Prefrontal neurons transmit signals to parietal neurons that reflect executive control of cognition

David A Crowe[1–3], Shikha J Goodwin[1,2], Rachael K Blackman[1,2,4], Sofia Sakellaridi[2,5], Scott R Sponheim[2,6], Angus W MacDonald III[5–7] & Matthew V Chafee[1,2,5]

Prefrontal cortex influences behavior largely through its connections with other association cortices; however, the nature of the information conveyed by prefrontal output signals and what effect these signals have on computations performed by target structures is largely unknown. To address these questions, we simultaneously recorded the activity of neurons in prefrontal and posterior parietal cortices of monkeys performing a rule-based spatial categorization task. Parietal cortex receives direct prefrontal input, and parietal neurons, like their prefrontal counterparts, exhibit signals that reflect rule-based cognitive processing in this task. By analyzing rapid fluctuations in the cognitive information encoded by activity in the two areas, we obtained evidence that signals reflecting rule-dependent categories were selectively transmitted in a top-down direction from prefrontal to parietal neurons, suggesting that prefrontal output is important for the executive control of distributed cognitive processing.

Prefrontal cortex is anatomically situated at the center of a complex array of projections that link it to other cortical association areas, one of which is the posterior parietal cortex[1–4]. Understanding what information is conveyed by the physiological signals prefrontal cortex transmits to other cortical areas is likely to hold the key to understanding the function of prefrontal cortex, as these signals are the primary mechanism by which this area can influence distributed information processing and behavior. At present we know relatively little about the nature of the behavioral information that is coded and transmitted by prefrontal output signals and hence what role prefrontal output may play in shaping computations that take place in the numerous cortical association areas that receive prefrontal input. We approached that question by simultaneously recording the activity of small groups of neurons in regions of the dorsolateral prefrontal and posterior parietal cortex that are anatomically interconnected[3] and that are jointly engaged to process visuospatial information[5,6]. Perhaps as a consequence of the corticocortical pathways that link them, neurons in prefrontal and parietal cortex exhibit parallel changes in firing rate in relation to behavioral events and therefore encode very similar types of behavioral information. For example, neurons located in both prefrontal and parietal cortex cooperatively sustain distributed representations of space to support working memory[7,8], motor planning[9,10] and attention[11] while jointly encoding more abstract cognitive variables as well, such as number[12,13], proportion[14], length[15], spatial category[16,17] and rule-dependent spatial category[18]. That suggests that neural representations generated by either prefrontal or parietal cortex are mediated by patterns of neural activity that are distributed between them, raising the question of the neural basis of localized function in cortical association networks, if it is not localized patterns of neural activity.

One possibility is cortical areas that are linked together in the same distributed network generate distinct neural signals that are rapidly transmitted to other areas along corticocortical pathways. In that case, neural signals of local and remote origin would intermingle within each cortical area. Determining the direction in which signals are transmitted between cortical areas could provide insight into where distributed signals originate in cortical networks, potentially revealing the unique functional contribution made by each participating area. Toward that end, we recorded neural activity in parietal and prefrontal cortex simultaneously and used pattern classification to decode spatial category from firing rates in a sequence of 50-ms time bins. This produced a time series of posterior probabilities quantifying the strengths of category representation in the two areas. We then investigated whether the time series in one cortical area could be used to predict the other (at a given lag). If so, we could infer that information about categories had been transmitted between cortical areas. We provide evidence that physiological signals encoding rule-dependent categories are transmitted selectively in a top-down direction from prefrontal to parietal neurons. This identifies a neural mechanism by which prefrontal output could rapidly adapt computations performed by distributed cortical networks to changing environmental demands.

RESULTS

We trained two monkeys to perform a rule-based spatial categorization task[18] (**Fig. 1**). In this task, monkeys viewed a spot visual stimulus

[1]Department of Neuroscience, University of Minnesota, Minneapolis, Minnesota, USA. [2]Brain Sciences Center, Veterans Affairs Medical Center, Minneapolis, Minnesota, USA. [3]Department of Biology, Augsburg College, Minneapolis, Minnesota, USA. [4]Medical Scientist Training Program (MD/PhD), University of Minnesota, Minneapolis, Minnesota, USA. [5]Center for Cognitive Sciences, University of Minnesota, Minneapolis, Minnesota, USA. [6]Department of Psychiatry, University of Minnesota, Minneapolis, Minnesota, USA. [7]Department of Psychology, University of Minnesota, Minneapolis, Minnesota, USA. Correspondence should be addressed to M.V.C. (chafee001@umn.edu).

