pre-response activity (250 ms).

We performed an omnibus test to assess for any relationship between neural activity and stochastic RT variability using a multi-variate linear regression model. The dependent variable was stochastic RT variability across all trials, and the independent variables were each of the neural features described above. We measured the predictive power of this model using the sum squared error of residuals (SSE). We also assessed specific relationships between each neural feature and stochastic RT variability using Spearman correlations.

We assigned non-parametric *z*-statistics and *p* values for each of these tests by comparing each of these "true" statistics (SSE of multi-variate model, and Spearman's *rho* for each neural features) to null distributions generated for each electrode by misaligning RTs and neural data (using a circular shift procedure, 1000 iterations, to account for any autocorrelation in RTs). For the omnibus test, we assigned a one-tailed *p* value (i.e., if the true SSE was <5% of null SSE values, we assigned a p value of 0.05). For Spearman correlations, we assigned a two-tailed *p* value (i.e., if *rho* was >2.5% of null *rho* values, we assigned a *p* value of 0.05).

<u>Measuring network</u> <u>dynamics between stimulus and response</u>

For each subject, we assessed latent neural dynamics by analyzing simultaneous HFA measurements across all implanted electrodes (excluding electrodes with major noise components, Fig. 4, 4.S1). We measured network state (*n*) at a given time sample as a vector representing neural activity from all electrodes at that time sample (*z*-scored HFA after removing stimulus- and response-locked components). The length of *n* equaled the number of electrodes in the given subject. On each correct trial with RT > 250 ms, we measured *n* at stimulus onset (*s*) and response onset (*r*) and summarized the trajectory of the intervening dynamics using the following measures, that were motivated by rise-to-bound model dynamics (O'Connell et al., 2018; Ratcliff et al., 2016; Figs. 4 and 4.S1): 1) *s*-*r* distance, the Euclidean distance between *s* and *r*; 2) s-r speed, the mean change in distance traveled per time sample; 3) pre-response trajectory, the ratio between *s*-*r* distance and *n*-*r* distance at 200 ms after stimulus onset (values > 1 indicate relatively direct trajectories, whereas values < 1 indicate relatively indirect trajectories); and 4) *s*-*r* variability, the variance in *n* for all time samples between *s* and *r* (measured as distance from ***r***). We used Principal Components Analysis (PCA) to visualize these high-dimensional stimulus-response trajectories in a reduced three-dimensional space.

Fig. 4a illustrates how brain-state vectors were constructed in an example subject (Subject #179). For this subject, brain state varied in a high-dimensional space (83 components needed to explain 95% of the variance using PCA; Fig. 4.S1). Across subjects, the dimensionality of brain state variability increased with increasing number of implanted electrodes (Spearman's *rho*= 0.98, $p<0.001$, Fig. 4.S1), but with a slope of <1 (0.59 dimensions/electrode), suggesting somewhat, but not completely, independent contributions from each electrode.

Relating network dynamics to foreperiod delay

For each subject, we assessed whether network dynamics showed categorical differences between the two foreperiod delay conditions. We measured differences in the mean of each neural-state metric (Fig. 4, 4.S1) observed during long- and short-delay trials and assigned non-parametric *z*-statistics and two-tailed *p* values by comparing these difference scores to their respective null distributions generated by randomly re-assigning foreperiod delay condition labels via permutation procedure (1000 iterations). Additionally, we measured the discriminability between observations of *s* on long- and short-delay trials using multi-variate binary classifiers ("peri-stimulus state"), and also performed the same procedure for observations of *r* ("peri-response state"; Fig. 4.S1). In each case, we trained a logistic regression model to predict the trial type (long vs. short delay) using the activity pattern observed across all electrodes during stimulus presentation ("*s* discriminability") and at the time of response ("*r* discriminability"). We performed L2 regularization and evaluated model performance using the area under the receiver operating characteristic curve (AUC) via 10-fold cross validation. We tuned the regularization strength hyperparameter (controlling model complexity) over 10 log-spaced values (ranging from 0.0001 to 10,000). We assigned non-parametric *z*-statistic and two-tailed *p* values to each classifier by comparing each AUC measure to their respective null distributions generated by randomly re-assigning foreperiod delay condition labels via permutation procedure (1000 iterations).

We assessed whether these various measures of global neural-state dynamics between stimulus and response (Figs. 4, 4.S1) were significantly modulated by delay condition via permutation procedure ($p<0.05$). We assessed how frequently we observed significant delay-

related differences in each of these measures across subjects and assessed whether these rates significantly exceeded the false-positive rate (5%, Fig. 4.S2). All measured neural dynamics showed significant differences between delay conditions (χ^2 test, FDR-corrected ps < 0.001) and were significantly modulated in at least 5 individual subjects (false-positive rate, 5% = 1.15 subjects). These results suggest that several measures of network dynamics were modulated by foreperiod delay. We assessed the behavioral relevance of these modulations for anticipatory modulation of RTs by studying relations with subject-to-subject variability in behavior.

Identifying behaviorally relevant network dynamics

We related individual differences in delay-related changes in behavior, as quantified by best-fitting rise-to-bound model parameters (Δ predictive bias and Δ reactive bias), with delay-related changes in various neural-state dynamics via Spearman correlation. We also performed partial correlation analyses to assess correlations between network dynamics and best-fit model parameters after accounting for delay-related changes in RTs (model-free measures).

