Distributed evidence accumulation across macaque large-scale neocortical networks during perceptual decision making

Licheng Zou[1,2], Nicola Palomero-Gallagher[3], Douglas Zhou[4], Songting Li[4,*], Jorge F. Mejias[1,*]

[1]Cognitive and Systems Neuroscience Group, Swammerdam Institute of Life Sciences, University of Amsterdam, Netherlands.

[2]Zhiyuan College, Shanghai Jiao Tong University, China.

[3]Institute of Neuroscience and Medicine (INM-1), Research Center Julich, Germany.

[4]School of Mathematical Sciences, MOE-LSC, Institute of Natural Sciences, Shanghai Jiao Tong University, China.

[*]Corresponding authors: Songting Li (songting@sjtu.edu.cn), Jorge F. Mejias (j.f.mejias@uva.nl).

Abstract

Despite the traditional view of parietal cortex as an important region for perceptual decision-making, recent evidence suggests that sensory accumulation occurs simultaneously across many cortical regions. We explored this hypothesis by integrating connectivity, cellular and receptor density datasets and building a large-scale macaque brain model able to integrate conflicting sensory signals and perform a decision-making task. Our results reveal sensory evidence accumulation supported by a distributed network of temporal, parietal and frontal regions, with flexible sequential decision pathways which depend on task difficulty. The model replicates experimental lesioning effects and reveals that the causal irrelevance of parietal areas like LIP for decision performance is explained by compensatory mechanisms within a distributed integration process. The model also reproduces observed temporal gating effects of distractor timing during and after the integration process. Overall, our work hints at perceptual integration during decision-making as a broad distributed phenomenon and provides multiple testable predictions.

Introduction

Decision-making is a fundamental cognitive function in the brain, involving the processing of sensory information towards a categorical decision and a subsequent motion response (Hanks and Summerfield, 2017). Electrophysiological recordings have characterized DM as an evidence accumulation process, where a decision is made when the firing rate of specific neuronal populations, typically in association parietal areas such as LIP, exceeds a threshold (Roitman and Shadlen, 2002; Shadlen and Newsome, 2001). This principle of sensory integration within local (parietal) circuits has been used to understand a vast range of decision-making paradigms, from simple perceptual decision-making tasks (Brunton et al., 2013; Roach et al., 2023; Wang, 2002; Wong and Wang, 2006) to context-dependent choices (Mante et al., 2013; Siegel et al., 2015; Song et al., 2016), multisensory integration (Chandrasekaran, 2017; Meijer et al., 2020; Raposo et al., 2014; Wierda et al., 2023) or even perceptual choices in the context of consciousness or arousal modulation (Beerendonk et al., 2023; Joglekar et al., 2018; Klatzmann et al., 2023; Mashour et al., 2020).

In recent years, however, compelling evidence has emerged against the local (parietal) sensory integration theory. Notably, lesioning and optogenetic studies in rodents have revealed that the inactivation of parietal cortex has little effect on the performance of decision tasks, even though the activity of those areas is highly correlated with decision speed and accuracy (Erlich et al., 2015; Lohuis et al., 2022). Similar experiments in macaques led to similar conclusions, suggesting that sensory areas like MT have a greater influence on achieving good performance than LIP (Katz et al., 2016). It seems implausible, therefore, that sensory integration necessary to guide decisions would occur solely on parietal circuits. Recent studies have explored whether the involvement of multiple brain regions could support perceptual decision-making (Pinto et al., 2022; Siegel et al., 2015; Steinmetz et al., 2019; Wilming et al., 2020), revealing a wide range of differentiated roles across cortical regions (Akrami et al., 2018; Hanks et al., 2015) and emergent interactions across different areas (Siegel et al., 2015). Promising conceptual advances have been facilitated by computational modeling work involving two or three interacting areas (Jaramillo et al., 2019; Murray et al., 2017; Wimmer et al., 2015). However, the lack of a data-constrained, brain-wide framework to compare and evaluate the contributions of individual areas involved in decision-making tasks has hindered further progress.

In this work, we integrated multiple brain-wide data sets and used them to develop a computational model of the large-scale network of the macaque neocortex with the capacity to simulate decision-making tasks. The data integrated into the model included cortex-wide variations in neuron density, autoradiography-derived NMDA and $GABA_A$ receptor densities, and average dendritic spine count per cell (Elston et al., 2011; Elston, 2007; Froudist-Walsh et al., 2023b), which contribute to capture part of the existing regional heterogeneity in cortex. By replicating a classical perceptual decision-making task (random dot motion discrimination), our model showed the involvement of multiple cortical areas in sensory accumulation, in contrast with the traditionally assumed local integration in parietal cortex. We also observed different roles, such as sensory, accumulators and categorizers, emerging across cortical areas. In addition, the model revealed the existence of two distinct dynamical regimes governing the flow of information during sensory integration, with the level of stimulus coherence determining which of them drove the decision process. By simulating lesions in different cortical regions, we reproduced the results of experimental inactivation studies, supporting the hypothesis of the causal irrelevance of area LIP for decision making and proposing a more important role for temporal lobe regions. Lastly, we propose that temporal gating during decision-making, a phenomenon observed in mice so far, may also be present in macaques due to the robustness of the distributed network supporting integration and maintenance of choice memory (Mejias and Wang, 2022).

Results

A large-scale macaque cortex model incorporating data from multiple macroscopic gradients

We combined multiple neuroanatomical data sets with local dynamical models to build a large-scale model of the macaque cortex (Figure 1A; see Materials and Methods for further details). We first considered a local circuit of two stimulus-selective excitatory populations and one non-selective inhibitory population. This local circuit was our template to describe each of the 40 cortical areas considered in our large-scale model, with properties of each individual circuit varying across areas following available area-specific data. This ensures that each cortical area

was unique in its configuration and dynamics, and reflected the high level of neural and circuit heterogeneity found in real brains. The area-specific properties revealed macroscopic gradients across cortex (Wang, 2020), and included variations in the effective strength of synapses inferred from local properties such as the number of dendritic spines per pyramidal cell (Elston, 2007; Mejias and Wang, 2022) and the area-specific density of NMDA and $GABA_A$ receptors per neuron (Froudist-Walsh et al., 2023b, 2023a)(Figure 1B).

Before assembling these areas into a large-scale brain model, we analyzed the organization of the receptor density per neuron data when compared to the position of each area in the anatomical hierarchy (Markov et al., 2013; Nikola T. Markov et al., 2014) and area-specific dendritic spine numbers (Elston, 2007; Mejias and Wang, 2022). A strong positive correlation was found between NMDA receptor density and the cortical hierarchical position inferred from laminar connectivity data (r=0.731) (Figure 1C), and NMDA receptor density had an almost linear relationship with GABA receptor density (Suppl. Fig. S1). As previously shown (Chaudhuri et al., 2015; Mejias and Wang, 2022), dendritic spine count positively correlated with cortical hierarchical (Figure 1D). We identified here some parietal areas (such as LIP, 7m, and 7a) as outliers based on their low pyramidal spine count but high cortical hierarchy value and NMDA receptor density. This suggested that NMDA-mediated synapses in those areas preferentially target inhibitory neurons.

The data-constrained local circuits were then connected between them to form an extended brain network, using tract-tracing connectivity data of the macaque neocortex (Froudist-Walsh et al., 2021; Markov et al., 2013; Nikola T. Markov et al., 2014; N. T. Markov et al., 2014), leading to a brain model constrained at both the local and large-scale levels by neuroanatomical data. Using the layer-specificity and projection-directionality provided by this dataset, and following previous work (Feng et al., 2023; Mejias and Wang, 2022), we implemented a counterstream inhibitory bias in our network, which assumes that feedforward and feedback projections along the cortical hierarchy slightly but preferentially target excitatory and inhibitory neurons, respectively. Background input to cortical areas was varied across areas to mimic the differentiated thalamocortical projections across cortex and to facilitate that all cortical areas displayed a similar spontaneous activity level.

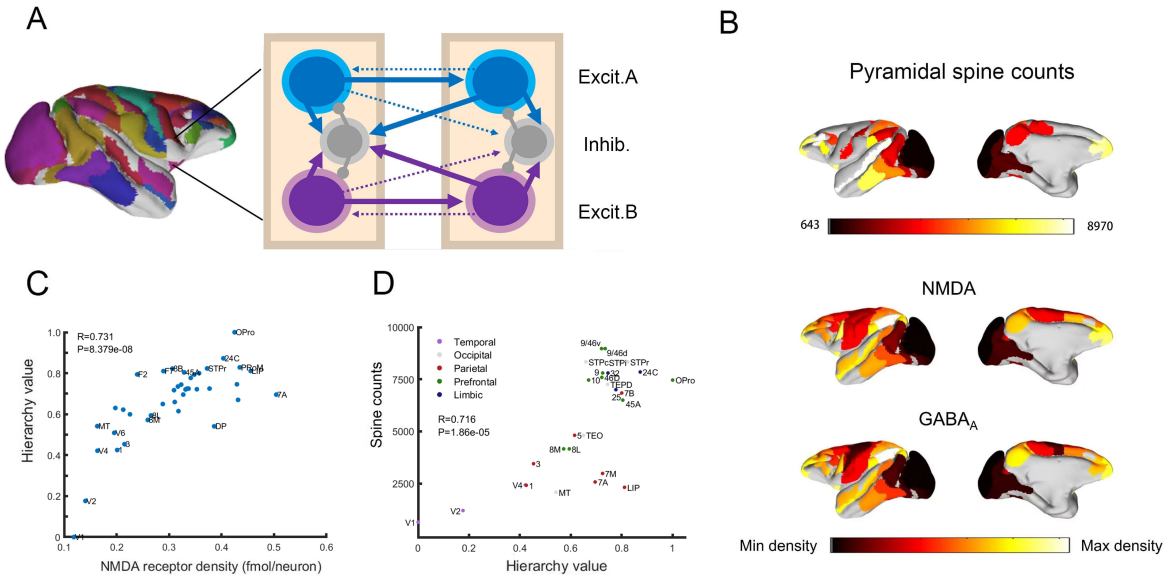

Figure 1: Scheme and anatomical basis of the large-scale macaque brain model. (A) Lateral view of the macaque cortical surface with modelled areas in color. In each area, cortical dynamics follow a local winner-take-all model, and areas are connected using anatomical connectivity data. (B) Area-specific values for the number of dendritic spines per neuron (top), the densities of NMDA receptors per neuron (middle) and $GABA_A$ receptors per neuron (bottom). (C) Correlation between NMDA receptor density and anatomical hierarchy as defined by layer-dependent connections. (D) Correlation between spine count data and anatomical hierarchy as defined by layer-dependent connections.

Distributed and heterogeneous evidence accumulation process in decision making

We simulated our large-scale macaque cortical model performing a classical two-choice random-dot motion discrimination task for decision making (Roitman and Shadlen, 2002; Shadlen and Newsome, 2001), where subjects have to choose which of two possible movement directions has a larger number of dots involved (Figure 2A). In our model, this visual input is modelled as two 700-ms-duration external currents entering both selective excitatory populations in V1 (Figure 2A and B), with dot movement coherence, or motion contrast, reflected in the difference between these currents (Figure 2B). This sensory input drives a competition between both selective excitatory V1 populations, and subsequently between the equivalent populations across other cortical areas.

Simulations of the large-scale brain model (Figure 2C) revealed that the competing visual stimuli induced the accumulation of sensory evidence and the resulting competition and winner-take-all dynamics in area LIP, as shown in experiments and classical models (Wang, 2002; Wong and Wang, 2006). However, such dynamics were also present in other brain areas, such as temporal lobe regions (TEO, TEpd), prefrontal cortex (9/46v), and many others (Suppl. Fig. S2). This suggests that the decision process, including sensory accumulation and winner-take-all dynamics, occurs across a distributed cortical network rather than in localized parietal regions. The dynamics presented important differences across the cortex: areas like LIP, TEO and 7M showed a slow and gradual ramping process, frontal areas like F5 and 9/46v underwent a sharp transition reflecting a categorical choice dynamic, and temporal areas like MT and TEpd

displayed mutual ramping after stimulus onset and then bifurcated to encode the final decision. As in classical models and experimental evidence on single areas, the accumulation process was dependent on the level of stimulus coherence (Figure 2D): as illustrated by this example in LIP, ramping activities bifurcated earlier in high-coherence trials compared to low-coherence ones, and experienced distinct trajectories for two situations. In both cases, the winning population displayed sustained activity after the integration process, due to the bistability of the system, and related to working memory capacity of the network (Mejias and Wang, 2022). When adopting a majority rule to infer the behavioral decision from the winning populations across all cortical areas, the large-scale brain model was also able to replicate psychometric curves in agreement with macaque behavioral output data (Figure 2E), as well as plausible chronometric curves (Figure 2F). The reaction time in chronometric curves was defined as the time at which populations displaying gradual ramping activity (like LIP) reached a threshold of 15 Hz, signaling a joint contribution of association areas towards a decision.