Figure 1 Dynamic spatial categorization task, behavioral performance and network representation of spatial categories. (**a**) Stimulus sequence on a trial employing the horizontal (left-right) categorization rule. (**b**) Stimulus sequence on a trial employing the vertical (above-below) categorization rule. (**c**,**d**) Division of the circular sample array into horizontal categories under the horizontal rule (**c**) and vertical categories under the vertical rule (**d**). Sample stimuli were 0.25–0.5° in diameter and were presented at 13° eccentricity. (**e**) Proportion correct performance (black) and mean reaction time (gray) of monkey 1 (solid lines) and monkey 2 (dashed lines) under the horizontal and vertical categorization rules (error bars, ± s.e.m.). Response accuracy was significantly lower for vertical relative to horizontal categorization, both in monkey 1 (z-test of proportions; $n = 2,506$ trials, $z = 7.48$, $P < 0.0001$) and monkey 2 ($n = 8,988$ trials, $z = 11.67$, $P < 0.0001$). Responses were also significantly slower for vertical relative to horizontal categorization, in both monkey 1 (2-tailed unpaired t-test; $n = 1,037$ trials, $t = 2.04$, $P < 0.05$) and monkey 2 ($n = 3,536$ trials, $t = 6.67$, $P < 0.0001$). (**f**–**i**) Accuracy of decoding spatial categories on the basis of the activity of neuronal ensembles (**f**,**h**) or neuronal populations (**g**,**i**) in prefrontal cortex (PFC; black) and posterior parietal cortex (PAR; gray). S, sample presentation; R, rule cue presentation, C, choice. Black filled circles indicate time bins for which the proportion of correctly decoded trials in prefrontal cortex significantly exceeded that in parietal cortex (z-test of proportions, $P < 0.05$; $n = 888, 124, 1,294, 190$ observations per bin for **f**–**i**). Gray filled circles indicate the converse. (**f**,**g**) Accuracy of decoding vertical categories on the basis of the activity of neural ensembles (**f**; $n = 5$ parietal and prefrontal ensembles, 74 trials per ensemble) or neural populations (**g**; $n = 14$ prefrontal and 18 parietal neurons, 74 trials). (**h**,**i**) Accuracy of decoding horizontal categories on the basis of the activity of neural ensembles (**h**; $n = 6$ ensembles in each area, 74 trials per ensemble) or neural populations (**i**; $n = 16$ prefrontal and 15 parietal neurons, 74 trials).

(the sample) and then a line serving as a category boundary. The task was to categorize the sample on the basis of its spatial relationship to the boundary. When the boundary was vertical, monkeys assigned sample stimuli to the spatial categories 'left' and 'right'. Alternatively, when the boundary was horizontal, monkeys reassigned the same sample stimuli to the categories 'above' and 'below'. Changing the orientation of the category boundary over trials imposed different rules by which samples were assigned to categories, placing categorization under executive control.

Monkeys selected choices that accurately matched the spatial category of the sample according to the rule in force (**Supplementary Fig. 1**). Both monkeys responded with greater accuracy and faster reaction times when assigning sample positions to horizontal categories than to vertical categories (**Fig. 1e**), suggesting that vertical categorization was more difficult. Differences in behavioral performance as a function of the categorization rule applied were significant for both response accuracy and reaction time (**Fig. 1**).