Hierarchical clustering of neural populations based on functional properties

We used data-driven unsupervised clustering to group all electrodes based on similar functional properties (Fig. 5.S1). This approach allowed us to study neural representations of specific task-related processing that were distributed across many regions and to avoid arbitrarily aggregating activity from nearby neural populations that might have heterogeneous functions. For each electrode, we defined a multi-variate feature representing the magnitude of (linear) relationships between each of the five features of the activation function listed above (e.g., baseline, post-stimulus buildup rate, etc.) and stochastic RT variability (i.e., z-scored RT computed separately for short- and long-delay trials, thus removing any mean difference in RT

for the two delay conditions) and foreperiod delay (a binary variable corresponding to short- or long-delay trials), using a multinomial distribution indicating the presence and direction of a significant effect (non-parametric $p<0.05$; 0 indicates no relation, 1 indicates relatively increased activity with long RTs or long-delay trials, and -1 indicates relatively increased activity with short RTs or short-delay trials). We used an agglomerative clustering algorithm (*scikit learn*, *Python*) to group all electrodes using Euclidean distance as a measure of pairwise similarity and a linkage function that merges clusters to minimize the variance within all clusters. We identified 20 clusters that were grouped by similar functional profiles using an objective function maximized the number of clusters that were observed in at least 5 subjects and 50 electrodes (Figs. 5.S1-3).

Assessing properties of electrode clusters

For each electrode cluster, our primary statistical test considered whether the anatomical distribution of electrodes in that cluster across brain regions differed from that expected based on overall sampling across the brain χ^2 test (Figs. 5.S1, 5.S3). We also studied several additional functional properties of each electrode cluster that were not considered by the clustering algorithm, including: 1) various activation function features as defined by the curve-fitting algorithm (i.e., the presence, increase, or decrease in post-target activity, pre-response activity during short-delay trials, or post-response activity), 2) spatial selectivity (post-target activity modulation by the spatial location of the target; one-way ANOVA, $p<0.05$), 3) post-error increase or decrease in activity (independent samples *t*-test comparing average activity 1000 ms post-response in error vs. non-error trials, $p<0.05$), 4) post-reward increase or decrease in activity (independent samples *t*-test comparing average activity 1000 ms post-response in RTs < 300 vs.

other trials, $p<0.05$), and 5) positive or negative correlation between post-target activity (500 ms) and timing of false alarms made after 500 ms (Spearman correlation, $p<0.05$).

Relating behaviorally relevant network dynamics to functionally-defined electrode groups

We assessed the relative contributions of various electrode clusters to across-subject correlations between delay-related changes in network dynamics and anticipatory behavior (as quantified by rise-to-bound model fits). Specifically, for each cluster, we repeated the correlation analysis without the respective electrode cluster and computed a relative contribution score for that electrode cluster as the ratio between the original correlation coefficient and the correlation coefficient obtained after removing the cluster. Scores >1 imply that including the cluster strengthened the across-subject correlation ("adds signal"), whereas scores < 1 imply that the cluster weakened the correlation ("adds noise"). We *z*-scored relative contribution scores across clusters (i.e., the relative contribution score of a functionally defined cluster was normalized against other functionally defined clusters) to identify outliers in relative contribution score (*z*-scores > 3). To account for the potential confound of different numbers of electrodes in each cluster, we performed the following procedure (prior to computing the z-score). For each across-subject correlation analysis, we generated a null distributions for relative contribution scores by removing a random sample of electrodes (2000 iterations each, removing 50–800 electrodes, randomly selected on each iteration). We fit a linear model to estimate the relation between number of observations and relative importance score. We corrected for this potentially confounding effect by measuring the relative contribution score relative to the score expected based on the number of electrodes in the cluster.

Identifying visuospatial and movement-related electrodes

We identified "visuospatial" electrodes that encoded the spatial position of the stimulus on the screen (post-target activity modulation by the spatial location of the target; one-way ANOVA, $p<0.05$). We identified "movement-related" electrodes that showed peaks in activity within 100 ms of button press based on Gaussian modeling of activation functions (we excluded electrodes that showed ramping activity without stimulus- or response-related peaks in activity).

Statistical tests

We performed non-parametric statistical tests as described above, when appropriate. Otherwise, we performed *t*-tests to compare continuous distributions and χ^2 tests to compare categorical distributions (counts data). We used False Discovery Rate (FDR) correction for multiple comparisons (Benjamini & Hochberg, 1995) and considered an FDR-corrected *p* value < 0.05 to be statistically significant. We report FDR-corrected *p* values when indicated, otherwise reported *p* values are uncorrected. We performed analyses using Python using both custom code and publicly available packages (e.g., *NumPy* for numerical computing, *SciPy* for statistics and signal processing, *MNE* for spectral analyses, and *statsmodels* for regression modeling).

Data sharing and Code Accessibility:

Behavioral and neural data used for this study will be made publicly available and have been submitted to Dryad. The Python code used to process these data are available online (https://github.com/ashwinramayya/code_RamaEtal_AntReact)

Table 1.

Subject characteristics

Subject ID	Gender	Age at implant	Number of electrodes
133	Female	52	51
136	Female	57	103
140	Female	48	39
142	Male	30	68
143	Female	30	97
145	Male	21	96
146	Male	19	96
150	Male	16	90
152	Female	33	79
153	Female	56	92
154	Female	43	199
157	Male	25	145
160	Female	46	148
165	Female	20	93
168	Male	27	170
171	Male	49	160
178	Female	38	140
179	Female	19	140
181	Female	31	138
182	Female	25	118
184	Male	22	161
187	Male	24	74
191	Female	32	112

Table 2.

