Differential Roles of Mediodorsal Nucleus of the Thalamus and Prefrontal Cortex in Decision-Making and State Representation in a Cognitive Control Task Measuring Deficits in Schizophrenia

Adele L. DeNicola,^{1,4} Min-Yoon Park,^{1,4} David A. Crowe,^{4,5} Angus W. MacDonald 3rd,^{2,3} and Matthew V. Chafee^{1,2,4}

¹Department of Neuroscience, ²Center for Cognitive Sciences, ³Department of Psychology, University of Minnesota, Minneapolis, Minnesota, 55455, ⁴Brain Sciences Center, Veterans Affairs Medical Center, Minneapolis, Minnesota, 55417, and ⁵Department of Biology, Augsburg University, Minneapolis, Minnesota, 55454

The mediodorsal nucleus of the thalamus (MD) is reciprocally connected with the prefrontal cortex (PFC), and although the MD has been implicated in a range of PFC-dependent cognitive functions (Watanabe and Funahashi, 2012; Mitchell and Chakraborty, 2013; Par-naudeau et al., 2018), little is known about how MD neurons in the primate participate specifically in cognitive control, a capability that reflects the ability to use contextual information (such as a rule) to modify responses to environmental stimuli. To learn how the MD-PFC thalamocortical network is engaged to mediate forms of cognitive control that are selectively disrupted in schizophrenia, we trained male monkeys to perform a variant of the AX continuous performance task, which reliably measures cognitive control deficits in patients (Henderson et al., 2012) and used linear multielectrode arrays to record neural activity in the MD and PFC simultaneously. We found that the two structures made clearly different contributions to distributed processing for cognitive control: MD neurons were specialized for decision-making and response selection, whereas prefrontal neurons were specialized to preferentially encode the environmental state on which the decision was based. In addition, we observed that functional coupling between MD and PFC was strongest when the decision as to which of the two responses in the task to execute was being made. These findings delineate unique contributions of MD and PFC to distributed processing for cognitive control and characterized neural dynamics in this network associated with normative cognitive control performance.

Key words: cognitive control; MD thalamus; neurophysiology; prefrontal; primate; schizophrenia

Significance Statement

Cognitive control is fundamental to healthy human executive functioning (Miller and Cohen, 2001) and deficits in patients with schizophrenia relate to decreased functional activation of the MD thalamus and the prefrontal cortex (Minzenberg et al., 2009), which are reciprocally linked (Goldman-Rakic and Porrino, 1985; Xiao et al., 2009). We carry out simultaneous neural recordings in the MD and PFC while monkeys perform a cognitive control task translated from patients with schizophrenia to relate thalamocortical dynamics to cognitive control performance. Our data suggest that state representation and decision-making computations for cognitive control are preferentially performed by PFC and MD, respectively. This suggests experiments to parse decision-making and state representation deficits in patients while providing novel computational targets for future therapies.

Introduction

Cognitive control is the ability to use internally stored rules or goals to modify mapping between sensory inputs and motor out-

puts, the flexible and purposeful resolution of action conflict. The prefrontal cortex (PFC) has been implicated as a key brain region involved in a number of aspects of cognitive control, including

Received July 16, 2019; revised Dec. 12, 2019; accepted Jan. 5, 2020.

Author contributions: A.L.D. and M.V.C. designed research; A.L.D. and M.-Y.P. performed research; A.L.D., D.A.C., and M.V.C. contributed unpublished reagents/analytic tools; A.L.D. and M.V.C. analyzed data; A.L.D., M.-Y.P., D.A.C., A.W.M., and M.V.C. wrote the paper.

This work was supported by the National Institutes of Health (R01MH107491, F31MH109238, and T32GM847121), Minnesota's Discovery, Research, and Innovation Economy Neuromodulation Fellowship (Mn-DRIVE), and the American Brain Sciences Chair. This material is the result of work supported with resources and the use of facilities at the Veteran's Affairs Medical Center in Minneapolis, MN. We thank David Redish and Sophia Vinogradov for their important insights regarding the relation of this work to state representation and cognitive

the neural representation of abstract internal state variables that govern flexible and context-appropriate action (Goldman-Rakic, 1987; Miller and Cohen, 2001; Schiffer et al., 2015). PFC neurons have been shown to encode the state of the task (Asaad et al., 2000), rules (Wallis et al., 2001; Buschman et al., 2012), and abstract categories (Christoff et al., 2009; Goodwin et al., 2012; Crowe et al., 2013), and rapidly adapt to changes in rules by adjusting how stimuli are mapped to categories (Roy et al., 2010; Goodwin et al., 2012; Mante et al., 2013; Crowe et al., 2014; Blackman et al., 2016). Although cognitive control is most commonly considered a cortical function, the mediodorsal nucleus of the thalamus (MD) has arisen as a potential contributor to cognitive control due to both its anatomical connectivity with the PFC (Goldman-Rakic and Porrino, 1985; Xiao et al., 2009) and its coactivation with PFC during cognitive processing. During working memory tasks, primate MD neurons exhibit persistent delay period activity and transmit signals encoding memory-dependent motor responses to PFC (Fuster and Alexander, 1973; Sommer and Wurtz, 2002, 2004a,b; Watanabe and Funahashi, 2004a,b, 2018; Watanabe et al., 2009). Furthermore, the PFC and MD physiologically interact during working memory, as indicated by the finding that PFC inactivation attenuates working memory signals in the MD in primates (Alexander and Fuster, 1973) and rodent studies report that MD inactivation reduces working memory signals in the medial PFC (Bolkan et al., 2017; Schmitt et al., 2017). Additionally, MD lesions in monkeys (Isseff et al., 1982; Zola-Morgan and Squire, 1985; Parker et al., 1997) and inactivation of the MD in rodents (Parnaudeau et al., 2013, 2015, 2018) produce working memory deficits and behavioral inflexibility and lesions in humans restricted to the MD produce executive control deficits (Van der Werf et al., 2000, 2003). However, to our knowledge, no prior study has recorded the activity of MD and PFC neurons simultaneously during a cognitive control task, particularly one that measures selective cognitive control deficits in patients with schizophrenia (MacDonald et al., 2005b; Henderson et al., 2012).

To relate the computations required for forms of cognitive control that are disrupted in schizophrenia to the activity of individual neurons in prefrontal networks, our lab recently translated a variant of the AX continuous performance task (AX-CPT) known as the dot-pattern expectancy (DPX) task that measures deficits in patients to nonhuman primates (Blackman et al., 2013, 2016; Zick et al., 2018). In the current study we use linear electrode arrays to record MD and PFC neural activity simultaneously during DPX task performance. The goals of the current experiment were to: (1) determine whether and how MD neurons contribute to cognitive control, (2) contrast computations performed by MD and PFC neurons for cognitive control, and (3) relate functional communication between MD and PFC neurons to cognitive control. This allowed us to characterize thalamocortical dynamics during normative performance, the disruption of which could lead to cognitive control failure in schizophrenia. We found that MD and PFC neurons made unique contributions to distributed processing for cognitive control: MD neurons were specialized for decision-making and re-

sponse execution, whereas PFC neurons were specialized to represent the environmental state on which the decision was based. Finally, functional coupling between MD and PFC neurons was strongest at the time when the decision was made. These data argue that MD and PFC make differential contributions to cognitive control, raising the question whether state representation or decision-making deficits contribute most directly to cognitive control in schizophrenia.

Materials and Methods

Subjects

We recorded neural activity in the dorsolateral PFC and the MD in two male monkeys (*Macaca mulatta*; 8–11 kg) during performance of the DPX task (Fig. 1A). All animal care and experimental procedures conformed to the National Institutes of Health guidelines and complied with protocols approved by the Animal Care and Use committee at the University of Minnesota and the Minneapolis Veterans Affairs Medical Center.

The DPX task and conditions

The DPX task was described in detail in a previous report from our laboratory (Blackman et al., 2013). Monkeys sat in a primate chair ~77.5 cm in front of a back-projection screen. An LCD projector (Dell 5100MP DLP) presented the visual stimuli. An infrared eye tracking system (ISCAN) recorded eye position at 60 Hz. Each trial consisted of a cue stimulus and probe stimulus presented in sequence, followed by a joystick response (left or right) depending on cue-probe sequence shown. Cue and probe stimuli consisted of patterns of dots: A-cues and X-probes each consisted of a single dot pattern, whereas B-cues and Y-probes were any of five dot patterns each (Fig. 1B). Trials began with a 500 ms period of gaze fixation of a central fixation target (cross). Gaze fixation within 3.3° of the fixation cross was required until the response was made. If monkeys broke gaze fixation before the response the trial was aborted. After the initial fixation period, a cue stimulus was presented at the center of the display for 1.0 s. Following the presentation of the cue there was a 1.0 s delay period that contained only the fixation cross. Next a probe stimulus was displayed for 0.5 s at the center of the display. The cue and probe stimuli subtended 2.7–4.4° of visual angle (Fig. 1B). From the time the probe stimulus was presented, the monkey had 1.0 s to move a hand-held joystick to the left or right. Trials in which an A-cue was followed by an X-probe (AX trials) were “valid” trials, requiring a leftward joystick movement (target response). All other cue-probe combinations (AY, BX, and BY) were “invalid” trials that required a rightward joystick movement (nontarget response; Fig. 1C). The probe was presented for the entirety of the 0.5 s probe period, regardless of the time of the response (analogous to the way the task is administered for human subjects). Monkeys were rewarded with a drop of juice (~0.1 ml) at the end of the intertrial interval (1.1 s after probe offset) for correctly performed trials.

We administered the DPX task in two different trial sets, balanced and prepotent, which differed in the proportion of trials with different cue-probe combinations. In balanced sets there were an equal proportion of the four cue-probe trial sequences (AX, AY, BX, BY). These included 80, 200, or 400 trials [in 20, 27, and 1 neural ensemble(s), respectively]. In prepotent trial sets the majority of trials (69%) presented the target cue-probe sequence (AX). The remaining trials (31%) presented nontarget cue-probe sequences (AY 12.5%, BX 12.5%, BY 6%). Prepotent trial proportions followed from the DPX task and AX-CPT that were administered to patients (MacDonald et al., 2005a; Jones et al., 2010). Prepotent sets included either 320 or 400 trials total (in 19 and 29 ensembles, respectively). For neural ensembles included in this study, we recorded neural activity during the performance of both a balanced and prepotent trial set. The prepotent sets presumably create a prepotent motor habit to produce the leftward joystick movement in response to the X-probe presentation, as this was the required response ~4/5 of the times the X-probe was presented. Overriding this habitual response to the X-probe on B-cue trials therefore required utilization of cognitive control to countermand a prepotent action.

Monkeys received an intramuscular injection of saline [Monkey 1 (Mk1): 0.3 ml; Monkey 2 (Mk2): 0.6 ml] before recording neurophysio-

control; C. Dean Evans for excellent technical assistance in surgeries, neural recording, and with animal care; Dale Boeff for computer programming and engineering support; and Matt Johnson for assistance with histological localization. The contents do not represent the views of the U.S. Department of Veterans Affairs or the United States Government.