Network representation of rule-dependent categories

Neural signals reflecting the spatial category of the sample as a function of the categorization rule provided a neurophysiological correlate of executive control over a spatial cognitive process. Rule-dependent and category-selective neural activity was distributed to neurons in both prefrontal and posterior parietal cortex. This included populations of neurons in both cortical areas with activity differentiating the categories 'above' and 'below' under the vertical but not the horizontal categorization rule, as well as neurons with activity differentiating
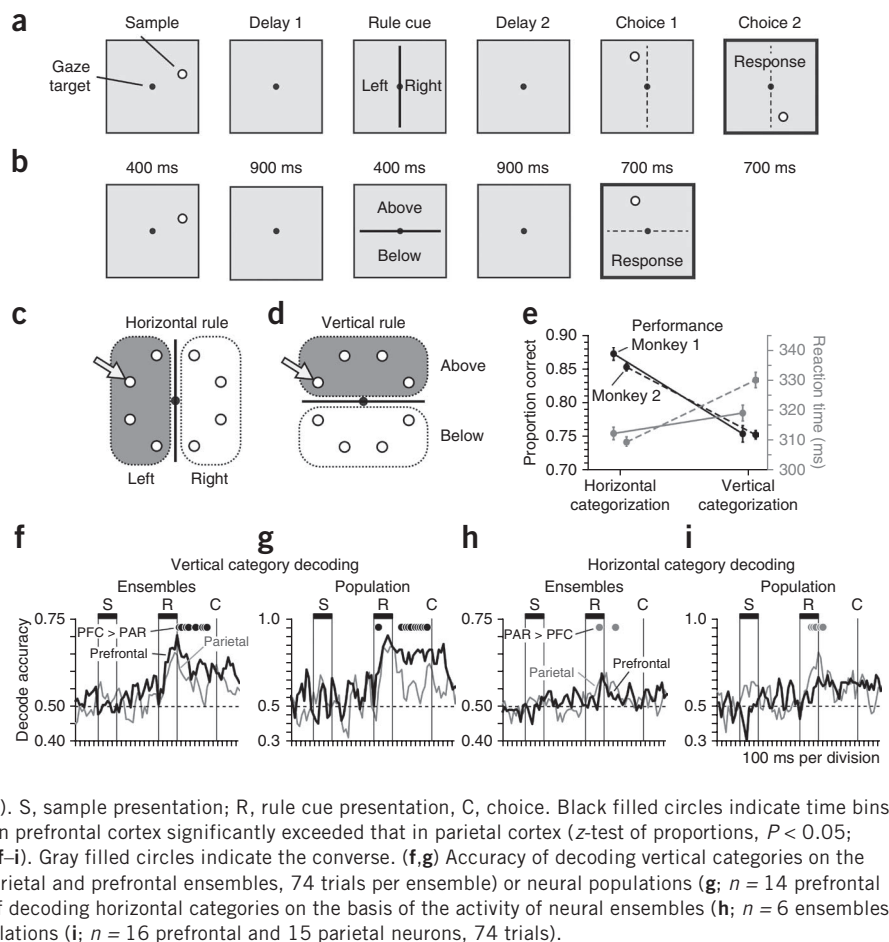
the categories 'left' and 'right' in a similarly rule-dependent fashion (**Supplementary Fig. 2**). Neural signals encoding rule-dependent categories did not encode the spatial position of the sample stimulus or the orientation of the category boundary (because the signal was a joint function of the two randomly associated task variables and was therefore dissociated from either of them considered individually). Rule-dependent category signals emerged after the presentation of the rule cue (**Supplementary Fig. 2**), when enough information had been provided in the trial to assign the sample to a rule-dependent category unambiguously.

To quantify the information about spatial categories that could be decoded from the pattern of activity in prefrontal and parietal neuronal populations, we applied pattern classification to firing rates in the two cortical areas measured in a sequence of 50-ms time bins throughout the trial. To compare decoding accuracy between cortical areas, we accumulated counts of accurate and erroneous category decoding (on a per-bin and per-trial basis) over the interval from the onset of the rule cue to the onset of the first choice stimulus and then tested the resulting distributions of dichotomous data using the z-test of proportions. The accuracy of decoding vertical categories was significantly greater when based on activity patterns in prefrontal cortex than in parietal cortex, both when decoding was based on the activity of neural ensembles (groups of neurons recorded simultaneously; **Fig. 1f**; $P < 10^{-8}$, $z = 6.41$, $n = 23,088$) and neural populations (accumulating significant neurons across recording sessions; **Fig. 1g**; $P < 10^{-15}$, $z = 8.15$, $n = 3,224$). Decoding the horizontal category of the sample on horizontal rule trials was, in contrast, significantly