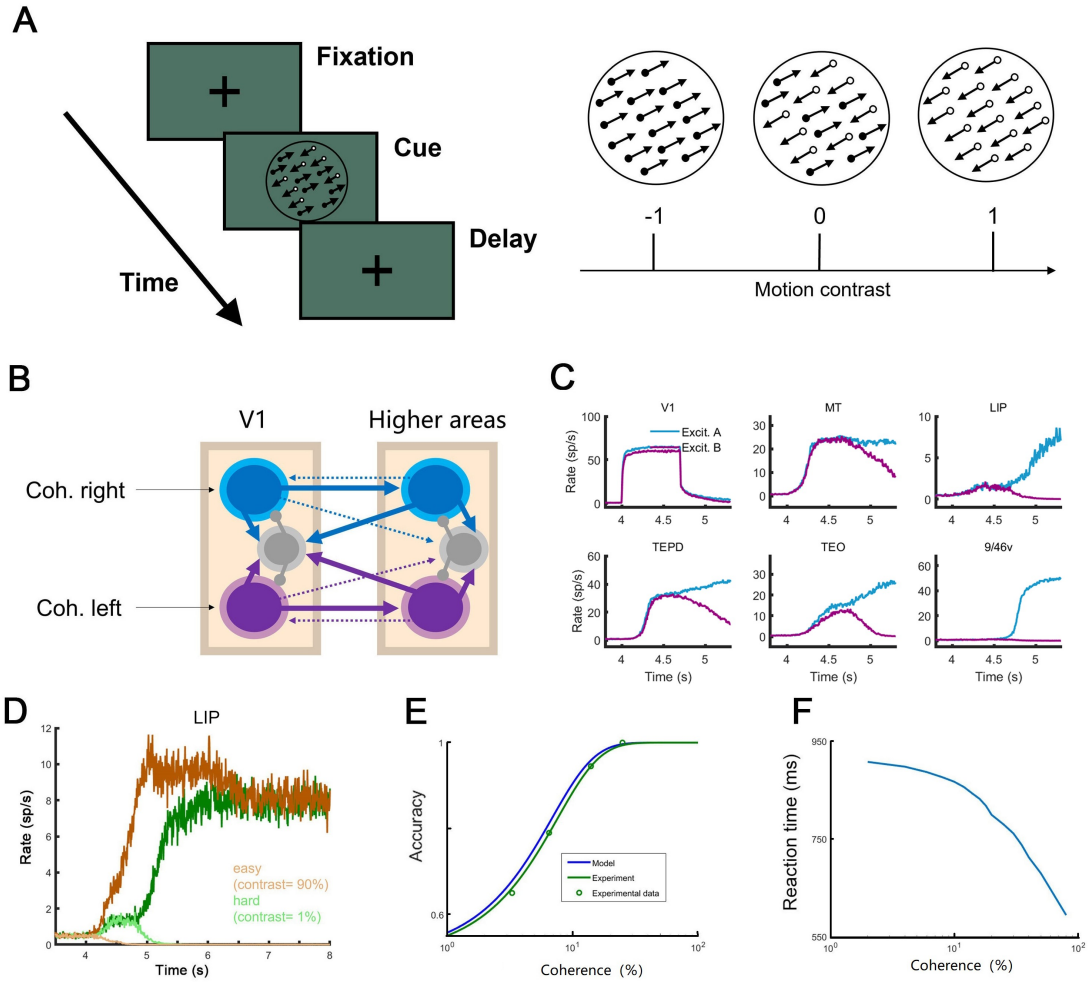

Figure 2: Large-scale brain model reveals a distributed sensory accumulation process for decision making. (A) Random-dot motion discrimination task simulated in our work (left panel), where monkeys must choose which is the preferred direction of movement of the visual stimuli. The stimuli can display any coherence level from 100% right to 100% left movement (right panel). (B) Circuit schematic of the simulation. Two external stimuli representing motion information enter the excitatory populations of V1. (C) Activity of selected cortical areas during the decision-making task, with a 5%-coherent selective visual input duration of 700 ms, displaying a sensory accumulation process distributed across areas. (D) Sensory accumulation activity in LIP, for two different coherence

levels. Green (orange) traces correspond to easy (hard) trials. Populations encoding the correct and incorrect choice are shown in thick and thin lines, respectively. (E) Psychometric and (F) chronometric curves as a function of stimulus coherence. Green dots and curves are experimental data from macaques (Roitman and Shadlen, 2002).

Winning onset times uncover motion information flow across the cortex

While evidence accumulation was observed across many cortical areas, integration times greatly varied across regions. We analyzed the temporal integration across multiple areas by comparing the normalized firing rate traces of the populations encoding the winning choice. For low coherence levels, temporal and prefrontal areas (like TEO and 9/46v, respectively) led the integration and were the first ones to reach their half-maximum level (Fig. 3A, left). However, the situation changed for trials with high coherence levels, as visual areas like V4 and MT were the first ones reaching their half-maximum level (Fig. 3A, right) and temporal and frontal areas evolved at a slower pace. This suggests that the information flow during decision-making is stimulus-dependent, and that the functional relationships between cortical areas (i.e. which area leads the integration), rather than being a fixed structure, flexibly depends on task difficulty.

To estimate the moment in which a winner emerges in the local competition between selective populations in each area, we defined the 'winning onset' in a cortical area as the moment in which the difference between their firing rates becomes larger than a certain threshold (zero, for simplicity) and continues evolving towards the decision threshold without returning back to zero (black dot in Fig. 3B). Intuitively, the winning onset signals the time at which a winner population gains an edge over the other one, and a differentiated integration emerges. The winning onset provides a simple way to evaluate which areas lead the integration process, which aligns reasonably well with decision sequences observed in categorical decision tasks in macaques (Siegel et al., 2015)(see Suppl. Fig. S3 for a direct comparison). We simulated, for a given coherence level, 2000 decision-making trials and computed the winning onset for each area. By sorting these winning onset times and ranking all areas from first/fastest to last/slower, we obtained a trial-averaged distribution of winning-onset ranks for each area (Fig. 3C).

For low coherence values, frontal areas like 9/46v and 45A, as well as temporal areas like TEO, STPc and STPi were usually ranked in the first places. Areas linking sensory and association areas, like MT and LIP, as well as vision-unrelated areas such as areas 2 and 5 (primary and secondary somatosensory cortex), fell behind in the ranking. This suggests that, under low coherence input, higher association areas led the sensory integration and decision process, which was later reflected in the activity of areas like LIP and MT. For high coherence values, the pathway for sensory integration drastically changed (Fig. 3C, right). In this condition, LIP, MT and similar areas connecting sensory and association cortex become fast integrators and reclaim the first places in the winning onset ranking. Association areas, on the other hand, became slower and decreased in the ranking (although some of them, like TEO, remain in moderately high ranks). Overall, high-ranked areas seemed to be hierarchically close to early sensory areas V1 and V2, indicating that sensory integration follows the anatomical hierarchy for high coherence input.

Two distinct mechanisms accounted for decision making in our model, depending on the coherence level. We focused on the rank order distribution of area MT to further illustrate the intrinsic mechanisms. Area MT experienced a remarkable transition from low to high winning-onset ranks when coherence increased from 0% to 100% (Fig. 3D). For low coherence, the differentiated sensory drive received by MT was weak, and due to the low self-coupling strength

of this area, it took some time for its firing rate to reach the winning onset. Association areas, which typically have stronger self-coupling, reached their winning onset points faster in this condition, and hence ranking higher. For higher coherence levels, the differentiated sensory drive was strong enough to trigger a winning onset (and therefore a choice) in MT and related areas, and this information was then passed to higher association areas which followed MT.

The rank switch phenomenon could be further illustrated with a toy model (Fig. 3E) including only V1, MT and 9/46v. Here we rescaled the anatomical connections while keeping their relative ratios, and kept other parameters constant (see Methods). As in Fig. 2C, we observed activity ramping in both MT and 9/46v (Suppl. Fig. S4). Winning onset times decreased with coherence level for both areas (as easier trials lead to faster decisions), and the intersection between both curves signaled a switch in the winning onset ranking.

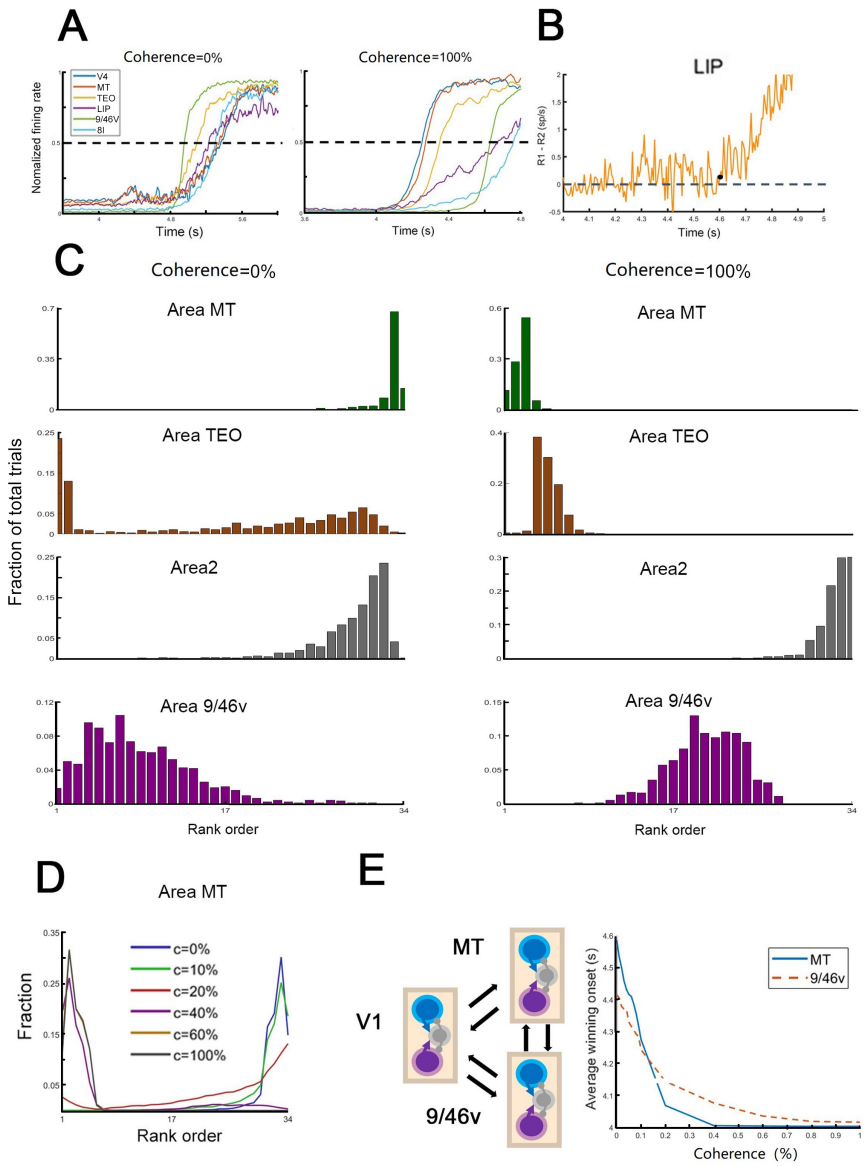

Figure 3: Winning onset analysis reveals information flow during decision making. (A) Normalized firing rate difference between both excitatory populations for selected cortical areas under zero (left) and full coherence (right).

(B) Winning onset was defined as the time from stimulus onset until the difference of firing rate between both excitatory populations rose from zero (black dot), without reverting to zero later in the trial. Dotted line indicates the baseline for excitatory firing rate difference. (C) Winning-onset rank distribution of selected areas under zero (left) and full coherence (right). (D) Winning-onset rank for area MT across six coherence levels. (E) Toy distributed model consisting of three areas (left) and the relation between coherence level and winning onset for MT and 9/46v. All datapoints in (C) and (D) were averaged over 2000 trials.

Inactivation reveals the functional irrelevance of LIP during decision making

Inactivation of specific brain regions, either by lesioning, pharmacology or optogenetics, has provided valuable insight for decision-making circuits. In macaques, inactivation via the $GABA_A$ receptor agonist muscimol revealed a profound impact of MT inactivation in behavioral performance, while LIP inactivation had no measurable impact on performance (Katz et al., 2016). We used our model to simulate inactivation of MT and LIP areas (Fig. 4A), and we compared our modeling results with experimental evidence. We found a strong agreement between our results and the experimental findings of Katz and colleagues (Katz et al., 2016), with LIP inactivation barely affecting the psychometric curve for our model and MT inactivation leading to a flattening of the sigmoidal, reducing the accuracy of the decision process (Fig. 4B). Our results therefore support the evidence of LIP playing a relatively minor role in decision making, despite accurately reflecting the sensory accumulation process occurring across other distributed cortical networks. The effects of inactivation were significantly smaller for higher values of the global synaptic coupling strength (G=0.52), as in this case the inactivation of any area could be easily covered by the joint contribution of other areas (Suppl. Fig. S5).