The authors declare no competing financial interests.

Correspondence should be addressed to Matthew V. Chafee at chafee001@umn.edu.

<https://doi.org/10.1523/JNEUROSCI.1703-19.2020>

Copyright © 2020 the authors

Figure 1. The DPX task and behavioral performance. **A**, Cue-probe sequence on AX and BX trials. The cue was presented for 1.0 s, followed by a 1.0 s delay period, followed by a 0.5 s presentation of the probe. The monkey could respond from probe onset until 0.5 s following probe offset. The intertrial interval was 1.1 s. Top Row, AX valid sequence requiring leftward (target) joystick movement. Bottom Row, BX invalid sequence requiring rightward (nontarget) movement. ITI, Intertrial interval. **B**, All cue and probe dot patterns used. A-cue and X-probe consisted of a single dot pattern each, whereas B-cues and Y-probes consisted of five different dot patterns each. **C**, The correct movement direction was dependent on the combination of the cue and probe stimuli. The combination of an A-cue and an X-probe required the target, left (L) response. All other combinations (AY; BX; BY) required a nontarget, right (R) response. **D**, Mean (± 2 SEM) proportion of trials correct. **E**, Mean (± 2 SEM) reaction time in seconds is shown split by trial type (cue-probe sequence) and by monkey (Mk 1, open circles and dashed lines; Mk 2, closed circles and solid line). Prepotent and balanced sets combined.

logical data for 18 of the 48 neural ensembles included in this study. These saline injections were part of a study to investigate the effects of the NMDA receptor antagonist phencyclidine on the neural activity in the mediodorsal nucleus of the thalamus and the PFC. All neural data included in the present study were recorded before monkeys had ever been exposed to phencyclidine.

Neural recording

Before surgery structural MRI images of monkeys were obtained at 3T to localize target recording regions: the banks of the principal sulcus in PFC (Brodmann area 46) and the MD (Fig. 2). Recording grid and electrode penetration locations were projected onto monkey-specific cortical surface reconstructions of the locations of cortical sulci obtained from MRI images (Fig. 2*A,B*). To visualize electrode trajectories in depth and target the MD, we used the Cicerone software package (Miocinovic et al., 2007) that generated 3D rotatable MRI volumes from the 2D image stack and visualized chamber placements and electrode angles. This allowed for more precise chamber placement to achieve specific electrode angles that intersected the MD in each monkey. Following recording, we obtained CT scans to visualize recording locations. We inserted stainless steel probes of approximately the same diameter as the recording electrodes passing through the center of locations and depths where MD neural recordings were conducted. The steel probe was visible in the CT scans. We uploaded the CT image stack into Cicerone and aligned the CT and prerecording MR image stacks. We confirmed the trajectory of the electrode penetrations that intersected the MD and plotted the targeted dorsal surface penetrations of the electrode in each monkey (Fig. 2*C–F*; eliminating data from 2 d of recording in which the electrode missed MD).

Monkeys underwent an aseptic surgical procedure under isoflurane gas anesthesia (1–2%). Two craniotomies were made in the skull, one positioned to allow recordings from the dorsolateral PFC and one positioned to allow recordings from the mediodorsal nucleus of the thalamus in the left cerebral hemisphere. Five titanium posts were screwed to the

surface of the skull using titanium screws. PEEK plastic recording chambers [PFC: 19 mm inner diameter (i.d.); MD: 16 mm i.d.] were placed over the craniotomies and fixed to the screws and posts using surgical bone cement. A halo was attached to the posts using metal tabs for head stabilization. Postsurgical analgesia was maintained for several days using an injectable analgesic (Buprenex, 0.05 mg/kg i.m., twice a day).

Neurophysiological recordings were obtained using dual 32-channel silicone vector arrays [NeuroNexus Technologies, Acute Vector Array (Deep Brain)] each advanced independently into the brain using a NAN motorized Microdrive (NAN Instruments, NAN C Drives). Two vector array electrode geometries were used throughout the experiment (“Edge”, 32 electrodes in a linear array, 100 μ m spacing; and the “Poly2”, 2 staggered 16 electrode linear arrays, 50 μ m spacing). Electrode arrays were fixed to the end of a 400 μ m outer diameter (o.d.) stainless steel tube support body. We advanced each electrode array through a stainless-steel guide tube (600 μ m o.d.), which penetrated the dura. The edge design was used for MD recording in 29 ensembles and the poly2 design was used in the remaining 19 ensembles. The poly2 design was used for all 48 ensembles in the PFC. Neural data were recorded via a 64-channel digital signal processing data acquisition system (Tucker-Davis Technologies). Recorded signals on each electrode were amplified and bandpass filtered (between 100 and 5000 Hz) and monitored during neural recording for the emergence of spiking neural activity once electrodes passed from white matter into the MD nucleus and to evaluate stability of spiking activity during recording. Amplified and filtered electrode signals were sampled at a frequency of 24,414.1 Hz and saved to disk along with time stamps indicating behavioral events. We sorted the action potentials of individual neurons based on a principle components analysis of recorded signals implemented by an open-source software package (KiloSort Suite; Pachitariu et al., 2016).

Data analysis

Behavioral analysis. We analyzed differences in proportion of trials performed correctly as a function of trial type (cue-probe sequence) using

Figure 2. Reconstructions of electrode array recording sites. **A, B**, Two-dimensional (2D) reconstructions of chamber and grid locations (red dots indicate grid locations used to record neural data and points of electrode entry at the cortical surface) superimposed on a dorsal view of sulcal anatomy in Mk1 (**A**) and Mk2 (**B**). PS, Principal sulcus; AS, arcuate sulcus; CS, central Sulcus; IPS, intraparietal sulcus. **C, D**, Images show superposition of electrode trajectory (gray line through blue chamber) targeting MD in three-dimensional (3D) reconstruction of MR image sequence with coregistered CT images in Mk1 (**C**) and Mk2 (**D**). Magenta chambers, prefrontal; blue chambers, thalamic chambers. Yellow region indicates location of MD in the Cicerone brain atlas after fitting the thalamus in atlas to each monkey's thalamus in the MR sequence. **E, F**, Volumetric reconstructions of MD thalamus from an axial view, based on primate brain atlas in Cicerone. Yellow dots are the targeted entry points of the electrode on the surface of the MD.

the χ^2 test of independence applied to contingency tables containing counts of correct and error trials per trial type ($p < 0.05$). We analyzed differences in response time as a function of trial type using a Kruskal–Wallis test followed by *post hoc* pairwise Tukey–Kramer tests.

ANCOVA-based classification of neural activity. We applied analysis of covariance (ANCOVA) to trialwise firing rates measured within cue, delay, probe and response periods to detect significant modulation of neural activity in relation to task variables. Cue, delay and probe periods were coextensive with their corresponding task epochs (Fig. 1). The response period was ± 200 ms centered on the time of the motor response which was indicated by the closure of a switch caused by movement of the joystick. We used firing rates in the pre-cue fixation period as a covariate to control for trial-by-trial fluctuations in baseline activity. We applied ANCOVA to firing rates combined across prepotent and balanced trial sets. To identify neurons in which activity varied as a function of the cue stimulus during the cue or delay periods we applied a one-way ANCOVA with cue (A- vs B-cue) as the single factor to firing rates in these time periods. To identify neurons in which activity varied as a function of the cue and probe stimulus during the probe period, we applied a two-way ANCOVA with the cue stimulus (A vs B) and the probe stimulus (X vs Y) as the factors. To identify neurons in which activity varied as a function of the response, we applied a one-way ANCOVA with response direction (target vs nontarget) as the single factor. Neurons were classified on the basis of the trial periods and the task factors yielding significance in the ANCOVAs ($p < 0.05$) as well as their preference for specific stimuli (A- vs B-cues, X- vs Y-probes) or response directions (target vs nontarget). Additionally, we differentiated between neurons that exhibited cue-

selective activity during the cue period (“early cue” neurons), delay period (“delay cue”), or probe period (“late cue”).

To determine whether the proportions of neurons with different types of task-related activity significantly differed between MD and PFC, we performed Z tests of proportions on neuron counts. To determine whether populations of cue, probe and response neurons were significantly biased in their preferences for specific stimuli or response directions, we performed a Z test of proportions on counts of neurons preferring different stimuli and responses during different task epochs ($p < 0.05$, two-tailed).

Analysis of population average activity. To obtain continuous estimations of firing rate over time [spike density functions (SDFs)], we convolved each neuron's spike train with a Gaussian kernel ($\sigma = 40$ ms) using the “ksdensity” function in MATLAB. We averaged single trial SDFs over groups of trials with the same cue-probe sequence (AX, AY, BX, BY) for each neuron (Fig. 4). We then averaged SDFs over neurons within the same ANCOVA-defined categories to visualize the diversity of population activity patterns present (Fig. 5).

We found that individual neurons often exhibited different stimulus or response preferences at different times in the trial (Fig. 5), producing multiple significant relationships to task variables in the ANCOVA. We identified seven neuronal groups each of which exhibited a distinct and consistent pattern of activity modulation and stimulus/response preference during the cue, delay, probe and response periods (Fig. 6). Neurons were assigned to one of the seven groups (so that the groups were non-overlapping) based on the pattern of significance obtained in the ANCOVA. Group 1 neurons (“Early A-cue”) significantly preferred the A-cue

during the cue period. Group 2 neurons significantly preferred the B-cue during the cue period or the Y-probe during the probe period (“Early B-cue” or “Y-probe”). Group 3 neurons significantly preferred the A-cue during the delay or probe periods (“Delay A-cue” or “Late A-cue”). Group 4 neurons significantly preferred the B-cue during the delay or probe periods (“Delay B-cue” or “Late B-cue”). Group 5 neurons significantly preferred the X-probe during the probe period. Groups 6 and 7 neurons significantly preferred the target (Group 6) and nontarget (Group 7) response directions during the response period, respectively. If a neuron was significantly modulated based on multiple task factors as indicated by the pattern of results in the ANCOVA above, assignment to one group was determined by the factor associated with the most significant p value. This grouping scheme is a modification of one we developed previously to describe the heterogeneity of neural responses in PFC (Blackman et al., 2016), elaborated in the present case to accommodate additional neuron types that were prominent in the MD.