Subject Behavior

Subject	n. trials	n. sessions	mean RT on long delay (ms)	std RT on long delay (ms)	mean RT on short delay (ms)	std RT on short delay (ms)	false alarm rate on short delay	false alarm rate on long delay	lapse rate on short delay	lapse rate on long delay
133	248	2	352.32	100.83	465.22	122.59	0.00	0.28	0.03	0.01
136	236	1	425.99	53.87	489.51	95.57	0.06	0.08	0.00	0.03
139	119	1	455.97	103.75	475.06	76.10	0.00	0.07	0.05	0.07
140	177	1	428.92	52.22	422.20	53.46	0.01	0.15	0.00	0.00
142	115	1	425.36	59.68	460.05	87.90	0.04	0.03	0.02	0.03
143	110	1	487.45	92.09	531.96	56.92	0.00	0.02	0.02	0.00
145	167	1	364.29	31.22	396.15	64.73	0.01	0.05	0.01	0.00
146	234	1	375.09	50.71	373.64	46.95	0.01	0.16	0.01	0.00
150	167	1	467.87	76.21	479.84	68.42	0.01	0.05	0.00	0.00
152	164	1	398.42	53.55	389.61	72.92	0.08	0.57	0.01	0.00
153	124	1	529.43	123.42	595.26	141.85	0.00	0.12	0.11	0.05
154	120	1	501.74	90.12	533.73	122.79	0.02	0.10	0.02	0.01
157	116	1	379.88	46.81	444.45	65.56	0.03	0.09	0.00	0.02
160	174	1	523.43	81.61	547.32	103.38	0.00	0.05	0.09	0.01
165	166	1	400.61	95.09	435.12	63.88	0.00	0.05	0.00	0.00
168	130	1	373.69	28.14	422.86	82.53	0.00	0.28	0.00	0.00
171	157	1	366.53	72.75	464.91	62.67	0.02	0.59	0.01	0.00
178	260	2	451.93	91.35	478.19	110.55	0.02	0.20	0.05	0.05
179	403	3	390.27	79.01	423.30	72.34	0.02	0.09	0.01	0.01
181	263	2	451.22	139.21	539.17	152.55	0.02	0.17	0.11	0.05
182	225	2	399.13	43.01	402.90	61.81	0.03	0.05	0.00	0.01
187	290	2	323.76	40.88	354.84	77.84	0.00	0.49	0.00	0.01
191	231	2	372.18	50.44	386.12	63.19	0.00	0.12	0.00	0.00

Table 3.

Model parameters

Subject	Mean target rise-to-bound)	Std. of target rise-to-bound	Mean of short delay rise-to-bound	Std. of short delay rise-to-bound	Premature response bias on short delay trials	Mean long delay rise-to-bound	Std. of long delay rise-to-bound	Premature response bias on long delay trials	R^2
133	0.69	0.09	0.51	0.03	0.00	0.51	0.05	0.60	0.55
136	0.68	0.06	0.50	0.02	0.07	0.50	0.02	0.07	0.68
139	0.97	0.01	0.51	0.02	0.00	0.51	0.02	0.04	0.90
140	0.73	0.10	0.52	0.01	0.05	0.52	0.01	0.05	0.99
142	0.79	0.10	0.51	0.02	0.04	0.51	0.02	0.04	0.86
143	0.57	0.10	0.49	0.01	0.00	0.49	0.03	0.00	0.86
145	0.73	0.00	0.53	0.02	0.01	0.53	0.01	0.00	0.86
146	0.69	0.10	0.53	0.01	0.07	0.53	0.01	0.07	0.99
150	0.70	0.00	0.51	0.02	0.01	0.51	0.02	0.00	0.92
152	0.68	0.09	0.53	0.02	0.28	0.53	0.02	0.32	0.86
153	0.61	0.03	0.48	0.03	0.05	0.48	0.03	0.03	0.72
154	0.67	0.09	0.49	0.03	0.14	0.49	0.02	0.06	0.87
157	0.82	0.10	0.51	0.02	0.03	0.51	0.02	0.05	0.73
160	0.87	0.00	0.49	0.02	0.00	0.49	0.02	0.01	0.86
165	0.62	0.06	0.52	0.02	0.01	0.52	0.03	0.06	0.80
168	0.71	0.09	0.52	0.02	0.06	0.52	0.01	0.37	0.89
171	0.74	0.10	0.51	0.02	0.04	0.51	0.05	0.82	0.91
178	0.66	0.07	0.51	0.03	0.14	0.51	0.02	0.08	0.82
179	0.71	0.10	0.52	0.01	0.05	0.52	0.02	0.12	0.86
181	0.71	0.10	0.49	0.04	0.02	0.49	0.04	0.18	0.67
182	0.67	0.06	0.53	0.02	0.06	0.53	0.01	0.01	0.93
187	0.72	0.10	0.55	0.03	0.04	0.55	0.01	0.48	0.76
191	0.69	0.10	0.53	0.01	0.08	0.53	0.01	0.08	0.94

Table 4.