To better explain this result, we evaluated how the percentage of valid trials, defined as those for which the model was able to make a decision, depended on the coherence level for different global coupling values. For small global coupling, lesioning MT prevents the sensory evidence to reach higher cortical areas, leading to noise-driven trials and poor performance. Larger global coupling values ensure instead that trials are evidence-driven, leading to a good performance for MT inactivation due to compensatory effects of other regions. The inactivation of LIP, however, can be compensated by sensory accumulation in other areas for both small and large global coupling, leading to a causal irrelevance of LIP for the decision process (Fig. 4C).

Our distributed computational model was also able to infer how lesioning one area would affect the dynamics of other areas, especially their ramping speed which is a key feature of the evidence accumulation process. To precisely define ramping speed, we introduced three area-specific metrics besides the winning onset: (i) the winning rate, defined as the firing rate of the winning population at the winning onset, (ii) the reaching rate, defined as 75% of the attractor firing rate minus 25% of the winning rate, and (iii) the reaching onset, defined as the time at which the winning population firing rate arrives at the reaching rate (Figure 4D). Ramping speed was defined as the change in firing rate over the ramping time window, and the lesioning effect of area j on area i was the percentage change between pre- and post-inactivation ramping speed. To better capture the ramping dynamics, we use a winning onset threshold of 0.5 instead of zero.

We then observed that lesioning MT had a stronger effect on other non-sensory areas than lesioning LIP (Figure 4E). Though lesioning LIP has a stronger impact in certain areas like 7B or STPi, the ramping speed of most areas (especially LIP, 7m, 8l and 10) changed more drastically when lesioning MT. Notably, lesioning LIP led to ramping speed increases in areas like F3 and TEpd; the reason was that long-range projections from LIP on those areas targeted inhibitory

populations more strongly than excitatory ones, so removing LIP had a positive effect on those areas. This effect was not observed when lesioning MT, which follows from its early position on the feedforward cortical hierarchy and its role in the excitatory feedforward pathway.

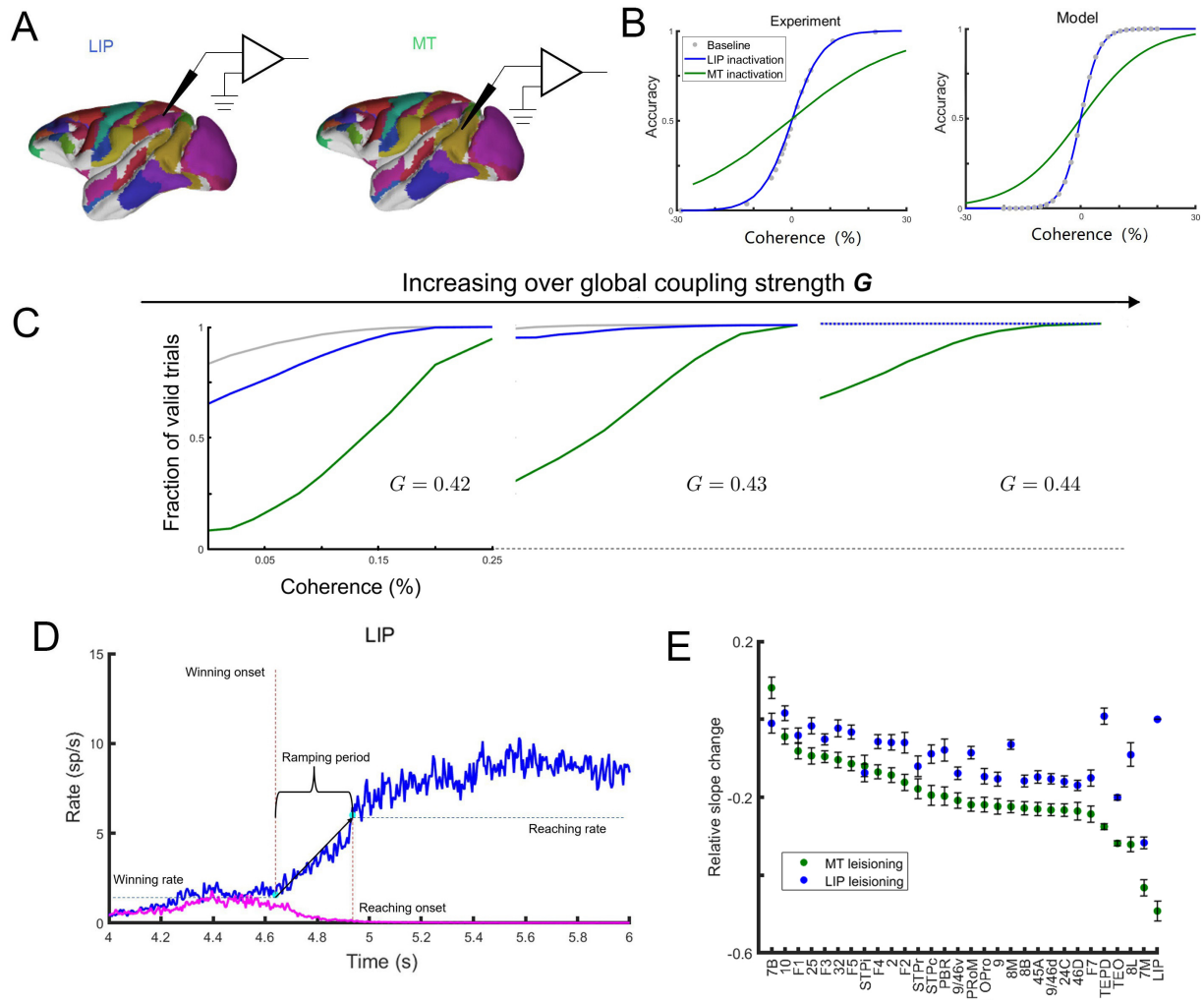

Figure 4: Comparison between MT lesion and LIP lesion in decision making. (A) Schematic of the inactivation protocol. (B) Psychometric curves for baseline, MT lesion and LIP lesion in macaques (left) and in our computational brain model with global coupling strength of G= 0.42 and $\boldsymbol{\sigma_s = 0.018}$ (right). (C) Relationship between coherence and fraction of valid trials, for different global coupling strength values. (D) Illustration of winning rate, reaching onset, reaching rate, and ramping period. (E) Relative slope change (percentage of ramping speed change) across individual areas when lesioning MT or LIP, averaged over 300 trials.

Effects of lesioning areas across the whole cortex

We extended the study of the previous section by lesioning all non-sensory areas one by one (Fig. 5A) and analyzed how any particular inactivation affects ramping speeds across cortex. The full inactivation matrix is displayed in Fig. 5B. We observed, for example, that lesioning

prefrontal area 46d had a strong effect on the ramping speed of parietal areas LIP and 7m, indicating a supportive role of PFC on parietal cortex. Furthermore, lesioning TEO greatly decreased the ramping speed of TEpd and increases that of area 45A and the OPro. Lesioning temporal areas like STPc, STPi and MT affected virtually all association areas.

When we took mean values of the lesioning effect matrix, effectively applying the majority rule, we could infer how lesioning specific cortical areas globally affected the macaque's decision-making performance (Fig. 5C) and the mean ramping speed across areas (Fig. 5D). For example, frontal areas like OPro, PRoM, 32 and 25 had little effect on the psychometric curve (as indicated by changes in its slope), since their interareal connectivity with other regions was relatively weak. This link between structural connectivity and lesioning effects was reflected in a negative correlation of $r = -0.39$ between both factors (Fig. 5E). In contrast, inactivation of temporal areas led to strong effects, suggesting a more salient role of the temporal lobe in decision making.

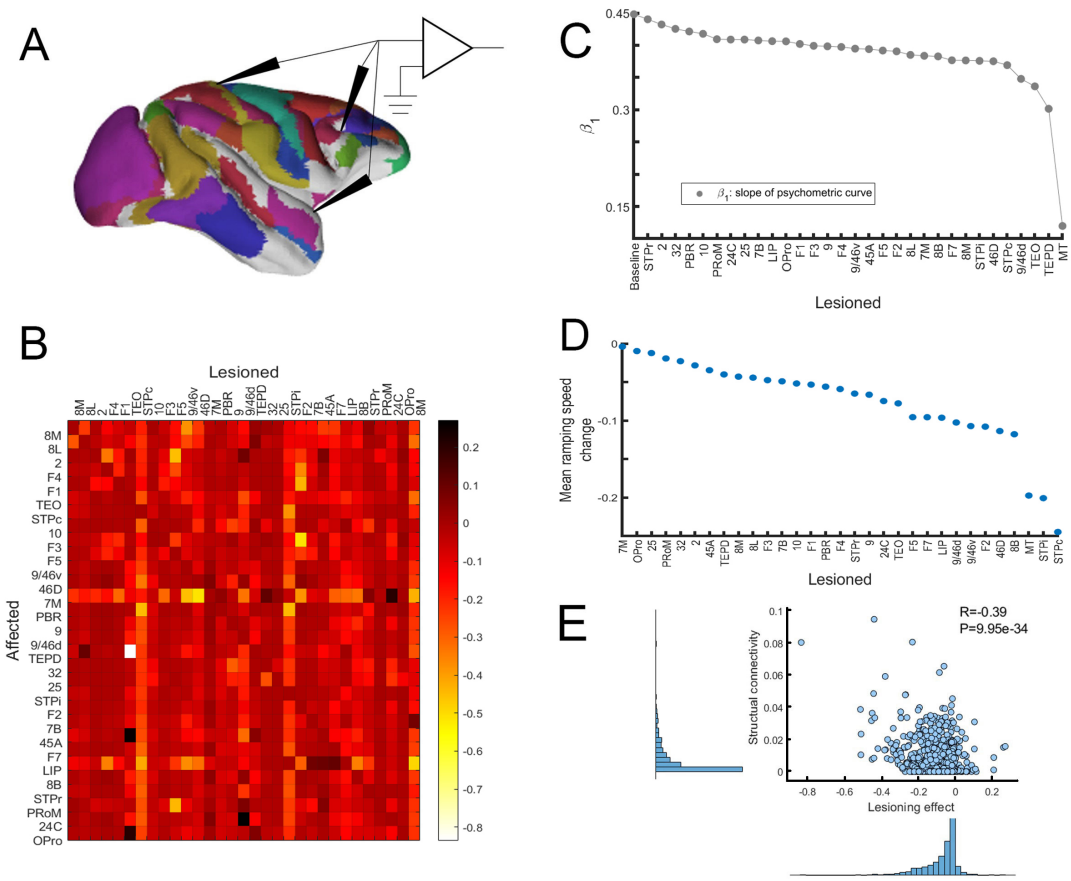

Figure 5: Model prediction of unrecorded cortical areas. (A) Schematic of the inactivation protocol for the neocortex. (B) Slopes of psychometric curve for baseline (0.45) and single area inactivation. (C) Inactivation matrix, with rows indicating the lesioned area, columns the affected areas, and color the relative slope change. (D) Mean of relative ramping speed changes when lesioning individual areas. (E) Relationship between lesioning effect (mean relative slope change) and structural connectivity. The distribution of corresponding variables is also displayed (left and bottom).

Temporal gating in the macaque brain

In our large-scale brain model, information about the decision taken is stored in working memory as distributed sustained activity patterns, as in previous work (Mejias and Wang, 2022) and experimental evidence (Christophel et al., 2017; Leavitt et al., 2017; Voitov and Mrsic-Flogel, 2022). For categorical decision-making tasks, the robustness of such memory representations is important for translating decision outcomes into motor output. This has been experimentally tested in rodents and macaques, by presenting distractors at different time points across the delay period, revealing that the robustness of decision outcomes greatly depends on the distractor timing, a phenomenon known as temporal gating (Finkelstein et al., 2021; Seidemann et al., 1998; Suzuki and Gottlieb, 2013). We tested the robustness of our decision-related working memory in a decision-making task with distractors (Fig. 6A). Two versions of the task were considered: one in which distractors have a fixed onset but variable duration, and one in which they have a fixed duration but variable onset: at the sampling, early delay, or late delay phases (Fig. 6B).

We first studied the system in the fixed duration version of the task, and calculated the fraction of trials in which the decision was changed as a consequence of a distractor at different onsets (Fig. 6C). We observe that early distractors, especially those presented during the sampling phase, have stronger effects on changing the decision output than late ones, in agreement with experimental observations (Finkelstein et al., 2021; Seidemann et al., 1998). The effects were similar for weak and strong distractors, with the latter simply scaling up the fraction of trials with reversed decisions. We observed a positive relationship between the distractor onset and the minimal strength needed for the distractor to be effective, i.e. to change the decision outcome (Fig. 6D, left panel). For the fixed onset version of the task, we found an inverse relationship between the distractor strength and the minimal duration required for the distractor to be effective (Fig. 6D, right panel).