Cue preference index. A prominent pattern of activity (in “switch” neurons) consisted of increased activity during the cue period on B-cue trials, and increased activation during the probe period on A-cue trials. To quantify the tendency of neurons to switch their cue preference from B to A between the cue and probe periods, we first computed the cue preference indices contrasting activity on A-cue and B-cue trials using activity in the cue and probe periods, as follows:

$$Index_{cue} = \frac{(A_{cue} - B_{cue})}{(A_{cue} + B_{cue})},$$

$$Index_{probe} = \frac{(A_{probe} - B_{probe})}{(A_{probe} + B_{probe})}.$$

A is the average firing rate of a given neuron during A-cue trials and B is the average firing rate of the same neuron during B-cue trials. Subscripts indicate whether firing rates were taken from the cue or probe periods. Then to capture the switch in cue preference from the cue to probe period, we computed the difference in corresponding cue preference indices:

$$Index_{switch} = Index_{probe} - Index_{cue}.$$

We considered the switch index for a given neuron significant if it was >97.5 or <2.5 percentile of a bootstrap distribution of switch indices computed after randomly shuffling firing rates between A-cue and B-cue conditions ($p < 0.05$, two-tailed). We applied the Z test of proportions ($p < 0.05$) to contrast numbers of neurons that switched from early B-cue to late A-cue preference to the number of neurons showing the reverse preference (early A-cue to late B-cue) in each area. We additionally applied the Z test of proportions to contrast the number of switch neurons (early B-cue to late A-cue) in the MD and the PFC.

Analysis of cognitive control load on neural activity. On prepotent trial sets, during the cue period, B-cues were associated with high cognitive control load (because they countermanded the habitual target response) and A-cues with low cognitive control load. The difference in firing rate between B- and A-cue trials during the cue period therefore provides a measure of cognitive control encoding. Similarly, during the probe period, AY trials were associated with high cognitive control load (because the Y-probe countermanded the habitual target response implied by the A-cue), and AX trials with low cognitive control load. To measure the influence of cognitive control on neural activity, we computed the difference in firing rate during the cue period on B-cue and A-cue trials, as well as the difference in firing rate during the probe period on AY and AX trials. We then compared these rate differences on a per neuron basis between balanced and prepotent trial sets and evaluated differences between task sets using a permutation test in which we shuffled trials of neural activity between the two sets and recomputed rate differences (100 iterations). We considered the difference in cognitive control effects on neural activity between balanced and prepotent datasets significant if it exceeded the 95th percentile of the bootstrap distribution ($p < 0.05$).

Sliding-window regression analysis. We applied a sliding-window regression analysis to evaluate the strength and significance of the relationship between neuronal activity and task variables as a function of time

within the trial. We measured firing rates of single neurons on correct trials of prepotent (see Fig. 9) trial sets, separately, within a sliding 100 ms window (advanced in 20 ms steps), and at each time step regressed the trial-by-trial firing rates onto the cue, probe, and response by fitting the following linear models:

$$R = \beta_0 + \beta_1 Cue + \beta_2 Probe + \beta_3 Cue * Probe + \varepsilon,$$

$$R = \beta_0 + \beta_1 Response + \varepsilon.$$

R is firing rate, Cue (A vs B), $Probe$ (X vs Y), and $Response$ (target vs nontarget) were dummy coded categorical variables, and E is the error (residual). The results of the regression analysis were then expressed as the proportion of explainable variance (PEV) associated with each regressor (Brincat and Miller, 2016), computed using the formula of (Olejnik and Algina, 2003):

$$\omega^2 = \frac{SS_{between\ groups} - df \times MSE}{SS_{total} + MSE}.$$

To visualize the sequential recruitment of neurons, we ranked neurons according to the time within the trial that they were maximally activated as measured by the time to each neuron’s peak PEV. To carry out the ranking, we divided trials into odd and even trial subsets (based on the trial number) and performed separate regression analyses for each neuron on the two trial subsets. We then determined the time to peak PEV using the regression results obtained from one-half of the trials, and ranked neurons on this basis to define their sequential activation. We then plotted PEV time series for each neuron based on this ranking but using the results of the regression analysis applied to the other half of the trials (see Fig. 9A–C). This cross-validated the ranking procedure to increase its robustness and reduced the influence of noise in the PEV time series on the ranking. To compare time and strength of recruitment between PFC and MD, we restricted the analysis to equal numbers of neurons in the two areas.

To determine whether there was a significant difference in the timing of signals across brain areas, we applied a Kolmogorov–Smirnov test to the cumulative distributions of the time to peak PEV. Additionally, we evaluated the significance of differences in population average PEV across the trial period between areas using a permutation test. We randomly shuffled regression results (PEV time series) between brain areas on a per neuron basis (1000 iterations) and compared the original difference between MD and PFC to the differences obtained from the bootstrap distribution. We considered area differences significant if they exceeded the 95th percentile of the bootstrap distribution.

We contrasted PEV time series derived from neural activity on correct and error trials to identify changes in single-neuron encoding of task variables that predicted behavioral outcome (see Fig. 11B, C).

Population decoding analysis. To quantify information encoded by population activity patterns in the MD and PFC about the cue, probe, and response, we performed time-resolved pattern classification (Kleck, 1980; Johnson and Wichern, 1998; Crowe et al., 2010; see Fig. 10). We used the “classify” MATLAB function (using empirical prior probabilities and “diaglinear” covariance matrix estimation). We measured the firing rates of all significant neurons or neurons in specific functional groups in a sequence of 50 ms time bins throughout the trial. At each 50 ms time step, we used population activity patterns to decode the cue, probe, and response on individual trials by constructing a vector of firing rates representing the population activity pattern at that time point. To construct this vector, we matched trials over recording days according to the trial repetition number of each of the decoded task variables (the first A-cue trial for each neuron, for example) and included firing rates in three consecutive 50 ms time bins from each neuron comprising a 150 ms sliding window. We used leave-one-out cross validation to perform the decoding. All trials (except the trial being decoded) were used as the training data to compute the mean (centroid) and covariance matrix of population activity patterns for subsets of trials corresponding to the potential values of the decoded variable (A- and B-cue when decoding cue, for example). We then computed the Euclidean distance between the population activity pattern on the decoded trial in a specific time bin and

the two group centroids computed using neural activity in the same time bin in the training data. That is, we allowed the group centroids defining the population codes for the decoded variable to change over time, and always compared population activity patterns to group centroids computed using training data at the same time point. We then converted the distance between the population activity pattern and the group centroids to posterior probabilities of the trial belonging to each of the two groups under the assumption that the distribution of population activity patterns in each group was multivariate normal. We assigned the trial to the group with the higher posterior probability. Repeating the decoding over time bins produced a time series of posterior probabilities quantifying fluctuations in the strength of population signals encoding specific items of task-defined information.

We included neurons in the populations if they had been recorded for a minimum number of trials in each set (70 trials for balanced and 230 trials for prepotent sets). We then ranked neurons in the two areas according to the significance of the p value defining their group membership, and selected matching numbers of the most significant neurons in both the MD and the PFC. Equalizing neuron numbers and trial numbers removed differences between them as potential factors that would confound comparison between the strength of population encoding in the MD and the PFC. We used a permutation test repeating the decoding analysis after randomly shuffling trials of posterior probability time series between brain areas (1000 permutations). We considered original differences in decoding between brain areas significant if they exceeded the 95th percentile of the bootstrap distribution of differences obtained after shuffling trials.

To determine whether changes in population representation were predictive of behavioral errors, we used the above procedure to decode trial outcome (success/error) on BX trials as a binary variable from population activity patterns (see Fig. 11A). We restricted the analysis to include neurons across all functional groups (Fig. 6, Groups 1–7) with neural activity on seven error trials and seven correct trials. We also expanded the sliding window in this trial outcome decoding analysis to consist of six consecutive 50 ms time bins, comprising a 300 ms sliding window.

To determine whether changes in population representation correlated with the speed of the behavioral response, we ranked all trials for each neuron according to response time (RT) on that trial, and median split the trials into fast and slow RT categories. We then applied the above decoding procedure to decode RT speed as a binary variable (fast/slow) from population activity patterns (see Fig. 12).

Signal transmission analysis. We applied a signal transmission analysis (Crowe et al., 2013) to measure temporal covariation in the information about task variables encoded by groups of simultaneously recorded neurons in MD and PFC. The logic of this analysis is that lagged correlation in coded information is likely to reflect physiological interaction between the groups of neurons. First, we identified subsets of neurons within each ensemble that encoded task variables of interest (as determined by ANCOVA-based classification above). Next, we applied time-resolved pattern classification to decode task variables using activity patterns in the neural subsets. This is analogous to the population decoding above except that neurons were limited to the sets of neurons in the PFC and MD that were recorded simultaneously. The decoding produced two posterior probability time series that capture moment-to-moment fluctuation in the strength of neural signals encoding different task variables in different brain structures. The question we addressed is whether these fluctuations were correlated in time, and if so at what lag, to infer the direction of flow of information between the groups of neurons. To obtain unbiased estimates of correlation between the two time series, we first removed linear trends and autocorrelation in the time series by fitting an ARIMA (autoregressive, integrated moving average) model to the probability time series. ARIMA models capture the variation in each time series explainable by its own prior history. The first step in ARIMA modeling is to difference the time series to remove time trends. Next a regression is performed to predict the value at each time step as a weighted sum of prior values. Finally, additional regression coefficients are obtained to improve the prediction of the value at each time step by including a weighted sum of prior errors in prediction. We used ARIMA models of order [10,2,2], including regressors for the 10 prior values in

the time series, fitting the model after differencing the times series twice, and including regressors to capture the influence of errors in estimation for the prior two values. We then used the time series of residual posterior probabilities after fitting the ARIMA model to evaluate correlation between time series. This reduced the influence of nonstationarity and autocorrelation in the time series on estimates of their correlation. We regressed one residual posterior probability time series onto the other within a sliding window of 500 ms advanced in 50 ms time steps (at a lag of either plus or minus one 50 ms time bin). This produced a time series of F statistics that measured fluctuation in the strength of functional coupling between groups of neurons that encoded task variables of interest in the same area or in different brain areas. When evaluating information transmission within area, the two groups of neurons were exclusive, eliminating the chance that the same neuron contributed to both times series at the final regression step.

Temporal covariation in the information coded by activity patterns in MD and PFC neurons could reflect their common entrainment to external events (such as stimuli or the motor response) rather than physiological interaction between the neurons. To factor out the potential influence of neural entrainment to external events, we conducted a permutation test. We randomly shuffled trials of PFC and MD posterior probability decoding time series and repeated the sliding-window regression using the trial shuffled data (100 iterations). Any temporal correlation between brain areas present in the trial shuffled data would likely reflect entrainment to external events (because the temporal relationship between neural signals and external events was preserved in the trial shuffling), rather than physiological interaction between neurons (because the trial shuffling broke the simultaneity of the neural signals in PFC and MD precluding fluctuations in coded information from reflecting real time physiological interactions between neurons in the 2 areas). We corrected functional coupling time series (see Fig. 13) by subtracting the mean F statistic of the trial-shuffled permutation distribution from the original F statistic at each time point and considered transmission significant at that time point if the original F statistic value exceeded the 95th percentile of the bootstrap distribution. We applied an FDR-corrected method to maintain the overall type 1 error rate at 0.05 (Fujisawa et al., 2008).