Anatomical labels used to define Regions of Interest

Occipital'	MOG middle occipital gyrus', 'SOG superior occipital gyrus', 'OFuG occipital fusiform gyrus', 'IOG inferior occipital gyrus','LiG lingual gyrus','Occipital Lobe_wm','Optic Radiation L', 'Optic Radiation R','Forceps Major', 'Vertical Occipital Fasciculus L', 'Vertical Occipital Fasciculus R'
'Parietal'	AnG angular gyrus', 'PCu precuneus', 'SPL superior parietal lobule', 'SMG supramarginal gyrus', 'PO parietal operculum', 'Parietal Lobe_wm'
'Temporal'	FuG fusiform gyrus', 'ITG inferior temporal gyrus', 'PT planum temporale', 'MTG middle temporal gyrus', 'FuG fusiform gyrus/"BA36', 'PHG parahippocampal gyrus', 'STG superior temporal gyrus', 'TTG transverse temporal gyrus', 'PP planum polare', 'TMP temporal pole', 'Temporal Lobe_wm', 'Acoustic Radiation L', 'Acoustic Radiation R'
'ILF-MLF WM'	Inferior Longitudinal Fasciculus L', 'Inferior Longitudinal Fasciculus R', 'Middle Longitudinal Fasciculus L', 'Middle Longitudinal Fasciculus R'
'MTL'	Hippocampus', 'Amygdala', ' entorhinal area', 'Fornix L', 'Fornix R'
'Perirolandic-CST'	'PoG postcentral gyrus', 'CO central operculum', 'PrG precentral gyrus','Corticospinal Tract L', 'Corticospinal Tract R'
'IFOF WM'	Inferior Longitudinal Fasciculus L', 'Inferior Longitudinal Fasciculus R', 'Middle Longitudinal Fasciculus L', 'Middle Longitudinal Fasciculus R'
'SLF WM'	Superior Longitudinal Fasciculus 1 L', 'Superior Longitudinal Fasciculus 1 R', 'Superior Longitudinal Fasciculus 2 L', 'Superior Longitudinal Fasciculus 2 R', 'Superior Longitudinal Fasciculus 3 L', 'Superior Longitudinal Fasciculus 3 R'
'Cingulate'	ACgG anterior cingulate gyrus', 'PCgG posterior cingulate gyrus', 'MCgG middle cingulate gyrus', 'Cingulum subsection: Dorsal L', 'Cingulum subsection: Dorsal R', 'Cingulum subsection: Peri-genual L', 'Cingulum subsection: Peri-genual R', 'Cingulum subsection: Temporal L', 'Cingulum subsection: Temporal R'
'Striatum'	Putamen', 'Caudate_wm', 'Caudate', 'Putamen_wm'
'Arc/Unc Fasiculus'	Arcuate Fasciculus L', 'Arcuate Fasciculus R', 'Uncinate Fasciculus L', 'Uncinate Fasciculus R'
'Insula'	PIns posterior insula', 'AIns anterior insula', 'Insula_wm'
'Thalamocortical WM'	Anterior Thalamic Radiation R', 'Anterior Thalamic Radiation L', 'Superior Thalamic Radiation L', 'Superior Thalamic Radiation R', 'Frontal Aslant Tract L', 'Frontal Aslant Tract R'
'Prefrontal'	FO frontal operculum', 'OpIFG opercular part of the inferior frontal gyrus', 'OrIFG orbital part of the inferior frontal gyrus', 'MFG middle frontal gyrus', 'TrIFG triangular part of the inferior frontal gyrus', 'SMC supplementary motor cortex', 'POrG posterior orbital gyrus', 'SFG superior frontal gyrus', 'GRe gyrus rectus', 'MOrG medial orbital gyrus', 'MFC medial frontal cortex', 'MSFG superior frontal gyrus medial segment', 'AOrG anterior orbital gyrus', 'FRP frontal pole', 'Forceps Minor', 'Anterior Commissure', 'Frontal Lobe_wm'

Acknowledgments: We thank Drs. Michael Kahana, Michael Beauchamp, Tahra Eissa, Matthew Nassar, Long Ding, Kareem Zaghloul, and Daniel Yoshor for helpful feedback on the manuscript. We thank Cameron Brandon, Everett Prince, John Bernabei, Jacqueline Boccanfuso, and Joel Stein for technical assistance in data collection.

Funding: National Institutes of Health grant 6T32NS091006 (AGR)

References:

Avants, B. B., Tustison, N. J., Song, G., Cook, P. A., Klein, A., & Gee, J. C. (2011). A reproducible evaluation of ANTs similarity metric performance in brain image registration. *NeuroImage*, *54*(3), 2033–2044. https://doi.org/10.1016/j.neuroimage.2010.09.025

Barack, D. L., & Krakauer, J. W. (2021). Two views on the cognitive brain. *Nature Reviews Neuroscience*, *22*(June). https://doi.org/10.1038/s41583-021-00448-6

Bechara, A., Damasio, H., & Damasio, A. R. (2000). Emotion, Decision Making and the Orbitofrontal Cortex. *Cerebral Cortex*, *10*(3), 295–307. https://doi.org/10.1093/CERCOR/10.3.295

Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, *57*(1), 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Bettus, G., Wendling, F., Guye, M., Valton, L., Régis, J., Chauvel, P., & Bartolomei, F. (2008). Enhanced EEG functional connectivity in mesial temporal lobe epilepsy. *Epilepsy Research*, *81*(1), 58–68. https://doi.org/10.1016/j.eplepsyres.2008.04.020