The results were similar for networks with a stronger global coupling (Fig. 6E), although with more abrupt limits in distractor effectiveness. For example, in the fixed duration version of the task (left panel), the decision was nearly impossible to be reversed with distractors of very late onset, irrespectively of the distractor strength. Likewise, decisions could not be reversed with very weak distractors (right panel), irrespectively of the distractor duration. Changes in distractor duration (for the fixed onset version) or distractor strength (for the fixed duration version) were determinant factors for the success of the distractor in overturning the decision (Suppl. Fig. S6).

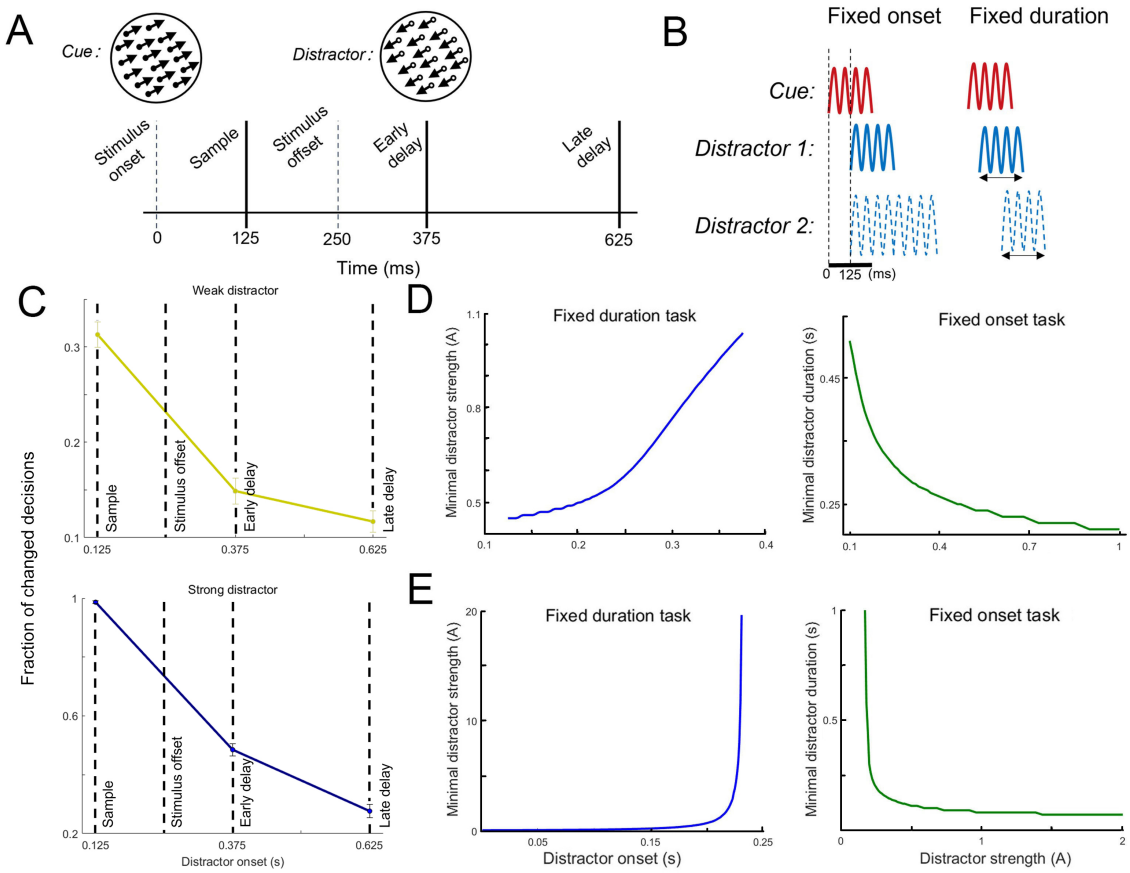

Figure 6: Temporal gating effects during decision making. (A) Decision making task with distractor. The model was presented with a 100%-coherence stimulus of strength 0.1 and duration 250ms, and then a distractor was presented in one of three time points (sample, early delay, or late delay). (B) Distractors in this task can be considered as either fixed onset (here at 125 ms) and variable duration, or as fixed duration (here, 250 ms) and variable onset. (C) Behavioral impact was quantified as the fraction of trials in which the network changed its decision due to the distractor, for the case of a weak (I=0.17, top panel) or strong distractor (I=0.80, bottom panel) and G=0.42. (D) Minimal effective distractors in terms of duration and strength, for the fixed duration (left) and fixed onset (right) versions of the task, global coupling G=0.42 and input noise of $\boldsymbol{\sigma_s = 0.017}$. (E) Same as panel D, but for a network with a stronger global coupling G=0.52 and $\boldsymbol{\sigma_s = 0.01}$.

Discussion

In this work we investigated the distributed computations underlying decision-making tasks, by simulation of large-scale cerebral network models. Traditionally, decision making has been characterized by a gradual evidence accumulation process reflected in ramp-up activity in local cortical circuits (Roxin and Ledberg, 2008; Wang, 2002; Wong and Wang, 2006; Ye and Li, 2021). While a few computational studies have approached this topic considering two interconnected areas (Jaramillo et al., 2019; Murray et al., 2017; Wimmer et al., 2015), our work provides a formal leap as a novel and detailed computational study of decision making encompassing a large-scale, data-constrained cortical network. Additionally, our study opens the door to explore cognitive functions within large cortical networks in an integrative manner, as its architecture and properties align with a recent proposal regarding the distributed nature of working memory (Mejias and Wang, 2022).

A salient feature of our work is the integration of multiple data sets of gradients of biophysical properties across the cortical network: (i) a pyramidal cell dendritic spine gradient informing on local synaptic density, (ii) a tract-tracing connectome embedding the model with directed and weighted long-range connections, (iii) a neuroanatomical cortical hierarchy used to infer targeted populations in long-range projections, and (iv) gradients of receptor density per neuron for the NMDA and $GABA_A$ receptors which allowed to constrain the excitatory/inhibitory balance and strength. This level of detail allowed us to identify outliers like LIP or 7A, which had a high hierarchical rank and NMDA receptor density per neuron but a low level of pyramidal dendritic spines. This pointed at an unexpected abundance of NMDA receptors on inhibitory neurons, which contributed to the observed slow ramping activity in parietal cortex compared to other areas such as prefrontal cortex. The combination of data sets also allowed our model to explain sustained activity as reported in areas V4 and MT (Bisley et al., 2004; Hayden and Gallant, 2013), which previous models had missed. Combining these data sets with additional refinements, such as causality-reconstructed connectivity maps and hierarchical interactions (Chaudhuri et al., 2015; Li and Wang, 2022; Mejias et al., 2016; Nunes et al., 2021; Zhou et al., 2014) or additional receptor per neuron densities (Froudist-Walsh et al., 2023b, 2021; Hansen et al., 2022; Zilles and Palomero-Gallagher, 2017) may further increase the model's predictive capacity.

Winning onsets uncover multiple information pathways

The sequential order of winning onsets observed in our model was a reflection of how motion information may be processed and broadcasted in the brain. Interestingly, the information pathways revealed by the model were not fixed, and rather dependent on the input coherence and therefore the task difficulty. We observed that higher cognitive and association areas like TEO and 9/46d led the decision process for hard trials, while early sensory areas led for easier trials. The alteration in winning onset order could indicate a change in the role of certain areas when the macaque encountered a tough vs simple task. This suggests that an efficient performance on difficult decision-making tasks might benefit from top-down modulation, a prediction aligned with conscious perception theories (Joglekar et al., 2018; Klatzmann et al., 2023; Mashour et al., 2020; Vugt et al., 2018). There is also evidence of monkeys taking different strategies depending on task difficulty (Brunet and Jagadeesh, 2019; Hyafil and Moreno-Bote, 2017).

In addition, previous experimental attempts to uncover the information flow during categorical perceptual decision tasks (Siegel et al., 2015) revealed a decision-related cortical pathway given by MT→ LIP→ V4→ IT→ FEF→ IPFC. By setting the input to full coherence, therefore approximating a categorical perceptual decision task, our model led to a similar cortical pathway for information flow attending to winning onset order: V4→ MT→ IT(TEO)→ LIP→ FEF (8I)→ IPFC (9/46v) (Suppl. Fig. S3), with only minor misalignments with the data for the positions of V4 and TEO/IT. Furthermore, the activation of this sequence occurs within 0.2 seconds for both experimental data and our model predictions. This indicated that data-constrained large-scale models are well-suited to explore decision-related information pathways in cortical networks, and to uncover the mechanisms behind sequential decision stages (Finkelstein et al., 2021; Guo et al., 2014).

Causal irrelevance of LIP is replicated by simulated inactivation

Despite the traditional focus on parietal cortex as a circuit involved in sensory accumulation, recent evidence on rodents and primates has disputed this idea (Akrami et al., 2018; Erlich et al.,

2015; Hanks et al., 2015; Katz et al., 2016; Pinto et al., 2022). However, a mechanistic explanation as of why parietal cortex would play a relatively minor role in decision making despite its activity's high correlation with behavior was crucially lacking. By simulating lesioning effects across different areas, our model demonstrates that a distributed network of redundant evidence accumulation is sufficient to explain the experimentally observed causal irrelevance of LIP, as its inactivation may be compensated by the integration process occurring in parallel in other areas. Inactivation of other areas such as MT led by contrast to a much more substantial impairment in performance, as also observed experimentally. This is mostly due to the relatively low hierarchical position of MT, acting as an important bridge in the sensory pathway. When further exploring currently unknown effects of lesioning other areas, the model predicted an important role of temporal lobe areas such as TEO and TEpd, a prediction that requires experimental validation.

Temporal gating in decision-related memory

In decision making tasks, monkeys are sometimes expected to maintain their choice in memory for a short time, which posits the robustness of choice-related memories as fundamental for performance. We varied global coupling strength to individually simulate both strongly and weakly distributed circuit models. As in (Finkelstein et al., 2021), distractors with different duration and strength were fed to the model, revealing distinct reactions to distractors and suggesting that early distractors are most effective than late ones. In weakly distributed models, it was easier for distractors to disturb the ramping activity, perhaps reflecting the performance of poorly trained monkeys. Future experiments could address how monkeys at different trained stages react to distractors, to test the existence of these two effects behind temporal gating (Treue and Maunsell, 1999).

Materials and Methods

The large-scale brain model presented here uses previous modeling work on distributed working memory as a basis (Mejias and Wang, 2022), and substantially improves upon it with the introduction of (i) a larger number of cortical areas interconnected by neuroanatomical data, (ii) area-specific data on the density of NMDA and GABA receptors across the brain, obtained via autoradiography experiments, and (iii) novel theoretical assumptions on the relative impact on the effects of the density of dendritic spines and postsynaptic receptors on intra- and inter-areal interactions.

We describe our model in three steps: the local circuit employed, the differentiation of such circuit across the whole neocortex via synaptic gradients, and the inter-areal projections. We then briefly define several metrics for behavioral performance and inactivation effects, ending with a short summary on our simplified 'toy model' for winning onset results.

Computational model: local neural circuit

We employed the Wong-Wang model (Wong and Wang, 2006) to characterize the neural dynamics of a local microcircuit representing a cortical area. This model, in its three-variable version, captures the temporal evolution of the firing rates of two input-selective excitatory populations, as well as the firing rate dynamics of an inhibitory population. The populations are

interconnected with each other, as depicted in Figure 1A. The model is governed by the following equations:

$$\frac{dS_A}{dt} = -\frac{S_A}{\tau_N} + \gamma\,(1 - S_A)\,r_A \qquad \text{(Eq. 1)}$$

$$\frac{dS_B}{dt} = -\frac{S_B}{\tau_N} + \gamma\,(1 - S_B)\,r_B \qquad \text{(Eq. 2)}$$

$$\frac{dS_C}{dt} = -\frac{S_C}{\tau_G} + \gamma_I\,r_C \qquad \text{(Eq. 3)}$$

Here, S_A and S_B are the NMDA conductances of selective excitatory populations A and B respectively, and S_C is the GABAergic conductance of the inhibitory population. Values for the constants are τ_N=60 ms, τ_G=5 ms, γ=1.282 and γ_I=2. The variables r_A, r_B and r_C are the mean firing rates of the two excitatory and one inhibitory population, respectively.

They are obtained by solving, at each time step, the transcendental equation $r_i = \phi_i(I_i)$ (where ϕ is the transfer function of the population, detailed below), with I_i being the input to population 'i', given by

$$I_A = J_s S_A + J_c S_B + J_{EI} S_C + I_{0A} + I_{net}^A + x_A(t) \qquad \text{(Eq. 4)}$$

$$I_B = J_c S_A + J_s S_B + J_{EI} S_C + I_{0B} + I_{net}^B + x_B(t) \qquad \text{(Eq. 5)}$$

$$I_C = J_{IE} S_A + J_{IE} S_B + J_{II} S_C + I_{0C} + I_{net}^C + x_C(t) \qquad \text{(Eq. 6)}$$

In these expressions, J_s, J_c are the self- and cross-coupling between excitatory populations, respectively, J_{EI} is the coupling from the inhibitory populations to any of the excitatory ones, J_{IE} is the coupling from any of the excitatory populations to the inhibitory one, and J_{II} is the self-coupling strength of the inhibitory population. The parameters I_{0i} with i=A, B, C are background inputs to each population. Fixed parameters across the cortex are J_c=0.0107 nA, J_{EI}=-0.31 nA, J_{II}=-0.20 nA, and I_{0C}=0.26 nA (background currents for populations A and B are specified later). The term I^i_{net} denotes the long-range input coming from other areas in the network and will be detailed later. Sensory stimulation can be introduced here as extra pulse currents of strength I_{stimA} =0.3(1+c) and I_{stimB}=0.3(1-c) to the V1 excitatory populations A and B respectively, with a duration of T_{pulse} =0.7 sec, unless specified otherwise.