Figure 2 The transmission analysis applied to hypothetical ensembles containing two category-selective neurons in prefrontal and parietal cortex. (**a**,**b**) Each open circle (numbered) illustrates the activity state of the prefrontal (**a**) and parietal (**b**) ensemble in a single 50-ms time bin as a point in a two-dimensional rate space, with the firing rates observed in neurons 1 and 2 plotted along the vertical and horizontal axes. Arrows connecting successive open circles illustrate the trajectory of each ensemble through its rate space over a short time span. Black filled circles indicate the mean firing rate observed in neurons 1 and 2 on 'above' and 'below' spatial category trials in the training data. Gray shading around each black circle represents the modeled Gaussian density of two-neuron activity patterns within each spatial category. Dashed lines (A–D) indicate distances in the rate space between activity patterns observed in each time bin on a single trial and the mean activity pattern based on the training data associated with the correct category for this trial (the correct category is 'below'). (**c**,**d**) Posterior probabilities associated with decoding the correct spatial category ('below' on this trial) on the basis of the activity pattern in each time bin in the prefrontal ensemble (**c**) and the parietal ensemble (**d**). A–D associate posterior probabilities with corresponding distances in the rate space in **a**,**b**.
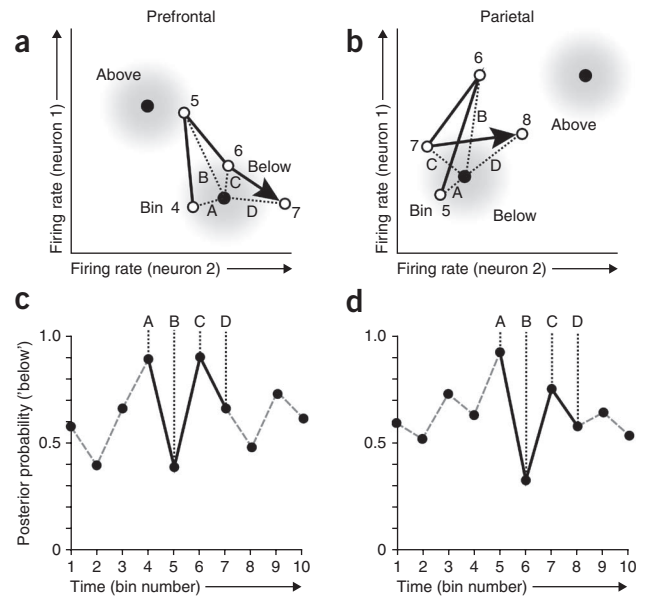
more accurate when based on neuronal activity patterns in parietal than prefrontal cortex, both at the ensemble level (**Fig. 1h**; $P < 0.02$, $z = 2.79$, $n = 33,644$) and at the population level (**Fig. 1i**; $P < 0.03$, $z = 1.98$, $n = 4,940$). However, the differences in decoding accuracy between cortical areas were smaller in this case (compare **Fig. 1h,i** to **Fig. 1f,g**).

We also compared the strength of vertical and horizontal category signals within each cortical area. Vertical decoding accuracy was significantly greater than horizontal decoding accuracy, both when based on neuronal activity in parietal cortex (**Supplementary Fig. 3a,b**; ensembles: $P < 10^{-10}$, $z = 6.70$, $n = 28,340$; population: $P < 0.02$, $z = 2.26$, $n = 4,056$) and in prefrontal cortex (**Supplementary Fig. 3c,d**; ensembles: $P < 10^{-50}$, $z = 15.00$, $n = 28,392$; population: $P < 10^{-35}$, $z = 12.54$, $n = 4,108$). These results are evidence that the cortical representation of vertical categories was stronger than the cortical representation of horizontal categories.

Measuring signal transmission between cortical areas
To determine whether neural signals encoding rule-dependent spatial categories were transmitted between prefrontal and parietal neurons during the trial, we evaluated whether rapid fluctuations in the strength of these signals were correlated between cortical areas over time. To measure the strength of category signals, we used the time series of posterior probabilities provided by the decoding analysis (as greater posterior probabilities correspond to more certain category membership based on neuronal activity). We then determined whether the time series of posterior probabilities in the two cortical areas were significantly correlated at different lags.