Results

Behavioral performance

The DPX task consists of four trial types based on four possible combinations of the cue (A, B) and probe (X, Y) stimuli (AX, AY, BX, BY; Fig. 1A–C) in sequence. Mk1 performed an average of 90% of DPX trials correctly and Mk2 performed an average of 92% of trials correctly. For both monkeys, performance varied as a function of trial type (χ^2 test: Mk1: AX = 90%, AY = 92%, BX = 83% BY = 99%, $\chi^2 = 364.81$, $p < 0.0001$; Mk2: AX = 93%, AY = 96%, BX = 80%, BY = 99%, $\chi^2 = 269.41$, $p < 0.0001$). There was an increased error rate on BX trials in both monkeys, indicating they occasionally released the prepotent target response to the X-probe. However, performance was maintained at 80% correct or greater for each of the trial types (Fig. 1D). RT varied significantly based on the trial type for both monkeys (Fig. 1E; Kruskal–Wallis test; Mk1 $\chi^2 = 1401.61$, $p < 0.0001$, and Mk2 $\chi^2 = 2640.73$, $p < 0.0001$). *Post hoc* testing using multiple comparisons of mean ranks ($p < 0.01$) indicated that for Mk two RTs for A-cue trials were significantly longer than B-cue trials and RT was greatest on AX trials (BX and BY were not significantly different from each other). RTs for Mk1 were significantly different across trial types, with RT being greatest on AY trials.

Neural database

We recorded 48 neural ensembles in the PFC and MD simultaneously, during the performance of one balanced and one prepotent set each. A total of 1105 neurons in the PFC and 1402 neurons in the MD (Mk1: 29 ensembles of 577 PFC neurons and

Figure 3. Numbers of neurons selective for cue, probe, and response in MD and PFC. **A, B**, Venn diagrams depict the numbers of neurons exhibiting significant selectivity (ANCOVA, $p < 0.05$) for the cue (blue), probe (pink), and/or response (purple) in MD (**A**) and PFC (**B**). Area of circles is proportional to neuron number. **C, D**, Bars indicate the number of neurons in the PFC (**C**) and MD (**D**) that exhibit preference for specific cues (A or B), probes (X or Y), and responses (T or N). Asterisks indicate significant biases in preference for specific cue, probes, and responses in the neural population (Z test of proportions: $*p < 0.05$, $**p < 0.001$, $***p < 0.0001$).

887 MD neurons; Mk2: 19 ensembles of 528 PFC neurons and 515 MD neurons) were recorded. This included an average of 29 neurons in the MD (range 16–50) and 23 neurons in the PFC (range 9–52) per ensemble. To target the PFC and MD reliably and reproducibly, we inserted a grid into each recording chamber (Fig. 2*A, B*, small black circles within each chamber indicate grid hole locations). Each recording location (Fig. 2*A, B*, red circles) produced between 1–6 ensembles (average of 2 ensembles per recording day). Points of electrode entry at the cortical surface can be seen by comparing red dots to sulcal anatomy. Electrode angle from these points of entry intersected MD at depth. We confirmed that electrode penetrations successfully reached the MD by obtaining CT images of a metal probe inserted to the depth of the MD during neural recording. We coregistered the CT images visualizing the trajectory of the probe with the presurgical coronal MR images at the level of the MD nucleus (3D reconstructions shown; Fig. 2*C, D*). The location and volume of the MD nucleus was estimated from a nonhuman primate structural brain atlas integrated into the MR targeting system (Fig. 2*C, D*; yellow). We then projected the electrode penetrations from which MD neural activity was recorded onto the dorsal surface of the MD volume from the atlas to indicate the targeted locations in the MD for each electrode penetration. Neural recordings targeted the lateral (parvocellular) portion of the MD, primarily (Fig. 2*E, F*).

Neurons encoding cue, probe, and response in MD and PFC

We applied ANCOVA to identify neurons in which firing rate varied significantly ($p < 0.05$) as a function of the cue (A vs B; during the cue, delay or probe periods; Fig. 3, light blue), the probe (X vs Y; during the probe period; Fig. 3, pink), or the response (target (T) vs nontarget (NT); during the response period; Fig. 3, purple). Across the population of sampled neurons, 71% of PFC neurons and 84% of MD neurons exhibited cue, probe or response-selective activity. The proportion of sampled neurons that encoded task-related information was significantly higher in the MD than the PFC (Z test of proportions; $Z = -7.63$,

$p < 0.0001$). The PFC (86%) and the MD (84%) had similar proportions of neurons that encoded cue information (Fig. 3*A, B*, blue circles; $Z = 1.44$, $p = 0.15$). The proportion of probe-selective neurons were significantly higher in MD (40%; Fig. 2*A, B*, pink circles) than PFC (30%; $Z = -4.20$, $p < 0.0001$). Last, the MD had a significantly higher proportion of response neurons (66%) than the proportion in the PFC (48%; $Z = -7.57$, $p < 0.0001$; Fig. 3*A, B*, purple circles). These data show that relatively more MD than PFC neurons were recruited to encode the probe and response, whereas similar proportions of neurons in the two structures were recruited to encode the cue.

Bar graphs in Figure 3 illustrate the numbers of cue, probe, and response-selective neurons that preferred A versus B-cues, X versus Y-probes, and T versus NT responses in the PFC (Fig. 3*C*) and the MD (3*D*). Cue neurons were split into three groups depending on the trial epoch during which their activity was significantly modulated by the cue. Early cue neurons (selective during the cue period) were biased to prefer B-cues in the PFC ($Z = 2.40$, $p = 0.017$) and the MD ($Z = 10.16$, $p < 0.0001$). This bias switched during the probe period; late cue neurons were biased to prefer A-cues in both areas (PFC: $Z = 6.48$, $p < 0.0001$; MD: $Z = 7.89$, $p < 0.0001$). In the delay period, PFC neurons were biased to prefer A-cues ($Z = 5.35$, $p < 0.0001$), whereas MD neurons were biased to prefer B-cues ($Z = 2.14$, $p = 0.03$). In the PFC probe neurons were significantly biased to prefer Y-probes ($Z = 3.56$, $p = 0.0004$) and response neurons were significantly biased to prefer nontarget responses ($Z = 3.44$, $p = 0.0006$). MD probe and response neurons did not show these biases (probe: $Z = 1.38$, $p = 0.17$; response: $Z = 0.89$, $p = 0.37$).

Individual neuron activity patterns during the DPX task

Next, we examined the diversity of firing patterns of individual neurons by plotting rasters and SDFs of their activity separated by cue-probe sequence (trial type) in PFC (Fig. 4*A–C*) and MD (4*D–F*). Some neurons modulated their firing rate in response to only

PFC**A EARLY B** **B EARLY B and LATE A** **C EARLY B and RESP NT** **MD****D EARLY B and LATE A** **E PROBE X** **F EARLY B and PROBE Y**

Figure 4. Single-neuron activity during DPX task performance. Each row of panels illustrates the activity of a single neuron in PFC (**A–C**) or MD (**D–F**). Individual panels indicate activity on a single trial type defined by cue-probe sequence. Rasters indicate timing of action potentials, spike density functions (solid gray; $\sigma = 40$ ms; 10 Hz/div) show modulations in average firing rate. Red tick marks indicate the time of the response in each trial. **A**, Early B-cue neuron: PFC cell 4623 is an example of a neuron whose firing rate during the cue period is greater on B-cue trials than A-cue trials. **B**, Early B-cue and Late A-cue neuron: PFC cell 5175 is an example of a switch neuron that switches its cue preference from B-cues during the cue period to A-cues during the probe period. **C**, Early B-cue and NT-response neuron: PFC cell 5445 is an example of a neuron whose firing rate is greater on B-cue than A-cue trials during the cue period and is also greater on invalid NT-response than T-response trials during the response period. **D**, Early B-cue and Late A-cue neuron: MD cell 4488 is an example of a ‘switch’ neuron that switches its cue preference from B-cues during the cue period to A-cues during the probe period. **E**, X-probe neuron: MD cell 4586 is an example of a neuron whose firing rate during the probe period is greater on X-probe than Y-probe trials. **F**, Early B-cue and Y-probe neuron: MD Cell 5372 is an example of a neuron whose firing rate during the cue period is greater on B-cue than A-cue trials, and during the probe period is greater on Y-probe than X-probe trials.

one task factor during a single trial period. Other neurons modulated their firing rate in response to multiple task factors during multiple trial periods. For example, the PFC neuron in Figure 4A primarily modulated firing rate during the cue period (B-cue preference), whereas the PFC neuron in Figure 4B modulated firing rate during both the cue period (B-cue preference) and the probe period

(A-cue preference). We refer to neurons exhibiting early (cue period) B-cue preference and late (probe period) A-cue preference as switch neurons, as we reported previously (Blackman et al., 2016). We found that switch neurons exist also in the MD thalamus (Fig. 4D). Other neurons exhibited selectivity for the probe (Fig. 4E,F) and response (4C).

Figure 5. Task-defined population activity patterns on prepotent trial sets. Population SDFs ($\delta = 40$ ms) display average population activity as a function of time on subsets of trials defined by cue-probe sequence (AX, orange; AY, red; BX, purple; BY, blue). Neurons were assigned to populations based on their cue, probe and response preference as well as the trial epoch in which activity was modulated as determined by ANCOVA ($p < 0.05$). Each neuron could belong to multiple populations so defined. The left column of **A–E** illustrates the activity of populations that preferred valid stimuli and responses (A-cue, X-probe, T-response) in MD (left) and PFC (right). The right column of **F–J** illustrates the activity of populations that preferred invalid stimuli and responses (B-cue, Y-probe, NT-response) in MD (left) and PFC (right). All SDFs are aligned to the cue onset. The stimulus or response preference defining each population is illustrated on the left margin of each pair of panels.

Population neural activity patterns during DPX task performance

We first examined population activity of neurons separated according to the cue, probe, or response preference of the neuron in addition to the trial epoch when selectivity was evident as detected by ANCOVA (Fig. 5). This revealed some general trends in population activity. Many ANCOVA-defined neural populations exhibited early B-cue preference, either during the cue period or delay period. This gave over to higher levels of activity during the probe period on A-cue trials, demonstrating the general ‘switch’ pattern (Fig. 5C) in cue preference noted above. In addition, several of the neural populations that exhibited the switch pattern exhibited the overall highest activity during the probe period on AY trials, in particular (Fig. 5F). Generally, the same temporal profiles of neural population activity were present both in the MD and PFC (Fig. 5). Because we did not vary the relationship between A- and B-cues and the cognitive states they signified (to maintain the parallel between the DPX task we used and comparable AX-CPT paradigms typically used in patients with schizophrenia), we cannot be certain that state

signals as encoded by switch neurons in the MD-PFC network did not reflect in part the visual attributes of the stimuli that instructed the cognitive states. However, switch neurons preferred the B-cue and then the A-cue at different times within the same trial, and their activity was further modulated by combinations of cues and probes that together connoted the necessity to inhibit the prepotent response (specifically A-cues and Y-probes). These observations argue that rather than reflecting the visual attributes of cues and probes, switch neurons encoded the cognitive states they implied.