Bogacz, R., & Gurney, K. (2007). The basal ganglia and cortex implement optimal decision making between alternative actions. *Neural Computation*, *19*(2), 442–477. https://doi.org/10.1162/neco.2007.19.2.442

Bruhn, P., & Parsons, O. A. (1977). Reaction Time Variability in Epileptic and Brain-Damaged Patients. *Cortex*, *13*(4), 373–384. https://doi.org/10.1016/S0010-9452(77)80018-X

Burke, J. F., Zaghloul, K. A., Jacobs, J., Williams, R. B., Sperling, M. R., Sharan, A. D., & Kahana, M. J. (2013). Synchronous and asynchronous theta and gamma activity during episodic memory formation. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *33*(1), 292–304. https://doi.org/10.1523/JNEUROSCI.2057-12.2013

Callaway, E., & Yeager, C. L. (1960). Relationship between reaction time and electroencephalographic alpha phase. *Science*, *132*(3441), 1765–1766. https://doi.org/10.1126/science.132.3441.1765

Carpenter, R. H. S., & Williams, M. L. L. (1995). Neural computation of log likelihood in control of saccadic eye movements. *Nature*, *377*(6544), 59–62. https://doi.org/10.1038/377059a0

Cattell, J. M. (1886). The time taken up by cerebral operations. *Mind*, *11*(42), 220–242.

Churchland, M., Byron, M., Cunningham, J., Sugrue, L., Cohen, M., Corrado, G., Newsome, W., Clark, A., Hosseini, P., Scott, B., & Bradley, D. (2010). Stimulus onset quenches neural variability: a widespread cortical phenomenon. *Nature Neuroscience*, *13*(3). https://doi.org/10.3389/conf.neuro.06.2009.03.295

Coull, J. T., & Nobre, A. C. (1998). Where and when to pay attention: The neural systems for directing attention to spatial locations and to time intervals as revealed by both PET and fMRI. *Journal of Neuroscience*, *18*(18), 7426–7435. https://doi.org/10.1523/jneurosci.18-18-07426.1998

Desikan, R. S., Ségonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., Buckner, R. L., Dale, A. M., Maguire, R. P., Hyman, B. T., Albert, M. S., & Killiany, R. J. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *NeuroImage*, *31*(3), 968–980. https://doi.org/10.1016/j.neuroimage.2006.01.021

Dubey, A., & Ray, S. (2019). Cortical electrocorticogram (Ecog) is a local signal. *Journal of Neuroscience*, *39*(22), 4299–4311. https://doi.org/10.1523/JNEUROSCI.2917-18.2019

Dustman, R. E., & Beck, E. C. (1965). Phase of alpha brain waves, reaction time and visually evoked potentials. *Electroencephalography and Clinical Neurophysiology*, *18*(5), 433–440. https://doi.org/10.1016/0013-4694(65)90123-9

Ebitz, R. B., & Hayden, B. Y. (2021). The population doctrine in cognitive neuroscience. In *Neuron* (Vol. 109, Issue 19, pp. 3055–3068). Cell Press. https://doi.org/10.1016/j.neuron.2021.07.011

Eichenbaum, H. (2014). Time cells in the hippocampus: A new dimension for mapping memories. In *Nature Reviews Neuroscience* (Vol. 15, Issue 11, pp. 732–744). https://doi.org/10.1038/nrn3827

Fischl, B., Van Der Kouwe, A., Destrieux, C., Halgren, E., Ségonne, F., Salat, D. H., Busa, E., Seidman, L. J., Goldstein, J., Kennedy, D., Caviness, V., Makris, N., Rosen, B., & Dale, A. M. (2004). Automatically Parcellating the Human Cerebral Cortex. *Cerebral Cortex*, *14*(1), 11–22. https://doi.org/10.1093/cercor/bhg087

Fox, M. D., Snyder, A. Z., Vincent, J. L., & Raichle, M. E. (2007). Intrinsic Fluctuations within Cortical Systems Account for Intertrial Variability in Human Behavior. *Neuron*, *56*(1), 171–184. https://doi.org/10.1016/j.neuron.2007.08.023

Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*, *11*(2), 127–138. https://doi.org/10.1038/nrn2787

Genon, S., Eickhoff, S. B., & Kharabian, S. (2022). Linking interindividual variability in brain structure to behaviour. *Nature Reviews Neuroscience*. https://doi.org/10.1038/s41583-022-00584-7

Ghose, G. M., & Maunsell, J. H. R. (2002). Attentional modulation in visual cortex depends on task timing. *Nature*, *419*(6907), 616–620. https://doi.org/10.1038/nature01057

Glasser, M. F., Coalson, T. S., Robinson, E. C., Hacker, C. D., Harwell, J., Yacoub, E., Ugurbil, K., Andersson, J., Beckmann, C. F., Jenkinson, M., Smith, S. M., & Van Essen, D. C. (2016). A multi-modal parcellation of human cerebral cortex. *Nature*, *536*(7615), 171–178. https://doi.org/10.1038/NATURE18933

Goel, A., & Buonomano, D. V. (2014). Timing as an intrinsic property of neural networks: Evidence from in vivo and in vitro experiments. In *Philosophical Transactions of the Royal Society B: Biological Sciences* (Vol. 369, Issue 1637). https://doi.org/10.1098/rstb.2012.0460

Gold, J. I., & Shadlen, M. N. (2007). The neural basis of decision making. *Annual Review of Neuroscience*, *30*(1), 535–574. https://doi.org/10.1146/annurev.neuro.29.051605.113038

Gonzalez-Castillo, J., Saad, Z. S., Handwerker, D. A., Inati, S. J., Brenowitz, N., & Bandettini, P. A. (2012). Whole-brain, time-locked activation with simple tasks revealed using massive averaging and model-free analysis. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(14), 5487–5492. https://doi.org/10.1073/pnas.1121049109

Hanes, D., & Schall, J. (1996). Neural control of voluntary movement initiation. *Science*, *274*(5286), 427--430.