The last term $x_i(t)$ with i=A, B, C is an Ornstein-Uhlenbeck process, which introduces some level of stochasticity in the system. It is given by

$$\tau_{noise}\,\frac{dx_i}{dt} = -x_i + \sqrt{\tau_{noise}}\,\sigma_i\,\xi_i(t) \qquad \text{(Eq. 7)}$$

Here, $\xi_i(t)$ is a Gaussian white noise, the time constant is τ_{noise}=2 ms and the noise strength is $\sigma_{A,B}$=0.01 nA for excitatory populations and σ_C=0 for the inhibitory one.

The transfer function $\phi_i(t)$ which transform the input into firing rates takes the following form for the excitatory populations(Abbott and Chance, 2005):

$$\phi_{A,B}(I) = \frac{1}{2}\,\frac{aI-b}{1-\exp[-d\,(aI-b)]} \qquad \text{(Eq. 8)}$$

The values for the parameters are a=135 Hz/nA, b=54 Hz and d=0.308 s. For the inhibitory population a similar function can be used, but for convenience we choose a threshold-linear function:

$$\phi_C(I) = \left[\frac{1}{g_I}\,(c_b I - c_a) + r_0\right]_+ \qquad \text{(Eq. 9)}$$

The notation $[x]_+$ denotes rectification. The values for the parameters are g_I=4, c_b=615 Hz/nA, c_a=177 Hz and r_0=5.5 Hz. Finally, it is sometimes useful for simulations (although not a requirement) to replace the transcendental equation $r_i = \phi_i(I_i)$ by its analogous differential equation, of the form:

$$\tau_r \frac{dr_i}{dt} = -r_i + \phi_i(I_i) \qquad \text{(Eq. 10)}$$

The time constant can take a typical value of τ_r=2 ms.

<u>Computational model: Gradient of synaptic strengths</u>

Before considering the large-scale network and the inter-areal connections, we look into the area-to-area heterogeneity to be included in the model.

Our large-scale cortical system consists of N=40 local cortical areas, for which inter-areal connectivity data is available. Each cortical area is described as a Wong-Wang model of three populations like the ones described in the previous section. Instead of assuming areas to be identical to each other, here we will consider some of the natural area-to-area heterogeneity that has been found in anatomical studies. For example, work from Elston (Elston, 2007) has identified a gradient of pyramidal cell dendritic spine density, from low spine numbers (~600) found in early sensory areas to large spine counts (~9000) found in higher cognitive areas. From an electrophysiological point of view, excitatory postsynaptic potentials (EPSP) have similar values both in early sensory (~1.7$\pm$1.3 mV) and higher cognitive areas (~0.55$\pm$0.43 mV). The combination of these findings suggests an increase of local recurrent strength as we move from sensory to association areas. In addition, cortical areas are distributed along an anatomical hierarchy(Felleman and Van Essen, 1991; Nikola T. Markov et al., 2014). The position of a given area 'i' within this hierarchy, namely h_i, can be computed with a generalized linear model using data on the fraction of supragranular layer neurons (SLN) projecting to and from that area. In particular, we assigned hierarchical values to each area such that the difference in values predicts the SLN of a projection. Concretely, we assign a value H_i to each area A_i so that SLN($A_j \rightarrow A_i$) ~ f (H_i-H_j), with 'f' being a logistic regression. The final hierarchical values are then obtained by normalizing $h_i=H_i/H_{max}$. Further details on the regression are provided elsewhere (Chaudhuri et al., 2015; Nikola T. Markov et al., 2014).

In the following, we will assign the incoming synaptic strength (both local and long-range) of a given area as a linear function of the pyramidal dendritic spine count values observed in anatomical studies, with age-related corrections when necessary. Alternatively, when spine count data is not available for a given area, we will use its position in the anatomical hierarchy, which displays a high correlation with the spine count data, as a proxy for the latter. After this process, the large-scale network will display a gradient of local and long-range recurrent strengths, with sensory/association areas showing weak/strong local connectivity, respectively. We denote the local and long-range strength value of a given area *i* in this gradient as h_i, and this value normalized between zero (bottom of the gradient, area V1) and one. We assume therefore a linear gradient of values of J_s, with its value going from J_{min} to J_{max}:

$$J_s(i) = J_{min} + (J_{max} - J_{min})\, h_i \qquad \text{(Eq. 11)}$$

The above equations consider the impact of the number of dendritic spines per neuron in the synaptic strength of each brain area, but the density of NMDA and $GABA_A$ receptors per neuron must also be taken into account. Since excitatory signals target not only other pyramidal neurons but also interneurons, we assume that the local NMDA receptor density has a positive correlation with the sum of all excitation signals in a given brain region. We use a saturating gradient, or logistic function, to model such area-specific parameters:

$$J_S^i + J_C + J_{IE}^i = \frac{a_1}{1+b_1 \exp\left(-k(NMDA^i - d_1)\right)} + c_1 \qquad \text{(Eq. 12)}$$

Here, the sum of the three excitatory projections in the left-hand side define the strength of excitatory transmission, which depends on $NMDA^i$, the experimental NMDA receptor density of brain area 'i', via a logistic function. The constants $a_1 = 0.5241, b_1 = 1.01, c_1 = 0.01846, d_1 = 0.12, k = 9.35$ are used constrain the slope and roof of the curve. Since receptor density ratio between NMDA and GABA measures the relative excitatory-inhibitory synaptic strength balance, GABAergic receptor density can be introduced via the following balance equation:

$$J_{EI}^i + \frac{1}{2} J_{II} = (J_S^i + J_C + J_{IE}^i)\frac{GABA_A^i}{NMDA^i} \qquad \text{(Eq. 13)}$$

where $GABA^i$ is the experimental $GABA_A$ receptor density of brain area 'i'.

If association areas have large values of J_S, it can influence their spontaneous activity, even without considering inter-areal coupling. To ensure that the spontaneous firing rate of these areas remains within physiologically realistic limits, a viable approach is to enforce that the fixed point of spontaneous activity is the same for all areas, which is a reasonable approximation. This can be done by adjusting the background currents I_0 (for both excitatory populations A and B) on an area-specific basis, which aligns with the differentiated thalamocortical input observed in real brains. We therefore use adaptive background current to balance spontaneous firing rate:

$$\zeta = \frac{\tau_G\, \gamma_I\, c_1}{g_I - J_{II}\, \tau_G\, \gamma_I\, c_1} \qquad \text{(Eq. 14)}$$

$$\beta = \tau_G\, \gamma_I\, \frac{c_1\, I_{0C} + g_I\, r_0 - c_0}{g_I - J_{II}\, \tau_G\, \gamma_I\, c_1} \qquad \text{(Eq. 15)}$$

$$I_0^i = I_0^{LIP} - \left(\left(J_S^i - J_S^{LIP}\right) + 2\left(J_{EI}^i J_{IE}^i - J_{EI}^{LIP} J_{IE}^{LIP}\right)\zeta\right) S_{sp} - \left(J_{EI}^i - J_{EI}^{LIP}\right)\beta \qquad \text{(Eq. 16)}$$

Here I_0^{LIP} is the background current in parietal area LIP and its value has been set $I_0^{LIP} = 0.3294$ as in previous work (Wong and Wang, 2006) to fit temporal ramping excitatory neuron firing rates in LIP. The parameter $S_{sp} = 0.03566$ is the constant value of spontaneous NMDA gating variable in a local disconnected circuit.

Unless specified otherwise, we assume a range of J_{min}=0.225 nA and $J_{max} = 0.42\ nA$ (i.e., below the critical value), so that the model displays distributed attractors as in the case of working memory (Mejias and Wang, 2022)

<u>Computational model: Inter-areal projections</u>

In the present study, the inter-areal projections that connect isolated areas contribute to the formation of the expansive cortical network. We assume that these inter-areal projections originate solely from excitatory neurons, as inhibitory projections typically exhibit a more localized pattern within real circuits and are selective for excitatory neurons. Thus, the network or long-range input term within a specific area "x" from all other cortical areas, can be expressed as follows:

$$I_{A,net}^x = G\ \sum_y W^{xy} SLN^{xy} S_A^y \qquad \text{(Eq. 18)}$$

$$I_{B,net}^x = G\ \sum_y W^{xy} SLN^{xy} S_B^y \qquad \text{(Eq. 19)}$$

$$I_{C,net}^x = \frac{G}{Z} \sum_y W^{xy} (1 - SLN^{xy})(S_A^y + S_B^y) \qquad \text{(Eq. 20)}$$

Here, G is the global coupling strength, which controls the overall long-range projection strength in the network (G=0.52 unless specified otherwise). $Z = 1.2$ is a factor that considers the relative balance between long-range excitatory and inhibitory projections.

Aside from global scaling factors, the effect of long-range projections from population *y* to population *x* is influenced by two factors. The first one, W^{xy}, is the anatomical projection strength as revealed by tract-tracing data (Markov et al., 2013). We use the fraction of labelled neurons (FLN) from population *y* to *x* to constrain our projections values to anatomical data. We rescale these strengths to translate the broad range of FLN values (over five orders of magnitude) to a range more suitable for our firing rate models. We use a rescaling that maintains the proportions between projection strengths, and therefore the anatomical information, that reads

$$W^{xy} = k_1\ (FLN^{xy})^{k_2} \qquad \text{(Eq. 21)}$$

Here, the values of the rescaling are k_1=1.2 and k_2=0.3. The same qualitative behavior can be obtained from the model if other parameter values, or other rescaling functions, are used as long as the network is set into a standard working regime (i.e. signals propagate across areas, global synchronization is avoided, etc.) FLN values are also normalized so that $\sum_y FLN^{xy} = 1$. While

in-degree heterogeneity might impact network dynamics (de Franciscis et al., 2011; Roxin, 2011), this was done to have a better control of the heterogeneity levels of each area, and to minimize confounding factors such as the uncertainty on volume injections of tract tracing experiments and the influence of potential homeostatic mechanisms. In addition, and as done for the local connections, we introduce a gradient of long-range projection strengths using the spine count data: $W^{xy} \rightarrow (J_s(x)/J_{\max})\, W^{xy}$, so that long-range projections display the same gradient as the local connectivity presented above.

The second factor that needs to be taken into account is the directionality of signal propagation across the hierarchy. Feedforward (FF) projections that are preferentially excitatory constitute a reasonable assumption which facilitate signal transmission from sensory to higher areas. On the other hand, having feedback (FB) projections with a preferential inhibitory nature contributes to the emergence of realistic distributed WM patterns (Figure 4) (see also previous work (Nikola T. Markov et al., 2014; Tsushima et al., 2006)). This feature can be introduced, in a gradual manner, by linking the different inter-areal projections with the SLN data, which provides a proxy for the FF/FB nature of a projection (SLN=1 means purely FF, and SLN=0 means purely FB). In the model, we assume a linear dependence with SLN for projections to excitatory populations and with (1-SLN) for projections to inhibitory populations, as shown above.

Following recent evidence of frontal networks having primarily strong excitatory loops (Markowitz et al., 2015), it is convenient to ensure that the SLN-driven modulation of FB projections between frontal areas is not too large, so that interactions between these areas are never strongly inhibitory. In practice, such constraint is only necessary for projections from frontal areas to 8l and 8m (which are part of the frontal eye fields) and has little effect on the behavior of our model otherwise. The introduction of this limitation has two minor consequences: (i) it allows area 8l and 8m to exhibit a higher level of persistent activity during distributed WM –as their hierarchical position and recurrent strength are not strong enough to sustain activity otherwise, as previously suggested in anatomical studies(Markov et al., 2013; Nikola T. Markov et al., 2014), and (ii) it slightly shifts the transition point in cortical space. Unless specified otherwise, we consider that the SLN-driven modulation of FB projections to 8l and 8m is never larger than 0.4.

Behavioral performance: psychometric curves

Psychometric curves were fitted with a binomial generalized linear regression model (GLM):

$$\log \frac{y}{1-y} = \beta_1 x + \beta_0$$

where y was the probability to choose population A as the final decision and x was the contrast strength. $\beta = (\beta_1, \beta_0)$ were the parameters of the fit, with β_1 representing the slope of the psychometric curve and β_0 the value at zero contrast.