In addition, because many neurons tended to respond to multiple task stimuli at different times in the trial, populations defined by significant ANCOVA results often exhibited similar activity profiles. Therefore, we combined populations with similar activity patterns into functional groups. This allowed for more simplified classification without the loss of information contained within each group, as we have done with previous datasets recorded during DPX performance (Blackman et al., 2016). Neurons were assigned to one of seven non-overlapping functional groups based on the task factor significantly modulat-

Figure 6. Population activity patterns in functionally defined neural groups in prepotent trial sets. Population SDFs ($\delta = 40$ ms) display average population activity patterns on subsets of trials defined by cue-probe sequence (AX, orange; AY, red; BX, purple; BY, blue). Neurons were divided into seven non-overlapping functional groups (see Results) depending on the specific preference for individual cues, probes and responses as well as the trial period that neurons exhibited selective activity as defined by ANCOVA. Population activity in MD (A–G) and PFC (H–N) is shown aligned to cue onset (left) and time of the response (right) of Groups 1–7.

ing their respective activity (cue, probe, or response), the trial period that activity modulation occurred (cue, delay, probe, or response), and the task condition associated with highest activity (A- or B-cue, X- or Y-probe, T- or NT-response).

Group 1 (Fig. 6A, MD, and H, PFC) consisted of neurons with an early A-cue preference. These neurons exhibited a higher activity rate on A-cue than B-cue trials. Interestingly, Group 1 neurons in both the MD and the PFC population exhibited primarily

a suppression of activity compared with the fixation period on B-cue trials.

Group 2 (Fig. 6B, MD, and I, PFC) and Group 3 (Fig. 6C, MD, and J, PFC) populations exhibited the switch in cue preference from B-cues during the cue period to A-cues during the probe period. Notably, probe period activity was better aligned to the onset of the cue (Fig. 6B,C,I,J, left) than the timing of the motor response (right), suggesting that these neurons do not play a direct role in motor control. Collectively switch neurons comprised a large fraction of task-related neurons in both areas. Group 2 was made up of neurons with an early B-cue preference or a late Y-probe preference, and generally exhibited a switch from early B-cue to late AY trial selective activity. Group 3 was made up of neurons with either delay or late A-cue preference. These neurons exhibited attenuated B-cue selective activity during the cue period that switched to A-cue selective activity during the probe period. The PFC population exhibited a decrease in delay period activity relative to baseline on B-cue trials (Fig. 6J), whereas the MD population did not (6C).

Group 4 (Fig. 6D, MD, and K, PFC) neurons were those with a preference for B-cues during the delay and probe periods. The MD and the PFC populations showed strong sustained delay period activity, classically associated with maintenance of information in working memory.

Group 5 (Fig. 6E, MD, and L, PFC) neurons exhibited X-probe preference during the probe period. This group was much more prominent in the MD than the PFC, providing a clear distinction between these areas. Activity in Group 5 neurons was better aligned with the response (Fig. 6E,L, right) than probe onset (left), suggesting a role in response control.

Group 6 (Fig. 6F, MD, and M, PFC) and Group 7 (Fig. 6G, MD, and N, PFC) were made up of neurons exhibiting response selectivity activity for T (Group 6) and NT (Group 7) motor responses, respectively. Neurons exhibiting response-selective activity were much more prominent in MD (Fig. 6F,G) than PFC (6M,N), both in terms of numbers of neurons and strength of signal. This suggests that MD plays a more direct role in response programming and or selection than PFC.

Quantification of the population of switch neurons

Activity in switch neurons was high under stimulus conditions associated with the inhibition of the prepotent target response (during the cue period when B-cues were presented, or during the probe period when the Y-probe followed the A-cue). To quantify the magnitude of the switch in cue preference between cue and probe periods in the MD and PFC, we first quantified the cue preference of single neurons during the cue period (Fig. 7A,B, x-axis) and probe period (y-axis). We identified neurons with switch indices either significantly greater than (Fig. 7, red) or less than (blue) a bootstrap distribution of indices computed after randomly shuffling firing rates between A-cue and B-cue trials ($p < 0.05$, two-tail). Larger switch indices are associated with B- to A-cue switch, smaller indices the converse). Switch neurons with significant switch indices in the direction associated with higher cognitive control load (stronger early B-cue and stronger late A-cue pref-

Figure 7. Quantification of switch in cue preference from the cue to the probe period in single neurons on prepotent trial sets. Each dot indicates the cue preference index ($(A-cue - B-cue)/(A-cue + B-cue)$) of a single neuron during the cue period (x-axis) and probe period (y-axis) in (A) PFC or (B) MD. Colored symbols indicate neurons in which the switch index (probe period cue preference index – cue period cue preference index) was either significantly greater than (red) or less than (blue) the mean of a bootstrap distribution of switch indices computed by randomly permuting firing rates across A-cue and B-cue trials ($p < 0.05$; two-tailed). Gray symbols indicate neurons that encoded the cue during both cue and probe periods but did not have a significant switch index.

ference) were significantly more numerous than neurons that switched cue preference in the opposite (stronger early A-cue and stronger late B-cue preference), both in MD (Fig. 7B; Z test of proportions: $Z = 6.95$, $p < 0.001$; 106 vs 26 of 248 neurons with cue-selective activity during both periods) and in PFC (Fig. 7A; $Z = 6.78$, $p < 0.001$; 70 vs 18 of 133 neurons). There was no significant difference in the proportion of switch neurons in the MD and PFC ($Z = 1.85$, $p = 0.064$).

Influence of cognitive control load on switch neuron activity

To investigate whether modulation of cognitive control load at the behavioral level influenced information processing in the MD-PFC network, we contrasted neural activity on balanced and prepotent trial sets. In balanced trial sets, all four cue-probe sequences occurred with equal probability. In prepotent trial sets, the majority of trials (69%) presented the AX cue-probe sequence, increasing the strength of the target response habit (and the level of cognitive control required to override it). During the cue period, B-cues were associated with high cognitive control load because they countermanded the habitual target response (whereas A-cues established expectancy of the habitual target response). Likewise, during the probe period, following the A-cue, Y-probes were associated with high cognitive control load because they countermanded the expectation of the target response, whereas X-probes were associated with low cognitive control load because they released the habitual target response. We computed the difference in firing rate on B-cue and A-cue trials during the cue period, and AY and AX trials during the probe period, to provide a measure of the strength of neural signals reflecting cognitive control. We then asked whether the strength of these neural signals differed between balanced and prepotent trial sets. We found that the B-cue signal (activity on B-cue minus A-cue trials; Fig. 8, blue) in Group 2 and 3 neurons was significantly larger (permutation test, $p < 0.05$) across the population of neurons on prepotent relative to balanced trial sets in both the PFC (Fig. 8A,C) and the MD (Fig. 8B,D). Similarly, we found that the AY signal (activity on AY minus AX trials) was significantly larger on prepotent relative to balanced trial sets again in both areas (Fig. 8A–D, pink). These data suggest that the strength of the physiological signals reflecting cognitive control at

Figure 8. Modulation of cognitive control neural signals with cognitive control load. Blue shading indicates difference in cue period firing rates on B-cue and A-cue trials. Pink shading indicates difference in probe period firing rates on AY and AX trials. Population activity of Group 2 neurons in PFC (**A**) and MD (**B**) on balanced (left) and prepotent (right) trial sets. Population activity of Group 3 neurons in PFC (**C**) and MD (**D**) on balanced (left) and prepotent (right) trial sets. *Indicates a significant difference in the difference in cue period firing rate between B-cue and A-cue trials on prepotent and balanced trial sets ($p < 0.05$, permutation test). #Indicates a significant difference in the difference in probe period firing rate on AY and AX trials on prepotent and balanced trial sets ($p < 0.05$).

the neural level in both the PFC and MD is augmented when cognitive control demand is increased at the behavioral level.

Timing and strength of single neuron recruitment in MD and PFC

We compared the timing and strength of cue, probe and response signals in MD and PFC neurons at the population level to provide insight into the pattern of information flow between the MD and the PFC during cognitive control processing. For that purpose, we performed a regression analysis in which we regressed single neuron firing rates in the MD and PFC onto the cue (A or B), probe (X or Y), or response (T or NT) over trials within a sliding window of 100 ms advanced in 20 ms steps. We quantify the strength of the relationship between single neuron firing rate and task variables by calculating PEV. We divided trials in half and performed separate regression analyses on the two trial subsets. We ranked neurons according to time to peak PEV using one trial subset and plot PEV as a function of time in the trial using the other trial set (thus we are showing regression results based on

half of the trials; Fig. 9). Heat maps (1 row = 1 neuron) illustrate the time course of PEV values associated with each predictor variable (cue, probe, and response) as a function of time in the trial (Fig. 9A–C). We restricted the PFC and MD populations to equal numbers of significant neurons to compare neural recruitment in the two areas (cue: $n = 290$, probe: $n = 160$, response: $n = 140$ neurons in PFC and MD). The bands of peak PEV values in the heat maps in both PFC (Fig. 9A–C, right) and MD (left) indicate a sequential recruitment of neurons taking place during the trial in both brain areas, with each neuron modulating its firing rate in relation to task variables for a restricted period of time consistent with a temporally dynamic population code (Crowe et al., 2010). Additionally, neurons exhibited more extended periods of firing rate modulation (longer horizontal bands of color in the heat maps) encoding the cue (Fig. 9A) and response (Fig. 9C) in both MD (left) and PFC (right), consistent with persistent activity reflecting the maintenance of task information in working memory. To compare the timing of neural recruitment in MD and PFC, we compared cumulative distributions of the time to peak PEV for each neuron and task variable. Cue signals emerged significantly earlier in the PFC compared with the MD [Fig. 9D; Kolmogorov–Smirnov (KS) test = 0.20, $p < 0.0001$]. Probe signals emerged earlier in the MD than the PFC (Fig. 9E; KS test = 0.48, $p < 0.0001$). Response signals also emerged earlier in the MD than the PFC (Fig. 9F; KS test = 0.51, $p < 0.0001$).

To compare the strength of neural recruitment in MD and PFC, we computed the population average PEV time course for each task variable and evaluated the significance of differences between brain

areas using a permutation test (see Materials and Methods; $p < 0.05$). The population average PEV attributable to the cue (Fig. 9G), probe (9H), and response (9I) was significantly larger in the MD than the PFC during the probe period. There was no significant difference in cue related signals between MD and PFC during the cue period (Fig. 9G). MD contains much stronger information about all task variables during the probe period when it is critical to combine cue and probe information to correctly select the response direction.

Timing and strength of population decoding in MD and PFC

To further elucidate differences in the strength and timing of neural signals between the MD and PFC, we conducted a time-resolved decoding analysis to quantify the information about task variables encoded by patterns of neural activity in the population. We decoded cue and probe from neurons in Groups 2 and 3 (switch neurons) and the response from neurons in Groups 6 and 7 (response neurons).