Hauser, C. K., Zhu, D., Stanford, T. R., & Salinas, E. (2018). Motor selection dynamics in FEF explain the reaction time variance of saccades to single targets. *ELife*, *7*. https://doi.org/10.7554/eLife.33456

Heitz, R. P., & Schall, J. D. (2012). Neural Mechanisms of Speed-Accuracy Tradeoff. *Neuron*, *76*(3), 616–628. https://doi.org/10.1016/j.neuron.2012.08.030

Helmholtz, H. Von. (1866). Treatise on Physiological Optics. In *Book* (Issues 3d ed, 1910).

Howard, M. W., Shankar, K. H., Aue, W. R., & Criss, A. H. (2015). A distributed representation of internal time. *Psychological Review*, *122*(1), 24–53. https://doi.org/10.1037/a0037840

Humphries, M. D. (2021). Strong and weak principles of neural dimension reduction. *Neurons, Behavior, Data Analysis, and Theory*. https://doi.org/10.51628/001c.24619

Janssen, P., & Shadlen, M. N. (2005). A representation of the hazard rate of elapsed time in macaque area LIP. *Nature Neuroscience*, *8*(2), 234–241. https://doi.org/10.1038/nn1386

Kriegeskorte, N., Mur, M., & Bandettini, P. (2008). Representational similarity analysis - connecting the branches of systems neuroscience. *Frontiers in Systems Neuroscience*, *2*(NOV). https://doi.org/10.3389/neuro.06.004.2008

Landman, B., & Warfield, S. (2012). MICCAI workshop on multi-atlas labeling. *Medical Image Computing and Computer Assisted Intervention Conference.*

Latimer, K. W., Yates, J. L., Meister, M. L. R., Huk, A. C., & Pillow, J. W. (2015). Single-trial spike trains in parietal cortex reveal discrete steps during decision-making. *Science*, *349*(6244), 184–187. https://doi.org/10.1126/SCIENCE.AAA4056/SUPPL_FILE/LATIMER-SM.PDF

Liu, S., & Parvizi, J. (2019). Cognitive refractory state caused by spontaneous epileptic high-frequency oscillations in the human brain. *Science Translational Medicine*, *11*(514), 1–14. https://doi.org/10.1126/scitranslmed.aax7830

Long, N. M., Burke, J. F., & Kahana, M. J. (2014). Subsequent memory effect in intracranial and scalp EEG. *NeuroImage*, *84*, 488–494. https://doi.org/10.1016/j.neuroimage.2013.08.052

Los, S. A., Kruijne, W., & Meeter, M. (2014). Outlines of a multiple trace theory of temporal preparation. In *Frontiers in Psychology* (Vol. 5, Issue SEP). https://doi.org/10.3389/fpsyg.2014.01058

Los, S. A., & Van Den Heuvel, C. E. (2001). Intentional and unintentional contributions to nonspecific preparation during reaction time foreperiods. *Journal of Experimental Psychology: Human Perception and Performance*, *27*(2), 370–386. https://doi.org/10.1037/0096-1523.27.2.370

Luce, R. D. (1986). Response Times: Their Role in Inferring Elementary Mental Organization. In *Response Times: Their Role in Inferring Elementary Mental Organization*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195070019.001.0001

MacDonald, C. J., Lepage, K. Q., Eden, U. T., & Eichenbaum, H. (2011). Hippocampal “time cells” bridge the gap in memory for discontiguous events. *Neuron*, *71*(4), 737–749. https://doi.org/10.1016/j.neuron.2011.07.012

Manning, J. R., Jacobs, J., Fried, I., & Kahana, M. J. (2009). Broadband Shifts in Local Field Potential Power Spectra Are Correlated with Single-Neuron Spiking in Humans. *Journal of Neuroscience*, *29*(43), 13613–13620. https://doi.org/10.1523/JNEUROSCI.2041-09.2009

Mazziotta, J., Toga, A., Evans, A., Fox, P., Lancaster, J., Zilles, K., Woods, R., Paus, T., Simpson, G., Pike, B., Holmes, C., Collins, L., Thompson, P., MacDonald, D., Iacoboni, M., Schormann, T., Amunts, K., Palomero-Gallagher, N., Geyer, S., … Mazoyer, B. (2001). A probabilistic atlas and reference system for the human brain: International Consortium for Brain Mapping (ICBM). *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *356*(1412), 1293–1322. https://doi.org/10.1098/rstb.2001.0915

Mesgarani, N., Cheung, C., Johnson, K., & Chang, E. F. (2014). Phonetic feature encoding in human superior temporal gyrus. *Science*, *343*(6174), 1006–1010. https://doi.org/10.1126/science.1245994

Mesulam, M. M. (1998). From sensation to cognition. *Brain*, *121*(6), 1013–1052.