Evaluation of lesioning effects

To replicate the results of the lesioning protocol, we employed a manual approach to induce lesioning effects in the target area's three populations. Specifically, we set the firing rates of all three populations to zero. We evaluated the impact of such lesioning by assessing the percentage change in ramping speed, which measures the speed at which a specific area encodes the final decision:

$$RS_i = \frac{RR_i - WR_i}{RO_i - WO_i} \quad \text{(Eq. 22)}$$

Here, RS_i is the ramping speed for area i in the decision-making task with the intact recurrent network. As mentioned in the main text, the winning onset (WO_i) is the time at which the difference between the firing rates of both excitatory populations becomes larger than a certain threshold (zero for Fig. 3, 0.5 for Figs. 4 and 5) and increases from there without returning to zero (see Fig. 3B). The winning rate (WR_i) is defined as the firing rate of the winning population at the winning onset, the reaching rate (RR_i) is defined as (0.75 x attractor firing rate – 0.25 x winning rate), and the reaching onset (RO_i) is defined as the time at which the winning population firing rate arrives at the reaching rate. We compare how lesioning one area affects the other area under same realization of Gaussian white noise with the following expression:

$$LE_{i,j} = \frac{RS_{i,j} - RS_i}{RS_i} \quad \text{(Eq. 23)}$$

where $LE_{i,j}$ is the degree by which lesioning area 'j' affects the ramping speed of area 'i', and $RS_{i,j}$ is the ramping speed in area 'i' when area 'j' is lesioned.

<u>A toy model accounting for the shift of winning onset order</u>

The toy model of Figure 3E consists of three cortical areas V1, MT and 9/46v, with V1 transmitting motion information, MT as an intermediate area connecting association areas with sensory areas, and 9/46v as an association area to sustain delay activity. The computational model structure is the same except for the long-range projections, which are now as follows:

$$I_{A,net}^{x} = \sum_y W_*^{xy} S_A^y \quad \text{(Eq. 24)}$$

$$I_{B,net}^{x} = \sum_y W_*^{xy} S_B^y \quad \text{(Eq. 25)}$$

$$I_{C,net}^{x} = \sum_y W_-^{xy} (S_A^y + S_B^y) \quad \text{(Eq. 26)}$$

Here, W_*^{xy}, W_-^{xy} are parameters assigning individual roles to the three areas. All parameter values of the model can be found in Supplementary Table S1.

Acknowledgments: We thank Kai Chen, Ziling Wang, Boran Yang, Chongming Liu for their support during the development of this work, Sean Froudist-Walsh for his help with data formatting and visualization, and Henry Kennedy for providing the connectivity dataset.

Funding: This work did not receive direct funding support.

Author contributions: JFM designed the study; NPG contributed datasets; LZ performed the research; LZ, DZ, SL and JFM analyzed and discussed the results; LZ, NPG, DZ, SL and JFM wrote the manuscript.

Competing interests: Authors declare not competing interests.

Data and materials availability: All information needed to reproduce the results of this manuscript are in the main text and Materials and Methods section, and the code used will be made available upon publication of this work.

References

Abbott LF, Chance FS. 2005. Drivers and modulators from push-pull and balanced synaptic input. *Prog Brain Res* **149**:147–155. doi:10.1016/S0079-6123(05)49011-1

Akrami A, Kopec CD, Diamond ME, Brody CD. 2018. Posterior parietal cortex represents sensory history and mediates its effects on behaviour. *Nature* **554**:368–372. doi:10.1038/nature25510

Beerendonk L, Mejías JF, Nuiten SA, Gee JW de, Fahrenfort JJ, Gaal S van. 2023. A disinhibitory circuit mechanism explains a general principle of peak performance during mid-level arousal. doi:10.1101/2023.07.28.550956

Bisley JW, Zaksas D, Droll JA, Pasternak T. 2004. Activity of Neurons in Cortical Area MT During a Memory for Motion Task. *Journal of Neurophysiology* **91**:286–300. doi:10.1152/jn.00870.2003

Brunet N, Jagadeesh B. 2019. Familiarity with visual stimuli boosts recency bias in macaques. *PeerJ* **7**:e8105. doi:10.7717/peerj.8105

Brunton BW, Botvinick MM, Brody CD. 2013. Rats and Humans Can Optimally Accumulate Evidence for Decision-Making. *Science* **340**:95–98. doi:10.1126/science.1233912

Chandrasekaran C. 2017. Computational principles and models of multisensory integration. *Current Opinion in Neurobiology*, Neurobiology of Learning and Plasticity **43**:25–34. doi:10.1016/j.conb.2016.11.002

Chaudhuri R, Knoblauch K, Gariel M-A, Kennedy H, Wang X-J. 2015. A Large-Scale Circuit Mechanism for Hierarchical Dynamical Processing in the Primate Cortex. *Neuron* **88**:419–431. doi:10.1016/j.neuron.2015.09.008

Christophel TB, Klink PC, Spitzer B, Roelfsema PR, Haynes J-D. 2017. The Distributed Nature of Working Memory. *Trends in Cognitive Sciences* **21**:111–124. doi:10.1016/j.tics.2016.12.007

de Franciscis S, Johnson S, Torres JJ. 2011. Enhancing neural-network performance via assortativity. *Phys Rev E* **83**:036114. doi:10.1103/PhysRevE.83.036114

Elston G, Benavides-Piccione R, Elston A, Manger P, Defelipe J. 2011. Pyramidal Cells in Prefrontal Cortex of Primates: Marked Differences in Neuronal Structure Among Species. *Frontiers in Neuroanatomy* **5**.

Elston GN. 2007. Specialization of the Neocortical Pyramidal Cell during Primate Evolution. *Evolution of Nervous Systems* 191–242. doi:10.1016/b0-12-370878-8/00164-6

Erlich JC, Brunton BW, Duan CA, Hanks TD, Brody CD. 2015. Distinct effects of prefrontal and parietal cortex inactivations on an accumulation of evidence task in the rat. *eLife* **4**:e05457. doi:10.7554/eLife.05457

Felleman DJ, Van Essen DC. 1991. Distributed hierarchical processing in the primate cerebral cortex. *Cereb Cortex* **1**:1–47. doi:10.1093/cercor/1.1.1

Feng M, Bandyopadhyay A, Mejias JF. 2023. Emergence of distributed working memory in a human brain network model. doi:10.1101/2023.01.26.525779

Finkelstein A, Fontolan L, Economo MN, Li N, Romani S, Svoboda K. 2021. Attractor dynamics gate cortical information flow during decision-making. *Nat Neurosci* **24**:843–850. doi:10.1038/s41593-021-00840-6

Froudist-Walsh S, Bliss DP, Ding X, Rapan L, Niu M, Knoblauch K, Zilles K, Kennedy H, Palomero-Gallagher N, Wang X-J. 2021. A dopamine gradient controls access to distributed working memory in the large-scale monkey cortex. *Neuron* **109**:3500-3520.e13. doi:10.1016/j.neuron.2021.08.024

Froudist-Walsh S, Niu M, Rapan L, Palomero-Gallagher N. 2023a. Neurotransmitter receptor densities per neuron across macaque cortex (v1.0). doi:10.25493/5HK3-S8M

Froudist-Walsh S, Xu T, Niu M, Rapan L, Zhao L, Margulies DS, Zilles K, Wang X-J, Palomero-Gallagher N. 2023b. Gradients of neurotransmitter receptor expression in the macaque cortex. *Nat Neurosci* **26**:1281–1294. doi:10.1038/s41593-023-01351-2

Guo ZV, Li N, Huber D, Ophir E, Gutnisky D, Ting JT, Feng G, Svoboda K. 2014. Flow of Cortical Activity Underlying a Tactile Decision in Mice. *Neuron* **81**:179–194. doi:10.1016/j.neuron.2013.10.020

Hanks TD, Kopec CD, Brunton BW, Duan CA, Erlich JC, Brody CD. 2015. Distinct relationships of parietal and prefrontal cortices to evidence accumulation. *Nature* **520**:220–223. doi:10.1038/nature14066

Hanks TD, Summerfield C. 2017. Perceptual Decision Making in Rodents, Monkeys, and Humans. *Neuron* **93**:15–31. doi:10.1016/j.neuron.2016.12.003

Hansen JY, Shafiei G, Markello RD, Smart K, Cox SML, Nørgaard M, Beliveau V, Wu Y, Gallezot J-D, Aumont É, Servaes S, Scala SG, DuBois JM, Wainstein G, Bezgin G, Funck T, Schmitz TW, Spreng RN, Galovic M, Koepp MJ, Duncan JS, Coles JP, Fryer TD, Aigbirhio FI, McGinnity CJ, Hammers A, Soucy J-P, Baillet S, Guimond S, Hietala J, Bedard M-A, Leyton M, Kobayashi E, Rosa-Neto P, Ganz M, Knudsen GM, Palomero-Gallagher N, Shine JM, Carson RE, Tuominen L, Dagher A, Misic B. 2022. Mapping neurotransmitter systems to the structural and functional organization of the human neocortex. *Nat Neurosci* **25**:1569–1581. doi:10.1038/s41593-022-01186-3

Hayden B, Gallant J. 2013. Working Memory and Decision Processes in Visual Area V4. *Frontiers in Neuroscience* **7**.

Hyafil A, Moreno-Bote R. 2017. Breaking down hierarchies of decision-making in primates. *eLife* **6**:e16650. doi:10.7554/eLife.16650

Jaramillo J, Mejias JF, Wang X-J. 2019. Engagement of Pulvino-cortical Feedforward and Feedback Pathways in Cognitive Computations. *Neuron* **101**:321-336.e9. doi:10.1016/j.neuron.2018.11.023

Joglekar MR, Mejias JF, Yang GR, Wang X-J. 2018. Inter-areal Balanced Amplification Enhances Signal Propagation in a Large-Scale Circuit Model of the Primate Cortex. *Neuron* **98**:222-234.e8. doi:10.1016/j.neuron.2018.02.031

Katz LN, Yates JL, Pillow JW, Huk AC. 2016. Dissociated functional significance of decision-related activity in the primate dorsal stream. *Nature* **535**:285–288. doi:10.1038/nature18617

Klatzmann U, Froudist-Walsh S, Bliss DP, Theodoni P, Mejías J, Niu M, Rapan L, Palomero-Gallagher N, Sergent C, Dehaene S, Wang X-J. 2023. A connectome-based model of conscious access in monkey cortex. doi:10.1101/2022.02.20.481230

Leavitt ML, Mendoza-Halliday D, Martinez-Trujillo JC. 2017. Sustained Activity Encoding Working Memories: Not Fully Distributed. *Trends in Neurosciences* **40**:328–346. doi:10.1016/j.tins.2017.04.004

Li S, Wang X-J. 2022. Hierarchical timescales in the neocortex: Mathematical mechanism and biological insights. *Proceedings of the National Academy of Sciences* **119**:e2110274119. doi:10.1073/pnas.2110274119

Lohuis MNO, Marchesi P, Pennartz CMA, Olcese U. 2022. Functional (ir)Relevance of Posterior Parietal Cortex during Audiovisual Change Detection. *J Neurosci* **42**:5229–5245. doi:10.1523/JNEUROSCI.2150-21.2022

Mante V, Sussillo D, Shenoy KV, Newsome WT. 2013. Context-dependent computation by recurrent dynamics in prefrontal cortex. *Nature* **503**:78–84. doi:10.1038/nature12742

Markov NT, Ercsey-Ravasz M, Van Essen DC, Knoblauch K, Toroczkai Z, Kennedy H. 2013. Cortical high-density counterstream architectures. *Science* **342**:1238406. doi:10.1126/science.1238406

Markov N. T., Ercsey-Ravasz MM, Ribeiro Gomes AR, Lamy C, Magrou L, Vezoli J, Misery P, Falchier A, Quilodran R, Gariel MA, Sallet J, Gamanut R, Huissoud C, Clavagnier S, Giroud P, Sappey-Marinier D, Barone P, Dehay C, Toroczkai Z, Knoblauch K, Van Essen DC, Kennedy H. 2014. A weighted and directed interareal connectivity matrix for macaque cerebral cortex. *Cereb Cortex* **24**:17–36. doi:10.1093/cercor/bhs270

Markov Nikola T., Vezoli J, Chameau P, Falchier A, Quilodran R, Huissoud C, Lamy C, Misery P, Giroud P, Ullman S, Barone P, Dehay C, Knoblauch K, Kennedy H. 2014. Anatomy of hierarchy: feedforward and feedback pathways in macaque visual cortex. *J Comp Neurol* **522**:225–259. doi:10.1002/cne.23458