Figure 9. PEV in firing rate over trials as a function of time in prepotent trial sets. Warmer colors in the heat maps represent greater PEV values associated with the cue (**A**), probe (**B**), and response (**C**). Trials were divided in half and separate regression analyses applied to the trial subsets. Neurons were ranked according to the time to peak PEV based on the regression analysis applied to one trial subset. PEV results from the regression analysis applied to the second trial subset are shown. Regression results are restricted to equal numbers of significant neurons in MD and PFC to facilitate comparison of neural recruitment between brain areas. PEV values were obtained by regressing the firing rate of each neuron onto the appropriate task factor over trials within a sliding 100 ms window (advanced in 20 ms steps). MD data are on the left and PFC data on the right. Cumulative distributions of time to peak PEV attributable to the cue (**D**), probe (**E**), and response (**F**) in MD neurons (blue) and PFC neurons (red). Population average PEV attributable to the cue (**G**), probe (**H**), and response (**I**) in MD neurons (blue) and PFC neurons (red). Significant differences in population average PEV time courses are indicated with circles (permutation test, $p < 0.05$).

In prepotent sets, cue decoding posterior probabilities rose in the MD and PFC in parallel but were sustained at significantly higher levels during the cue period in PFC, particularly on B-cue trials (Fig. 10A; $p < 0.05$, permutation test). During the probe period, the converse was true; cue posterior probability was significantly greater in MD than in PFC (Fig. 10A), perhaps reflecting the stronger late-cue signals in MD (Fig. 6, compare C, J). Similarly, probe posterior probabilities rose for the MD and PFC in parallel during the probe period (Fig. 10B). Last, response posterior probabilities were significantly greater in MD than PFC (Fig. 10C), rising first to an intermediate level during the cue period (because the cue carried partial information about the direction of the response) and rising further after the probe was presented (which determined response direction definitively). This is further support for the idea that the MD is more involved in response selection than the PFC.

Error trial analysis

To determine the extent to which neural activity in MD and PFC predicted success or failure in the task, we decoded the outcome of each trial (correct/error) from neural activity restricted to seven error and seven correct BX trials (Fig. 11) and neurons in all functional neural groups (Groups 1–7). Trial outcome was significantly better represented by population activity patterns in MD than PFC (Fig. 11A). To examine changes in neural encoding that could lead to errors at the single neuron level, we compared results of the sliding window regression analysis applied to neural

activity on correct and error trials. Cue PEV was profoundly suppressed on error trials comparably in both in MD and PFC neurons (Fig. 11B), suggesting that errors were associated with weaker encoding of the cue stimulus providing contextual information for the response. In contrast, error trials were associated with comparably enhanced response PEV in MD and PFC neurons (Fig. 11C), maintaining significantly greater response PEV in PFC relative to MD on error trials.

Response time decoding

To investigate the degree to which neural activity predicted the timing of the motor response, we ranked trials according to RT, median split the data into fast and slow trial categories, and decoded RT (fast/slow) from neural activity patterns. The posterior probability associated with RT decoding was significantly greater when using population activity patterns in MD compared with PFC, particularly at the time of the motor response (Fig. 12).

Functional coupling between neurons in MD and PFC

We next measured functional coupling between groups of neurons in MD and PFC to evaluate communication in thalamocortical networks during cognitive control. For that purpose, we decoded the cue, probe and response from neural activity patterns in a sequence of 50 ms time bins. The resulting time series of posterior probabilities captured fluctuations in the strength of the neural representation of the cue, probe and response. To evaluate functional coupling between PFC and MD signals, we

Figure 10. Time-resolved population decoding of cue, probe and response in MD and PFC. Functions plot the population average posterior probability associated with the neural representation of the cue (**A**), probe (**B**), and response (**C**) as a function of time in the trial, based on neural population activity of neurons in Groups 2 and 3 or Groups 6 and 7 in MD (blue) and PFC (red). Neural data were recorded on prepotent trial sets and trial numbers were equalized across brain regions (Cue = 256 neurons on 272 trials each, probe = 256 neurons on 279 trials each, response = 114 neurons on 271 trials each). Left, The posterior probability averaged over all trials; center and right, posterior probability averaged over the subsets of trials as indicated. Circles indicate significant differences in decoding strength between areas ($p < 0.05$, MD > PFC, blue; PFC > MD, red; permutation test, FDR-corrected; see Materials and Methods).

measured temporal covariation in the posterior probability time series associated with the neural representation of specific task variables at different lags, after correcting for autocorrelation in the time series (that can spuriously inflate estimates of the correlation between 2 time series), and covariation in the two posterior probability time series that could reflect common entrainment to external sensory or motor events (see Materials and Methods). We found that the most prominent functional coupling between MD and PFC took place during the probe period, between neurons encoding the probe (Fig. 13B), when information about the cue and probe stimuli had to be integrated to determine the motor response. During this period, both corticothalamic (Fig. 13B, red; PFC signals predicting MD signals at a lag of one 50 ms time bin) and thalamocortical (13B, blue; MD signals predicting PFC signals) transmission of probe signals was evident, suggesting bidirectional communication between the two structures. Lagged covariation in probe posterior probabilities in both corticothalamic and thalamocortical directions exceeded the 95th percentile of transmission values obtained after trial-shuffling the probe posterior probability time series (Fig. 13B; dots at the top indicate time points where transmission exceeded the bootstrap distribution at $p < 0.05$). This provided evidence that covariation in probe signals between MD and PFC exceeded the level that could be attributed to common entrainment to external input (e.g., onset of the probe stimulus) and was likely therefore to derive from physiological interaction between the structures. Weaker

and less often significant simultaneous covariation of probe signals was also observed, perhaps reflecting the limited ability of trial shuffling to disrupt simultaneous signal covariation (Fig. 13B, orange). Significant transmission of cue (Fig. 13A) and response (13C) signals was also observed to occur between MD and PFC, which was predominantly in the corticothalamic direction during the probe period (13A, C, red).

Discussion

Prior research has documented that the MD-PFC network malfunctions during cognitive control in schizophrenia (Minzenberg et al., 2009; Welsh et al., 2010; Woodward et al., 2012; Anticevic et al., 2014a,b; Woodward and Heckers, 2016), but how neurons in this network function to carry out the computations that mediate cognitive control is unknown. To address this, we translated a task measuring cognitive control deficits in patients directly to monkeys and recorded neural activity in both structures concurrently during task performance. We provide evidence that although neural signals relating to cognitive control are distributed across the MD-PFC network, each structure is specialized in terms of the kinds of neural signals it generates during the DPX task that indicate distinctive roles in cognitive control.

Most notably, MD neurons play a more direct role in the computations required for action selection than PFC neurons. In the DPX task, information about the remembered cue and the visible probe stimulus had to be integrated specifically during the probe period (when all this information is available) to decide whether to execute the target or nontarget motor response at the end of the trial. MD neurons, compared with PFC neurons, were more strongly activated during the probe period when this decision was made. A large phasic increase in firing rate during the probe period was a prominent feature of population activity in nearly all functional groups of MD neurons (Fig. 6A–G), whereas the phasic neural response during the probe period was attenuated in PFC population activity (6H–N). The activation of MD neurons during the probe period carried more information compared with PFC neurons about the identity of the cue stimulus (Figs. 9G, 10A), the identity of the probe stimulus (9H), and the direction of the forthcoming motor response (9I, 10C). This indicates that MD generated a stronger neural representation not only of the motor response but also the sensory information on which that response was based at the time that the decision about which way to move was made. Neural signals encoding the probe and response emerged on average earlier in MD than PFC as well (Fig. 9E,F). Additionally, neural activity in MD better predicted both trial outcome (correct/error; Fig. 11A), and response time (fast/slow; Fig. 12) on a trial-by-trial basis compared with neural activity in PFC. These findings provide evidence, somewhat surprisingly, that integration of sensory signals leading to a motor decision is most strongly represented by neural signals in MD. In addition, whereas the population representation of stimuli and responses in the PFC was strongly biased toward Y-probes and nontarget responses (Fig. 3C) and therefore encoding the infrequent occurrence when the habitual response had to be overridden, MD neurons did not show this bias (Fig. 3D), but more equally encoded both X- and Y-probe stimuli (Fig. 5D, I) as well as target and nontarget responses (Fig. 5E, J). This suggests that MD may play a more direct role in mediating the competition between target and nontarget responses because both were more strongly encoded there compared with PFC. Finally, the strongest functional coupling between neurons in MD and PFC involved the transmission of neural signals encoding the probe stimulus

and occurred during the probe period (Fig. 13B). During this period, information about the probe had to be combined with prior information about state (A-cue vs B-cue) to decide which way to move (Fig. 13B). Transmission of probe information between MD and PFC arose simultaneously in both the corticothalamic direction (with PFC probe signals predicting MD probe signals one 50 ms time bin later) and the thalamocortical direction (MD probe signals predicting PFC probe signals) during the probe period, without clear evidence of transmission in one direction leading the other (Fig. 13B). Timing differences in the initiation of transmission in the two directions may have been too small to detect with population decoding applied to activity patterns in 50 ms time bins (at the single neuron level, both probe and response signals emerged significantly earlier in MD than PFC; Fig. 9E,F). Alternatively, bidirectional transmission and reciprocal interactions between MD and PFC could reflect the fact that the DPX task was relatively easy and well learned by the monkeys, as we have observed that neural dynamics in PFC networks vary as a function of task difficulty in other behavioral contexts, with top-down transmission predominating when task conditions are particularly difficult (Crowe et al., 2013). However, the transmission data are informative about which neural signals were transmitted between structures and when in the trial. The observation that the strongest transmission between MD and PFC involved neural signals encoding the probe at the time that probe information was integrated with state information to determine the response direction, provides convergent evidence that the MD-PFC network is engaged particularly intensely around the time of the decision required by the task.