Miniussi, C., Wilding, E. L., Coull, J. T., & Nobre, A. C. (1999). Orienting attention in time. Modulation of brain potentials. *Brain*, *122*(8), 1507–1518. https://doi.org/10.1093/brain/122.8.1507

Montague, P. R., Dayan, P., & Sejnowski, T. J. (1996). A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *Journal of Neuroscience*, *16*(5), 1936–1947. https://doi.org/10.1523/jneurosci.16-05-01936.1996

Natraj, N., Silversmith, D. B., Chang, E. F., & Ganguly, K. (2022). Compartmentalized dynamics within a common multi-area mesoscale manifold represent a repertoire of human hand movements. *Neuron*, *110*(1), 154-174.e12. https://doi.org/10.1016/j.neuron.2021.10.002

Nickerson, R. S. (1965). Response time to the second of two successive signals as a function of absolute and relative duration of intersignal interval. *Perceptual and Motor Skills*, *21*(1), 3–10. https://doi.org/10.2466/pms.1965.21.1.3

Niemi, P., & Näätänen, R. (1981). Foreperiod and simple reaction time. *Psychological Bulletin*, *89*(1), 133–162. https://doi.org/10.1037/0033-2909.89.1.133

Nobre, A. C., & Van Ede, F. (2018). Anticipated moments: Temporal structure in attention. In *Nature Reviews Neuroscience* (Vol. 19, Issue 1, pp. 34–48). https://doi.org/10.1038/nrn.2017.141

Noorani, I., & Carpenter, R. H. S. (2016). The LATER model of reaction time and decision. In *Neuroscience and Biobehavioral Reviews* (Vol. 64, pp. 229–251). https://doi.org/10.1016/j.neubiorev.2016.02.018

O'Connell, R. G., Dockree, P. M., & Kelly, S. P. (2012). A supramodal accumulation-to-bound signal that determines perceptual decisions in humans. *Nature Neuroscience*, *15*(12), 1729–1735. https://doi.org/10.1038/nn.3248

O'Connell, R. G., Shadlen, M. N., Wong-Lin, K. F., & Kelly, S. P. (2018). Bridging Neural and Computational Viewpoints on Perceptual Decision-Making. *Trends in Neurosciences*, *41*(11), 838–852. https://doi.org/10.1016/j.tins.2018.06.005

Ollman, R. T., & Billington, M. J. (1972). The Deadline model for simple reaction times. *Cognitive Psychology*, *3*(2), 311–336. https://doi.org/10.1016/0010-0285(72)90010-2

Paraskevopoulou, S. E., Coon, W. G., Brunner, P., Miller, K. J., & Schalk, G. (2021). Within-subject reaction time variability: Role of cortical networks and underlying neurophysiological mechanisms. *NeuroImage*, *237*. https://doi.org/10.1016/j.neuroimage.2021.118127

Parvizi, J., & Kastner, S. (2018). Promises and limitations of human intracranial electroencephalography. *Nature Neuroscience*, *21*(4), 474–483. https://doi.org/10.1038/s41593-018-0108-2

Purcell, B. A., Heitz, R. P., Cohen, J. Y., Schall, J. D., Logan, G. D., & Palmeri, T. J. (2010). Neurally constrained modeling of perceptual decision making. *Psychological Review*, *117*(4), 1113–1143. https://doi.org/10.1037/a0020311

Ramayya, A. G., Pedisich, I., & Kahana, M. J. (2015). Expectation modulates neural representations of valence throughout the human brain. *NeuroImage*, *115*, 214–223. https://doi.org/10.1016/j.neuroimage.2015.04.037

Ramayya, A. G., Yang, A. I., Buch, V. P., Burke, J. F., Richardson, A. G., Brandon, C., Stein, J. M., Davis, K. A., Chen, H. I., Proekt, A., Kelz, M. B., Litt, B., Gold, J. I., & Lucas, T. H. (2021). Theta Synchrony Is Increased near Neural Populations That Are Active When Initiating Instructed Movement. *ENeuro*, *8*(1). https://doi.org/10.1523/ENEURO.0252-20.2020

Rao, R. P. N., & Ballard, D. H. (1999). Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, *2*(1), 79–87. https://doi.org/10.1038/4580

Ratcliff, R., Philiastides, M. G., Sajda, P., & Shiffrin, R. M. (2009). *Quality of evidence for perceptual decision making is indexed by trial-to-trial variability of the EEG*. *106*(16), 6539–6544. https://doi.org/10.1073/pnas.0812589106

Ratcliff, R., Smith, P. L., Brown, S. D., & McKoon, G. (2016). Diffusion Decision Model: Current Issues and History. In *Trends in Cognitive Sciences* (Vol. 20, Issue 4, pp. 260–281). https://doi.org/10.1016/j.tics.2016.01.007

Ray, S., & Maunsell, J. H. R. (2011). Different Origins of Gamma Rhythm and High-Gamma Activity in Macaque Visual Cortex. *PLoS Biology*, *9*(4), e1000610. https://doi.org/10.1371/journal.pbio.1000610

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning and the effectiveness of reinforcement and non-reinforcement. In *Classical Conditioning. 2. Current Research and Theory.* (pp. 64–69). https://ci.nii.ac.jp/naid/10024275851/

Rohenkohl, G., & Nobre, A. C. (2011). Alpha oscillations related to anticipatory attention follow temporal expectations. *Journal of Neuroscience*, *31*(40), 14076–14084. https://doi.org/10.1523/JNEUROSCI.3387-11.2011

Salet, J., Kruijne, W., van Rijn, H., Los, S. A., & Meeter, M. (2022). fMTP: A unifying computational framework of temporal preparation across time scales. In *Psychological Review*.