Markowitz DA, Curtis CE, Pesaran B. 2015. Multiple component networks support working memory in prefrontal cortex. *Proc Natl Acad Sci USA* **112**:11084–11089. doi:10.1073/pnas.1504172112
Mashour GA, Roelfsema P, Changeux J-P, Dehaene S. 2020. Conscious Processing and the Global Neuronal Workspace Hypothesis. *Neuron* **105**:776–798. doi:10.1016/j.neuron.2020.01.026
Meijer GT, Marchesi P, Mejias JF, Montijn JS, Lansink CS, Pennartz CMA. 2020. Neural Correlates of Multisensory Detection Behavior: Comparison of Primary and Higher-Order Visual Cortex. *Cell Reports* **31**:107636. doi:10.1016/j.celrep.2020.107636
Mejias JF, Murray JD, Kennedy H, Wang X-J. 2016. Feedforward and feedback frequency-dependent interactions in a large-scale laminar network of the primate cortex. *Sci Adv* **2**:e1601335. doi:10.1126/sciadv.1601335
Mejias JF, Wang X-J. 2022. Mechanisms of distributed working memory in a large-scale network of macaque neocortex. *eLife* **11**:e72136. doi:10.7554/eLife.72136
Murray JD, Jaramillo J, Wang X-J. 2017. Working Memory and Decision-Making in a Frontoparietal Circuit Model. *J Neurosci* **37**:12167–12186. doi:10.1523/JNEUROSCI.0343-17.2017
Nunes RV, Reyes MB, Mejias JF, de Camargo RY. 2021. Directed functional and structural connectivity in a large-scale model for the mouse cortex. *Network Neuroscience* **5**:874–889. doi:10.1162/netn_a_00206
Pinto L, Tank DW, Brody CD. 2022. Multiple timescales of sensory-evidence accumulation across the dorsal cortex. *ELife* **11**:e70263.
Raposo D, Kaufman MT, Churchland AK. 2014. A category-free neural population supports evolving demands during decision-making. *Nat Neurosci* **17**:1784–1792. doi:10.1038/nn.3865
Roach JP, Churchland AK, Engel TA. 2023. Choice selective inhibition drives stability and competition in decision circuits. *Nat Commun* **14**:147. doi:10.1038/s41467-023-35822-8
Roitman JD, Shadlen MN. 2002. Response of Neurons in the Lateral Intraparietal Area during a Combined Visual Discrimination Reaction Time Task. *J Neurosci* **22**:9475–9489. doi:10.1523/JNEUROSCI.22-21-09475.2002
Roxin A. 2011. The Role of Degree Distribution in Shaping the Dynamics in Networks of Sparsely Connected Spiking Neurons. *Front Comput Neurosci* **5**. doi:10.3389/fncom.2011.00008
Roxin A, Ledberg A. 2008. Neurobiological Models of Two-Choice Decision Making Can Be Reduced to a One-Dimensional Nonlinear Diffusion Equation. *PLOS Computational Biology* **4**:e1000046. doi:10.1371/journal.pcbi.1000046
Seidemann E, Zohary E, Newsome WT. 1998. Temporal gating of neural signals during performance of a visual discrimination task. *Nature* **394**:72–75. doi:10.1038/27906
Shadlen MN, Newsome WT. 2001. Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey. *J Neurophysiol* **86**:1916–1936. doi:10.1152/jn.2001.86.4.1916
Siegel M, Buschman TJ, Miller EK. 2015. Cortical information flow during flexible sensorimotor decisions. *Science* **348**:1352–1355. doi:10.1126/science.aab0551
Song HF, Yang GR, Wang X-J. 2016. Training Excitatory-Inhibitory Recurrent Neural Networks for Cognitive Tasks: A Simple and Flexible Framework. *PLOS Computational Biology* **12**:e1004792. doi:10.1371/journal.pcbi.1004792
Steinmetz NA, Zatka-Haas P, Carandini M, Harris KD. 2019. Distributed coding of choice, action and engagement across the mouse brain. *Nature* **576**:266–273. doi:10.1038/s41586-019-1787-x
Suzuki M, Gottlieb J. 2013. Distinct neural mechanisms of distractor suppression in the frontal and parietal lobe. *Nat Neurosci* **16**:98–104. doi:10.1038/nn.3282
Treue S, Maunsell JHR. 1999. Effects of Attention on the Processing of Motion in Macaque Middle Temporal and Medial Superior Temporal Visual Cortical Areas. *J Neurosci* **19**:7591–7602. doi:10.1523/JNEUROSCI.19-17-07591.1999
Tsushima Y, Sasaki Y, Watanabe T. 2006. Greater disruption due to failure of inhibitory control on an ambiguous distractor. *Science* **314**:1786–1788. doi:10.1126/science.1133197
Voitov I, Mrsic-Flogel TD. 2022. Cortical feedback loops bind distributed representations of working memory. *Nature* **608**:381–389. doi:10.1038/s41586-022-05014-3
Vugt B van, Dagnino B, Vartak D, Safaai H, Panzeri S, Dehaene S, Roelfsema PR. 2018. The threshold for conscious report: Signal loss and response bias in visual and frontal cortex. *Science* **360**:537–542. doi:10.1126/science.aar7186
Wang X-J. 2020. Macroscopic gradients of synaptic excitation and inhibition in the neocortex. *Nat Rev Neurosci* **21**:169–178. doi:10.1038/s41583-020-0262-x
Wang X-J. 2002. Probabilistic Decision Making by Slow Reverberation in Cortical Circuits. *Neuron* **36**:955–968. doi:10.1016/S0896-6273(02)01092-9

Wierda T, Dora S, Pennartz CMA, Mejias JF. 2023. Diverse and flexible behavioral strategies arise in recurrent neural networks trained on multisensory decision making. doi:10.1101/2023.10.28.564511

Wilming N, Murphy PR, Meyniel F, Donner TH. 2020. Large-scale dynamics of perceptual decision information across human cortex. *Nat Commun* **11**:5109. doi:10.1038/s41467-020-18826-6

Wimmer K, Compte A, Roxin A, Peixoto D, Renart A, de la Rocha J. 2015. Sensory integration dynamics in a hierarchical network explains choice probabilities in cortical area MT. *Nat Commun* **6**:6177. doi:10.1038/ncomms7177

Wong K-F, Wang X-J. 2006. A Recurrent Network Mechanism of Time Integration in Perceptual Decisions. *J Neurosci* **26**:1314–1328. doi:10.1523/JNEUROSCI.3733-05.2006

Ye L, Li C. 2021. Quantifying the Landscape of Decision Making From Spiking Neural Networks. *Frontiers in Computational Neuroscience* **15**.

Zhou D, Xiao Y, Zhang Y, Xu Z, Cai D. 2014. Granger Causality Network Reconstruction of Conductance-Based Integrate-and-Fire Neuronal Systems. *PLOS ONE* **9**:e87636. doi:10.1371/journal.pone.0087636

Zilles K, Palomero-Gallagher N. 2017. Multiple Transmitter Receptors in Regions and Layers of the Human Cerebral Cortex. *Frontiers in Neuroanatomy* **11**.

Supplementary figures

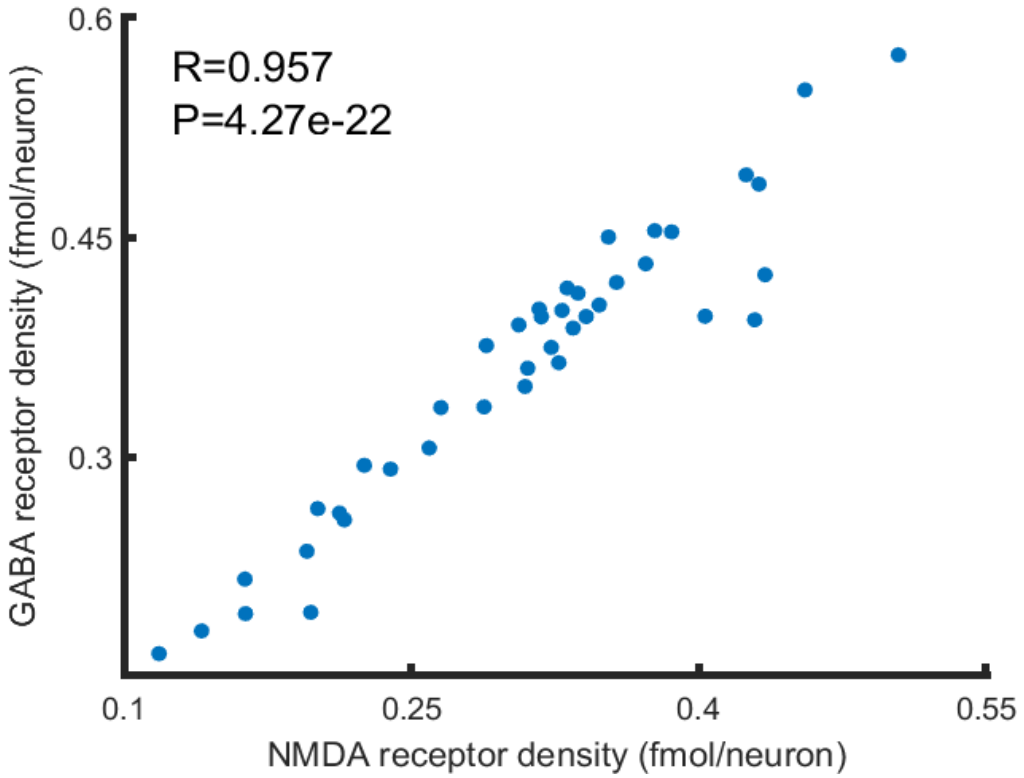

Figure S1: Correlation between the density of $GABA_A$ and NMDA receptors per neuron across all cortical areas considered. The high correlation allows to adopt several modeling assumptions (see Methods).

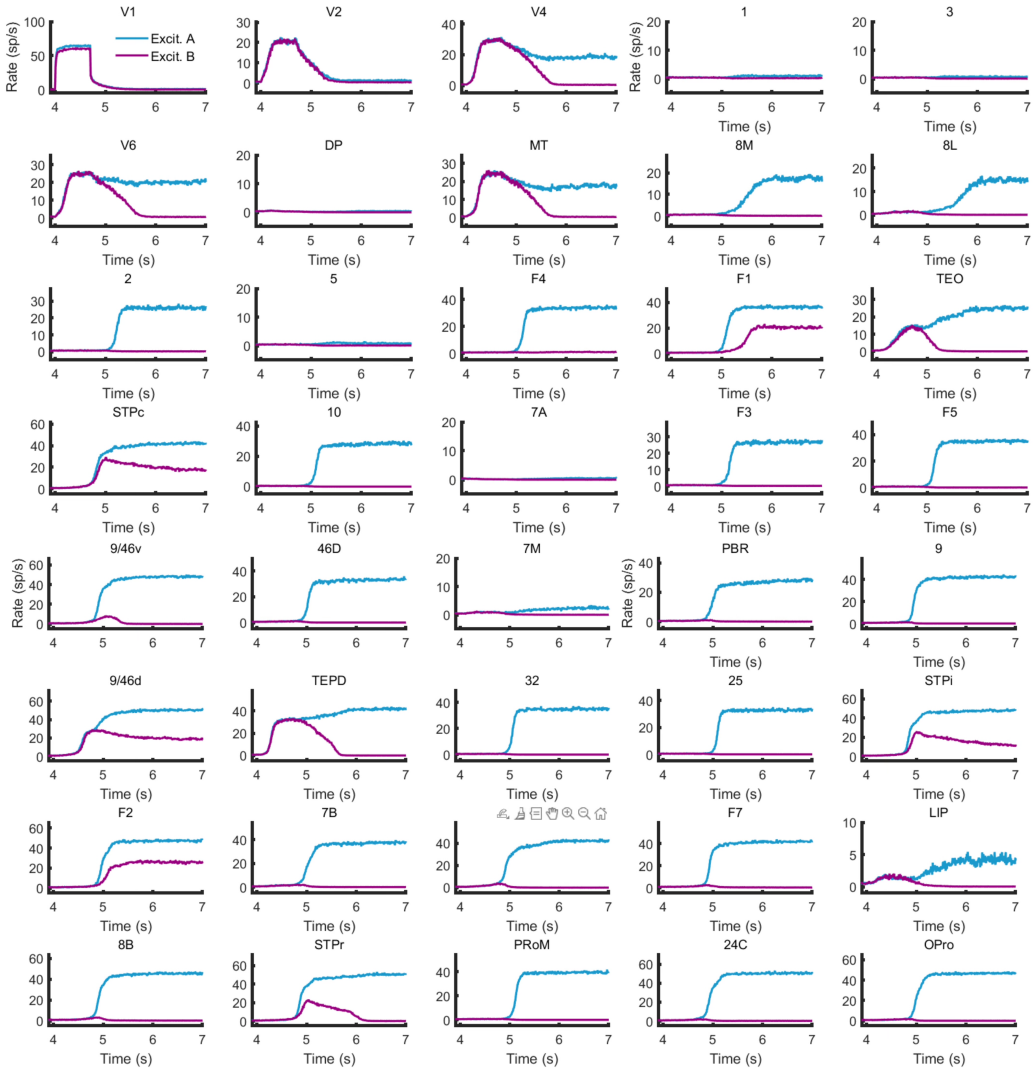

Figure S2: Evidence accumulation is distributed across multiple cortical areas. Integration of sensory evidence across all 40 cortical areas in the model, for a visual input of 5% coherence entering V1. We can distinguish between sensory-driven areas (such as V1 and V2), accumulators (like TEO, LIP and 9/46D) and classificators (like 45A and 24c) depending on their response profile and speed of integration.