Neural recording studies in primate MD and PFC during performance of the oculomotor delayed response task have provided consistent evidence of a stronger representation of the motor response by MD compared with PFC neurons (Sommer and Wurtz, 2004a,b; Watanabe and Funahashi, 2004a, 2018; Tanibuchi et al., 2009b; Watanabe et al., 2009). More MD neurons exhibit spatially tuned pre-saccadic activity and MD neurons transform working memory signals from cue-based to saccade-based spatial coordinates earlier in the delay period than PFC neurons (Watanabe and Funahashi, 2004a, 2018; Watanabe et al., 2009). MD neurons that receive input from the superior colliculus and project to the PFC are enriched in signals encoding the saccadic response and so may convey a corollary discharge of saccade commands from the colliculus to the PFC (Sommer and Wurtz, 2004a,b). Similarly, in a Go/No-Go task MD neurons that projected to the PFC encoded both stimulus and motor response information (Tanibuchi et al., 2009b). Anatomical inputs to the MD further suggest a preferential role in motor processing. In addition to inputs from the PFC, the MD also receives anatomical inputs directly from the basal ganglia, specifically the globus pallidus and the substantia nigra (Alexander et al., 1986; Haber and McFarland, 2001; Haber and Calzavara, 2009). This basal ganglia projection to the MD has been considered part of more extensive basal ganglia-thalamus-cortical loop circuits involved in response selection (Haber and McFarland, 2001; Tanibuchi et al., 2009b). It is possible that input from the basal ganglia to the MD conveys information regarding the habitual target response in the

Figure 11. Neural activity on error trials in MD and PFC. **A**, Functions plot the population average posterior probability associated with decoding trial outcome (correct/error) from population activity in MD (blue) and PFC (red). Blue circles indicate time points where the posterior probability was significantly greater in MD than PFC ($p < 0.05$; FDR-corrected). All neuronal groups (1–7) were included in the population decoding. Neural data are from prepotent trial sets and trial numbers were equated between MD and PFC (235 neurons per region on 14 trials, 7 correct and 7 error trials, each). Average population PEV attributable to **(B)** the cue, and **(C)** the response in MD (blue) and PFC (red). Blue circles indicate time points where the population PEV was significantly greater in MD than PFC ($p < 0.05$; FDR-corrected). All trials were used in the regression analysis. Results are plotted separately for correct trials (thin PEV functions) and error trials (thick PEV functions).

RESPONSE TIME DECODING

Figure 12. Decoding RT (fast/slow) from neural activity. RTs were median split to divide trials into fast and slow RTs. Functions plot the population average posterior probability obtained by decoding RT (fast/slow) from population activity patterns in MD (blue) and PFC (red). Time points at which RT posterior probability was significantly higher in MD than PFC are indicated by gray bars at the top of the figure ($p < 0.05$; FDR-corrected).

DPX task, contributing to the stronger representation of the target response in MD relative to PFC, as the basal ganglia have been implicated to play a central role in habit formation (Yin and Knowlton, 2006). However, whereas these data and our own are consistent in terms of indicating a preferential role in motor programming, our data provide new insight into the role of MD in the preceding decision-making process. We show a preferential engagement of the MD to encode stimuli on which the decision is based, in addition to the motor response itself, and that communication between MD and PFC is maximal at the time of decision processing. Therefore, the MD may be a node in a distributed response selection network where information about habitual responses from the basal ganglia and state-based responses from the PFC converge to mediate the competition between habit and cognitive control.

Whereas decision and response information processing were more strongly represented in the MD, state information was more strongly represented in the PFC. Specifically, the activity of switch neurons, which appear to encode an environmental state variable corresponding to the necessity to override the habitual response as signaled by different cue stimuli at different times in the trial, emerges first (Fig. 9D) and is stronger in the cue period (Fig. 10A, right) in PFC than MD. This suggests that PFC is preferentially engaged for state representation and MD for decision-making and response selection. Nonetheless, these data show MD participates in state representation, demonstrating a role in cog-

Figure 13. Functional coupling between neural signals encoding the cue, probe and response in PFC and MD. Functional coupling between (A) cue, (B) probe, and (C) response decoding time series in MD and PFC. Functional coupling is defined as temporal covariation in the posterior probability time series obtained when decoding the cue, probe and response from patterns of neural ensemble activity recorded simultaneously in the PFC and MD. Corticothalamic coupling (red) indicates the PFC decoding time series predicted the MD decoding time series one 50 ms time bin later. Thalamocortical coupling (blue) indicates the converse temporal relationship. Simultaneous covariation in posterior probabilities is indicated in orange. Coupling strength is expressed by the F statistic (y -axis) obtained when regressing the posterior probability time series in one area onto the posterior probability time series in the other within a sliding window at the lag indicated. The autocorrelation structure of the time series was estimated (by fitting ARIMA models), and coupling evaluated using the residual posterior probability values. This limited analysis of functional coupling between brain areas to fluctuations in the posterior probability time series that did not reflect their own history (autocorrelation), and therefore could reflect physiological communication between brain areas. To further isolate this coupling from covariation in PFC and MD neural signals that could result from their common entrainment to external events, we generated a permutation distribution of coupling functions after randomly shuffling trials of PFC and MD probability time series. Trial shuffling broke the simultaneity of the neural signals but maintained the temporal relationship of neural signals to external events. Illustrated transmission functions were bootstrap corrected by subtracting the mean F statistic in the permutation distribution at each time step. Dots at the top indicate time points at which the original F statistic exceeded the 95th percentile of the bootstrap distribution ($p < 0.05$; FDR-corrected). Numbers of ensembles (nEns) contributing to each transmission analysis are shown below each panel.

nitive control processing in parallel with PFC (Blackman et al., 2016).

Our data provide the first cell-level evidence that MD neurons are directly involved in state representation and decision-making for cognitive control. The MD has been implicated in behavioral flexibility by lesion studies in rodents (Block et al., 2007; Parnaudeau et al., 2013, 2015) and humans (Van der Werf et al., 2000, 2003) demonstrating that MD lesions cause an increase in perseverative behavior during tasks that require shifting of behavioral strategies. Lesions in monkeys have indicated dysfunctional strategy adaptation in reversal learning tasks (Chudasama et al., 2001; Chakraborty et al., 2016). However, these studies do not address the underlying neural mechanisms. Prior neurophysiological studies have focused on the role of MD neurons in working memory, by showing MD neurons, like their PFC counterparts, exhibit persistent delay period activity associated with the short-term storage of task-specific information (Fuster and Alexander, 1973; Tanibuchi and Goldman-Rakic, 2003, 2005; Watanabe and Tanahashi, 2004a,b, 2018; Tanibuchi et al., 2009a,b; Watanabe et al., 2009) and by showing that working memory signals in the two areas are reciprocally interdependent during the delay period (Alexander and Fuster, 1973; Parnaudeau et al., 2013; Bolkan et al., 2017; Schmitt et al., 2017). Although working memory signals were present in our data (Fig. 6D,K), we found several new forms of neural activity reflecting state and decision processes as well, and further that neural communication between MD and PFC was strongest during the probe period, when the response decision was made rather than the delay period (Fig. 11).

Conclusion

A key aspect of cognitive control is the ability to adjust one's response to a stimulus based on contextual information and to inhibit perseverative responses within tasks that produce/exploit habitual responding patterns, a capability that is selectively

disrupted in schizophrenia. MD is structurally abnormal in schizophrenia (Pakkenberg, 1992; Popken et al., 2000; Young et al., 2000; Dorph-Petersen and Lewis, 2017). Moreover, decreased activation of MD is correlated with decreased activation of PFC in patients during cognitive control tasks (Minzenberg et al., 2009). Functional coupling between MD and PFC is reduced in patients (Welsh et al., 2010; Woodward et al., 2012; Anticevic et al., 2014a,b) and improvements in MD-PFC functional coupling following cognitive remediation therapy correlate with the extent of cognitive improvement (Ramsay et al., 2017). Our data provide the first evidence that MD and PFC neurons provide unique contributions to distributed processing for cognitive control with the MD specialized for decision-making and response selection and the PFC specialized for state representation. These data lead to testable hypotheses about which of these computations mediated by the MD-PFC network may be most disrupted in schizophrenia leading to cognitive control failure in patients.

References

- Alexander GE, Fuster JM (1973) Effects of cooling prefrontal cortex on cell firing in the nucleus medialis dorsalis. *Brain Res* 61:93–105.
- Alexander GE, DeLong MR, Strick PL (1986) Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annu Rev Neurosci* 9:357–381.
- Anticevic A, Yang G, Savic A, Murray JD, Cole MW, Repovs G, Pearlson GD, Glahn DC (2014a) Mediodorsal and visual thalamic connectivity differ in schizophrenia and bipolar disorder with and without psychosis history. *Schizophr Bull* 40:1227–1243.
- Anticevic A, Cole MW, Repovs G, Murray JD, Brumbaugh MS, Winkler AM, Savic A, Krystal JH, Pearlson GD, Glahn DC (2014b) Characterizing thalamo-cortical disturbances in schizophrenia and bipolar illness. *Cereb Cortex* 24:3116–3130.
- Asaad WF, Rainer G, Miller EK (2000) Task-specific neural activity in the primate prefrontal cortex. *J Neurophysiol* 84:451–459.
- Blackman RK, Macdonald AW 3rd, Chafee MV (2013) Effects of ketamine on context-processing performance in monkeys: a new animal model of cognitive deficits in schizophrenia. *Neuropsychopharmacology* 38:2090–2100.
- Blackman RK, Crowe DA, DeNicola AL, Sakellaridis S, MacDonald AW 3rd, Chafee MV (2016) Monkey prefrontal neurons reflect logical operations for cognitive control in a variant of the AX continuous performance task (AX-CPT). *J Neurosci* 36:4067–4079.
- Block AE, Dhanji H, Thompson-Tardif SF, Floresco SB (2007) Thalamic-prefrontal cortical-ventral striatal circuitry mediates dissociable components of strategy set shifting. *Cereb Cortex* 17:1625–1636.
- Bolkan SS, Stujenske JM, Parnaudeau S, Spellman TJ, Rauffenbart C, Abbas AI, Harris AZ, Gordon JA, Kellendonk C (2017) Thalamic projections sustain prefrontal activity during working memory maintenance. *Nat Neurosci* 20:987–996.
- Brincat SL, Miller EK (2016) Prefrontal cortex networks shift from external to internal modes during learning. *J Neurosci* 36:9739–9754.
- Buschman TJ, Denovellis EL, Diogo C, Bullock D, Miller EK (2012) Synchronous oscillatory neural ensembles for rules in the prefrontal cortex. *Neuron* 76:838–846.
- Chakraborty S, Kolling N, Walton ME, Mitchell AS (2016) Critical role for the mediodorsal thalamus in permitting rapid reward-guided updating in stochastic reward environments. *eLife* 5:e13588.