Seeley, W. W., Menon, V., Schatzberg, A. F., Keller, J., Glover, G. H., Kenna, H., Reiss, A. L., & Greicius, M. D. (2007). Dissociable intrinsic connectivity networks for salience processing and executive control. *Journal of Neuroscience*, *27*(9), 2349–2356. https://doi.org/10.1523/JNEUROSCI.5587-06.2007

Shadlen, M. N., Kiani, R., Newsome, W. T., Gold, J. I., Wolpert, D. M., Zylberberg, A., Ditterich, J., De Lafuente, V., Yang, T., & Roitman, J. (2016). Comment on “Single-trial spike trains in parietal cortex reveal discrete steps during decision-making.” *Science (New York, N.Y.)*, *351*(6280), 1406. https://doi.org/10.1126/SCIENCE.AAD3242

Shankar, K. H., & Howard, M. W. (2012). A scale-invariant internal representation of time. In *Neural Computation* (Vol. 24, Issue 1, pp. 134–193). https://doi.org/10.1162/NECO_a_00212

Stachenfeld, K. L., Botvinick, M. M., & Gershman, S. J. (2017). The hippocampus as a predictive map. *Nature Neuroscience*, *20*(11), 1643–1653. https://doi.org/10.1038/nn.4650

Summerfield, C., & Egner, T. (2009). Expectation (and attention) in visual cognition. *Trends in Cognitive Sciences*. https://www.sciencedirect.com/science/article/pii/S1364661309001442?casa_token=EMs14nsV5yUAAAAA:drBcxAV1eu4AZBYurLNwKW0_VDmrxgm9A88W5GYzHUr4zObJzwyJkeQnETpgeSESLzzT5CHcDg

Tenenbaum, J. B., Kemp, C., Griffiths, T. L., & Goodman, N. D. (2011). How to grow a mind: Statistics, structure, and abstraction. *Science*, *331*(6022), 1279–1285. https://doi.org/10.1126/SCIENCE.1192788/SUPPL_FILE/TENENBAUM.SOM.REV1.PDF

Umakantha, A., Morina, R., Cowley, B. R., Snyder, A. C., Smith, M. A., & Yu, B. M. (2021). Bridging neuronal correlations and dimensionality reduction. *Neuron*, *109*(17), 2740-2754.e12. https://doi.org/10.1016/j.neuron.2021.06.028

Vallesi, A. (2010). Neuro-anatomical substrates of foreperiod effects. In *Attention and Time*. https://books.google.com/books?hl=en&lr=&id=LBoDED__q3gC&oi=fnd&pg=PA303&dq=vallesi+neuroanatomical+substrates&ots=8sKa_kv7Ih&sig=v1E1sodk_uTkVt9O1ELS-mW_v2E

Vallesi, Antonino, McIntosh, A. R., Shallice, T., & Stuss, D. T. (2009). When time shapes behavior: fMRI evidence of brain correlates of temporal monitoring. *Journal of Cognitive Neuroscience*, *21*(6), 1116–1126. https://doi.org/10.1162/jocn.2009.21098

Van Essen, D. C., Anderson, C. H., & Felleman, D. J. (1992). Information processing in the primate visual system: an integrated systems perspective. In *Science* (Vol. 255, Issue 5043, pp. 419–423). https://doi.org/10.1113/jphysiol.2010.201798

Vyas, S., Golub, M. D., Sussillo, D., & Shenoy, K. V. (2020). Computation through Neural Population Dynamics. In *Annual Review of Neuroscience* (Vol. 43, Issue 1, pp. 249–275). https://doi.org/10.1146/annurev-neuro-092619-094115

Walter, W. G., Cooper, R., Aldridge, V. J., McCallum, W. C., & Winter, A. L. (1964). Contingent negative variation : An electric sign of sensori-motor association and expectancy in the human brain. *Nature*, *203*(4943), 380–384. https://doi.org/10.1038/203380a0

Warrington, S., Bryant, K. L., Khrapitchev, A. A., Sallet, J., Charquero-Ballester, M., Douaud, G., Jbabdi, S., Mars, R. B., & Sotiropoulos, S. N. (2020). XTRACT - Standardised protocols for automated tractography in the human and macaque brain. *NeuroImage*, *217*, 116923. https://doi.org/10.1016/j.neuroimage.2020.116923

Wong, K. F., & Wang, X. J. (2006). A recurrent network mechanism of time integration in perceptual decisions. *Journal of Neuroscience*, *26*(4), 1314–1328. https://doi.org/10.1523/JNEUROSCI.3733-05.2006

Yang, G. R., & Wang, X. J. (2020). Artificial Neural Networks for Neuroscientists: A Primer. *Neuron*, *107*(6), 1048–1070. https://doi.org/10.1016/j.neuron.2020.09.005

Yeo, B. T., Krienen, F. M., Sepulcre, J., Sabuncu, M. R., Lashkari, D., Hollinshead, M., Roffman, J. L., Smoller, J. W., Zöllei, L., Polimeni, J. R., Fisch, B., Liu, H., & Buckner, R. L. (2011). The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal of Neurophysiology*, *106*(3), 1125–1165. https://doi.org/10.1152/JN.00338.2011