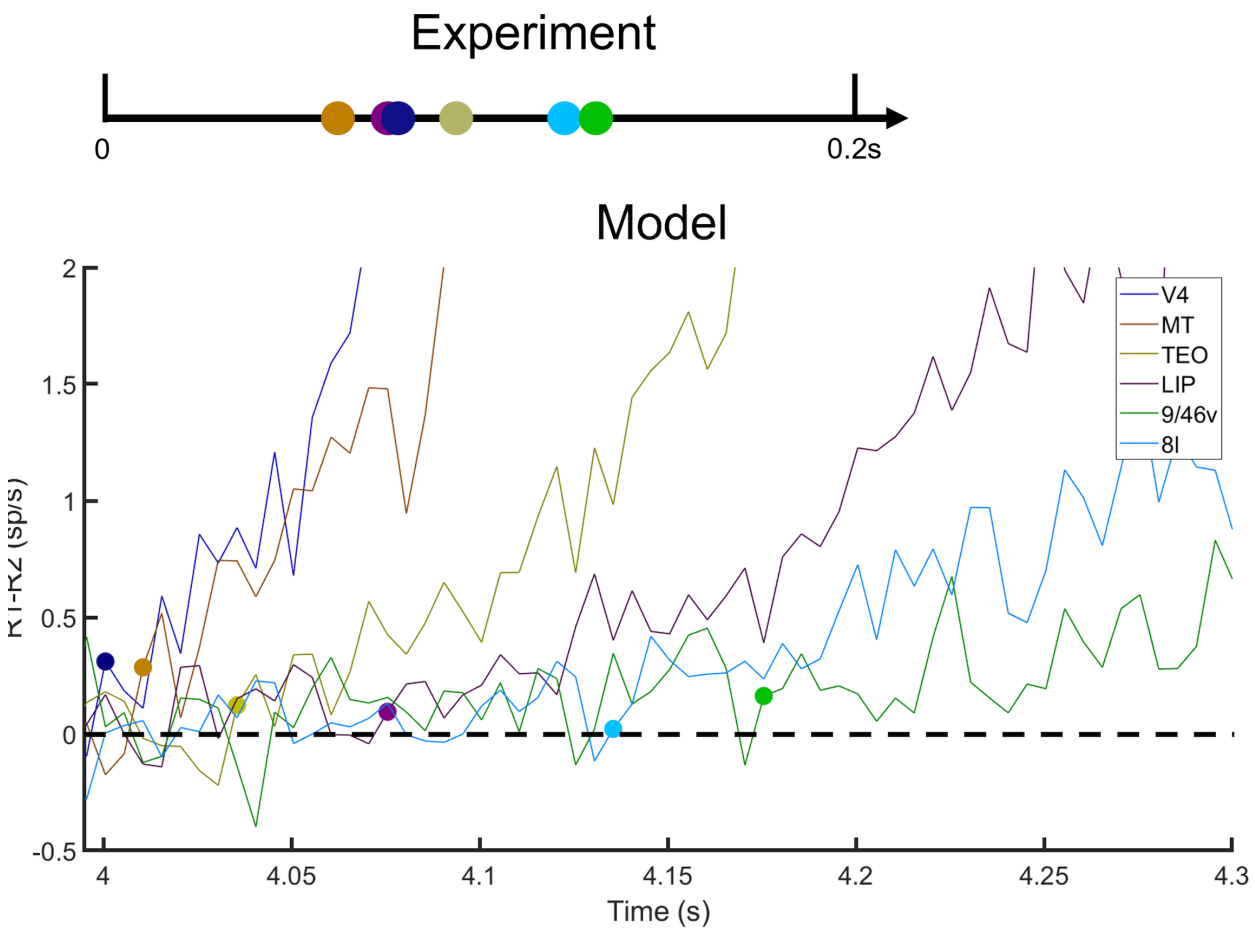

Figure S3: Comparison of area-specific decision times in experiments and model. Top: Timing of onsets in the cortical information flow in categorical decision-making as measured experimentally in macaques (Siegel et al., 2015). Bottom: Timing of the winning onsets for decision making as predicted by our model. Area legend applies to both panels. Both experimental data and model predictions have the same overall activation windonw of ~0.2 seconds and both follow the same order, except for areas V4 and TEO/IT.

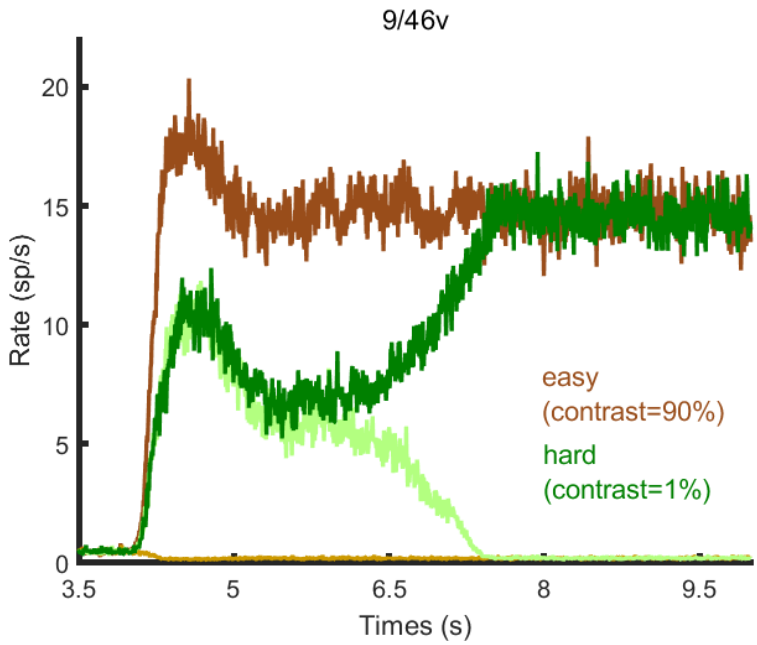

Figure S4: Evidence accumulation at prefrontal cortex (9/46v) during the decision-making task for the simplified three-area model (V1, MT, 9/46v). As in the full-network model, trials with low coherence lead to longer decision times.

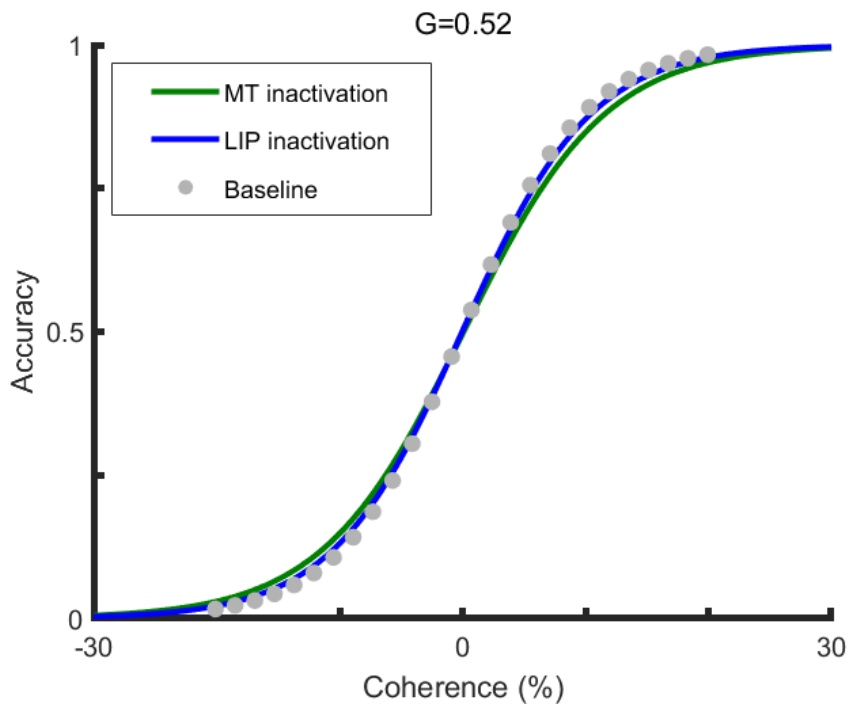

Figure S5: Lesioning effect on behavioral performance for strongly connected networks. Psychometric curve of the full-network model for the control case (baseline), inactivation of MT (green) and inactivation of LIP (blue) for global coupling of G=0.52. Due to the strong interaction between cortical areas, lesioning either MT or LIP does not lead to substantial performance drops in the task, in contrast to the case of weakly connected networks described in the main text.

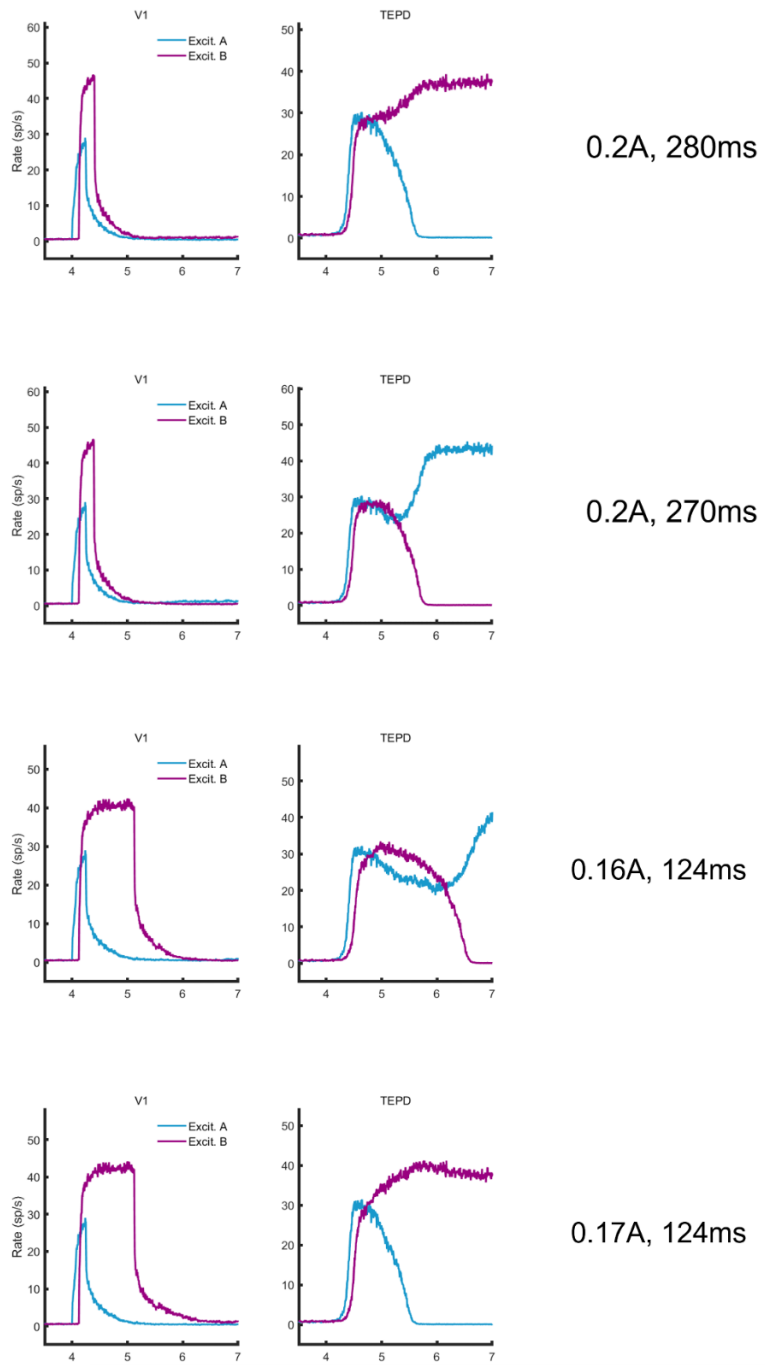

Figure S6: Temporal gating effects. Impact of temporal gating of distractor signals (purple) on the maintenance of cue information (blue) for V1 and TEpd. Top and middle-top: effects of changing the duration from 280 ms to 270 ms in a fixed onset task. Middle-bottom and bottom: effects of changing the distractor strength in a fixed duration task.

Supplementary tables

	V1	MT	PFC
J_s	0.25	0.42	0.29
J_{IE}	0.015	0.05	0.1
I_{0I}	0.26	0.26	0.26
I_{0E}	0.3195	0.3192	0.3172
W_E from V1 to area:	0	0.07	0.01
W_E from MT to area:	0.01	0	0.1
W_E from PFC to area:	0.01	0.07	0
W_I from V1 to area:	0	0.001	0.01
W_I from MT to area:	0.01	0	0.005
W_I from PFC to area:	0.01	0.05	0

Table S1*: Parameters of the toy model used to explain the shift in winning onset values.*