- Christoff K, Keramiatian K, Gordon AM, Smith R, Madler B (2009) Prefrontal organization of cognitive control according to levels of abstraction. *Brain Res* 1286:94–105.
- Chudasama Y, Bussey TJ, Muir JL (2001) Effects of selective thalamic and prelimbic cortex lesions on two types of visual discrimination and reversal learning. *Eur J Neurosci* 14:1009–1020.
- Crowe DA, Averbeck BB, Chafee MV (2010) Rapid sequences of population activity patterns dynamically encode task-critical spatial information in parietal cortex. *J Neurosci* 30:11640–11653.
- Crowe DA, Goodwin SJ, Blackman RK, Sakellaridi S, Sponheim SR, MacDonald AW 3rd, Chafee MV (2013) Prefrontal neurons transmit signals to parietal neurons that reflect executive control of cognition. *Nat Neurosci* 16:1484–1491.
- Crowe DA, Zarco W, Bartolo R, Merchant H (2014) Dynamic representation of the temporal and sequential structure of rhythmic movements in the primate medial premotor cortex. *J Neurosci* 34:11972–11983.
- Dorph-Petersen KA, Lewis DA (2017) Postmortem structural studies of the thalamus in schizophrenia. *Schizophr Res* 180:28–35.
- Fujisawa S, Amarasingham A, Harrison MT, Buzsaki G (2008) Behavior-dependent short-term assembly dynamics in the medial prefrontal cortex. *Nat Neurosci* 11:823–833.
- Fuster JM, Alexander GE (1973) Firing changes in cells of the nucleus medialis dorsalis associated with delayed response behavior. *Brain Res* 61:79–91.
- Goldman-Rakic PS (1987) Motor control function of the prefrontal cortex. *Ciba Found Symp* 132:187–200.
- Goldman-Rakic PS, Porrino LJ (1985) The primate mediodorsal (MD) nucleus and its projection to the frontal lobe. *J Comp Neurol* 242:535–560.
- Goodwin SJ, Blackman RK, Sakellaridi S, Chafee MV (2012) Executive control over cognition: stronger and earlier rule-based modulation of spatial category signals in prefrontal cortex relative to parietal cortex. *J Neurosci* 32:3499–3515.
- Haber S, McFarland NR (2001) The place of the thalamus in frontal cortical-basal ganglia circuits. *Neuroscientist* 7:315–324.
- Haber SN, Calzavara R (2009) The cortico-basal ganglia integrative network: the role of the thalamus. *Brain Res Bull* 78:69–74.
- Henderson D, Poppe AB, Barch DM, Carter CS, Gold JM, Ragland JD, Silverstein SM, Strauss ME, MacDonald AW 3rd (2012) Optimization of a goal maintenance task for use in clinical applications. *Schizophr Bull* 38:104–113.
- Isseroff A, Rosvold HE, Galkin TW, Goldman-Rakic PS (1982) Spatial memory impairments following damage to the mediodorsal nucleus of the thalamus in rhesus monkeys. *Brain Res* 232:97–113.
- Johnson RA, Wichern DW (1998) Applied multivariate statistical analysis. Saddle River, NJ: Prentice Hall.
- Jones JA, Sponheim SR, MacDonald AW 3rd (2010) The dot pattern expectancy task: reliability and replication of deficits in schizophrenia. *Psychol Assess* 22:131–141.
- Klecka KR (1980) Discriminant analysis. Newbury Park, CA: Sage.
- MacDonald AW, Goghari VM, Hicks BM, Flory JD, Carter CS, Manuck SB (2005a) A convergent-divergent approach to context processing, general intellectual functioning, and the genetic liability to schizophrenia. *Neuropsychology* 19:814–821.
- MacDonald AW 3rd, Carter CS, Kerns JG, Ursu S, Barch DM, Holmes AJ, Stenger VA, Cohen JD (2005b) Specificity of prefrontal dysfunction and context processing deficits to schizophrenia in never-medicated patients with first-episode psychosis. *Am J Psychiatry* 162:475–484.
- Mante V, Sussillo D, Shenoy KV, Newsome WT (2013) Context-dependent computation by recurrent dynamics in prefrontal cortex. *Nature* 503:78–84.
- Miller EK, Cohen JD (2001) An integrative theory of prefrontal cortex function. *Annu Rev Neurosci* 24:167–202.
- Minzenberg MJ, Laird AR, Thelen S, Carter CS, Glahn DC (2009) Meta-analysis of 41 functional neuroimaging studies of executive function in schizophrenia. *Arch Gen Psychiatry* 66:811–822.
- Miocinovic S, Noecker AM, Maks CB, Butson CR, McIntyre CC (2007) Cicerone: stereotactic neurophysiological recording and deep brain stimulation electrode placement software system. *Acta Neurochir Suppl* 97:561–567.
- Mitchell AS, Chakraborty S (2013) What does the mediodorsal thalamus do? *Front Syst Neurosci* 7:37.
- Olejnik S, Algina J (2003) Generalized eta and omega squared statistics: measures of effect size for some common research designs. *Psychol Methods* 8:434–447.
- Pachitariu M, Steinmetz NA, Caradini M, Harris KD (2016) Fast and accurate spike sorting of high channel count probes with KiloSort. *Adv Neural Inf Process Syst* 29:4455–4463.
- Pakkenberg B (1992) The volume of the mediodorsal thalamic nucleus in treated and untreated schizophrenics. *Schizophr Res* 7:95–100.
- Parker A, Eacott MJ, Gaffan D (1997) The recognition memory deficit caused by mediodorsal thalamic lesion in non-human primates: a comparison with rhinal cortex lesion. *Eur J Neurosci* 9:2423–2431.
- Parnaudeau S, O'Neill PK, Bolkan SS, Ward RD, Abbas AI, Roth BL, Balsam PD, Gordon JA, Kellendonk C (2013) Inhibition of mediodorsal thalamus disrupts thalamofrontal connectivity and cognition. *Neuron* 77:1151–1162.
- Parnaudeau S, Taylor K, Bolkan SS, Ward RD, Balsam PD, Kellendonk C (2015) Mediodorsal thalamus hypofunction impairs flexible goal-directed behavior. *Biol Psychiatry* 77:445–453.
- Parnaudeau S, Bolkan SS, Kellendonk C (2018) The mediodorsal thalamus: an essential partner of the prefrontal cortex for cognition. *Biol Psychiatry* 83:648–656.
- Popken GJ, Bunney WE Jr, Potkin SG, Jones EG (2000) Subnucleus-specific loss of neurons in medial thalamus of schizophrenics. *Proc Natl Acad Sci U S A* 97:9276–9280.
- Ramsay IS, Nienow TM, MacDonald AW 3rd (2017) Increases in intrinsic thalamocortical connectivity and overall cognition following cognitive remediation in chronic schizophrenia. *Biol Psychiatry Cogn Neurosci Neuroimaging* 2:355–362.
- Roy JE, Riesenhuber M, Poggio T, Miller EK (2010) Prefrontal cortex activity during flexible categorization. *J Neurosci* 30:8519–8528.
- Schiffer AM, Waszak F, Yeung N (2015) The role of prediction and outcomes in adaptive cognitive control. *J Physiol Paris* 109:38–52.
- Schmitt LI, Wimmer RD, Nakajima M, Happ M, Mofakham S, Halassa MM (2017) Thalamic amplification of cortical connectivity sustains attentional control. *Nature* 545:219–223.
- Sommer MA, Wurtz RH (2002) A pathway in primate brain for internal monitoring of movements. *Science* 296:1480–1482.
- Sommer MA, Wurtz RH (2004a) What the brain stem tells the frontal cortex: I. Oculomotor signals sent from superior colliculus to frontal eye field via thalamic relay nucleus. *J Neurophysiol* 91:1381–1402.
- Sommer MA, Wurtz RH (2004b) What the brain stem tells the frontal cortex: II. Role of the SC-MD-FEF pathway in corollary discharge. *J Neurophysiol* 91:1403–1423.
- Tanibuchi I, Goldman-Rakic PS (2003) Dissociation of spatial-, object-, and sound-coding neurons in the mediodorsal nucleus of the primate thalamus. *J Neurophysiol* 89:1067–1077.
- Tanibuchi I, Goldman-Rakic PS (2005) Comparison of oculomotor neuronal activity in paralamina and mediodorsal thalamus in the rhesus monkey. *J Neurophysiol* 93:614–619.
- Tanibuchi I, Kitano H, Jinnai K (2009a) Substantia nigra output to prefrontal cortex via thalamus in monkeys: I. Electrophysiological identification of thalamic relay neurons. *J Neurophysiol* 102:2933–2945.
- Tanibuchi I, Kitano H, Jinnai K (2009b) Substantia nigra output to prefrontal cortex via thalamus in monkeys: II. Activity of thalamic relay neurons in delayed conditional go/no-go discrimination task. *J Neurophysiol* 102:2946–2954.
- Van der Werf YD, Witter MP, Uylings HB, Jolles J (2000) Neuropsychology of infarctions in the thalamus: a review. *Neuropsychologia* 38:613–627.
- Van der Werf YD, Scheltens P, Lindeboom J, Witter MP, Uylings HB, Jolles J (2003) Deficits of memory, executive functioning and attention following infarction in the thalamus; a study of 22 cases with localised lesions. *Neuropsychologia* 41:1330–1344.
- Wallis JD, Anderson KC, Miller EK (2001) Single neurons in prefrontal cortex encode abstract rules. *Nature* 411:953–956.
- Watanabe Y, Funahashi S (2004a) Neuronal activity throughout the primate mediodorsal nucleus of the thalamus during oculomotor delayed-responses: I. Cue-, delay-, and response-period activity. *J Neurophysiol* 92:1738–1755.
- Watanabe Y, Funahashi S (2004b) Neuronal activity throughout the primate mediodorsal nucleus of the thalamus during oculomotor delayed-responses: II. Activity encoding visual versus motor signal. *J Neurophysiol* 92:1756–1769.
- Watanabe Y, Funahashi S (2012) Thalamic mediodorsal nucleus and working memory. *Neurosci Biobehav Rev* 36:134–142.

- Watanabe Y, Funahashi S (2018) Change of information represented by thalamic mediodorsal neurons during the delay period. *Neuroreport* 29:466–471.
- Watanabe Y, Takeda K, Funahashi S (2009) Population vector analysis of primate mediodorsal thalamic activity during oculomotor delayed-response performance. *Cereb Cortex* 19:1313–1321.
- Welsh RC, Chen AC, Taylor SF (2010) Low-frequency BOLD fluctuations demonstrate altered thalamocortical connectivity in schizophrenia. *Schizophr Bull* 36:713–722.
- Woodward ND, Heckers S (2016) Mapping thalamocortical functional connectivity in chronic and early stages of psychotic disorders. *Biol Psychiatry* 79:1016–1025.
- Woodward ND, Karbasforoushan H, Heckers S (2012) Thalamocortical dysconnectivity in schizophrenia. *Am J Psychiatry* 169:1092–1099.
- Xiao D, Zikopoulos B, Barbas H (2009) Laminar and modular organization of prefrontal projections to multiple thalamic nuclei. *Neuroscience* 161:1067–1081.
- Yin HH, Knowlton BJ (2006) The role of the basal ganglia in habit formation. *Nat Rev Neurosci* 7:464–476.
- Young KA, Manaye KF, Liang C, Hicks PB, German DC (2000) Reduced number of mediodorsal and anterior thalamic neurons in schizophrenia. *Biol Psychiatry* 47:944–953.
- Zick JL, Blackman RK, Crowe DA, Amirikian B, DeNicola AL, Netoff TI, Chafee MV (2018) Blocking NMDAR disrupts spike timing and decouples monkey prefrontal circuits: implications for activity-dependent disconnection in schizophrenia. *Neuron* 98:1243–1255.e5.
- Zola-Morgan S, Squire LR (1985) Amnesia in monkeys after lesions of the mediodorsal nucleus of the thalamus. *Ann Neurol* 17:558–564.