To illustrate the analytical approach, we consider hypothetical ensembles in prefrontal and parietal cortex, each of which contains two category-selective neurons (**Fig. 2**). The activity state of each ensemble can be represented by a point in a corresponding two-dimensional rate space (with the firing rate of one neuron along the horizontal axis and the other along the vertical). The evolving pattern of activity in each ensemble is captured by the sequence of firing rate measurements in the two neurons, comprising a trajectory through the rate space (**Fig. 2a,b**). The correct spatial category for this example trial, as based on the visual stimuli presented, is 'below'. As the firing rates of neurons 1 and 2 fluctuate over time, the pattern of ensemble activity in each time bin of this trial gets closer to (or farther from) the mean pattern of activity associated the category 'below' in the training data. (Activity on training trials is used to define the mean

and covariance of activity patterns associated with each category in the analysis.) Distances in the rate space (**Fig. 2a,b**) are converted to posterior probabilities (**Fig. 2c,d**) under the assumption that the distribution of activity patterns is multivariate Gaussian within each category (shorter distances in the rate space corresponding to greater probabilities of category membership). Note that for simplicity we illustrate the mean activity patterns for categories (**Fig. 2a,b**) as remaining fixed over bins. In the transmission analysis we adopted, we recomputed the mean and covariance of activity patterns associated with spatial categories within each time bin.

We evaluated whether the time series of posterior probabilities in the two cortical areas were correlated at different lags. For that purpose, we regressed the time series of posterior probabilities in one cortical area on the posterior probability time series in the other, keeping the two time series in temporal register (lag 0) or shifting one time series relative to the other by a variable number of time bins (lags 1 to 8). Before performing the regression analysis, we 'pre-whitened' the posterior probability time series to make them stationary over time and eliminate autocorrelation (**Supplementary Figs. 4 and 5**). This is advantageous because computing the cross correlation of time series that themselves contain strong time trends or are strongly autocorrelated can inflate estimates of cross-correlation[19–21]. After pre-whitening, we slid a 500-ms window (10 bins) through the trial in 50-ms time steps, aggregated values over trials and ensembles at each time step such that we maintained the pairing of data in parietal and prefrontal cortex to simultaneously recorded neurons in the two areas, and performed the regression using the posterior probability values in the window at each time step. This produced a time-varying estimate of signal transmission between cortical areas throughout the trial. We used the F-statistic associated with the regression at each time step to index the strength of the relationship.

Top-down transmission of rule-dependent category signals
We applied this analytical approach to separately evaluate the directional transmission of signals encoding vertical and horizontal categories at a lag of 1 time bin between the decoding data in the two cortical areas and found that vertical category signals were selectively transmitted in a top-down direction from prefrontal to posterior parietal neurons (**Fig. 3a**). The F-statistic associated with top-down transmission first exceeded the significance threshold (**Fig. 3a**) approximately

Figure 3 Lag 1 transmission of category signals between prefrontal and parietal neurons. Red and blue functions plot the time-varying F-statistic obtained by regressing the residual posterior probabilities within a sliding window in one cortical area on the corresponding probabilities in the other area shifted earlier by one 50-ms time bin. The dashed horizontal line indicates significance for the F-statistic ($P < 0.05$). (**a**) Top-down (red) and bottom-up (blue) transmission of vertical category signals. Black circles mark time bins in which the difference between top-down and bottom-up F values was significant ($P < 0.05$) in a permutation test randomly shuffling neurons across cortical areas and repeating the analysis. Ensembles containing at least two vertical category-selective neurons in parietal and prefrontal cortex were included ($n = 10$ ensembles, 32 neurons; parietal: 5 ensembles, 18 neurons; prefrontal: 5 ensembles, 14 neurons). Regression results at each point are based on 4,440 observations (posterior probabilities). Black and gray functions plot the mean posterior probability associated with the correct category when decoding was based on the activity of the prefrontal ensembles (PFC; black line) and parietal ensembles (PAR, gray line) included in the transmission analysis. R, duration of the rule cue. (**b**) Top-down and bottom-up transmission of neural signals encoding horizontal categories ($n = 12$ ensembles, 31 neurons; parietal: 6 ensembles, 15 neurons; prefrontal: 6 ensembles, 16 neurons). Regression results at each point are based on 5,250 observations (posterior probabilities).
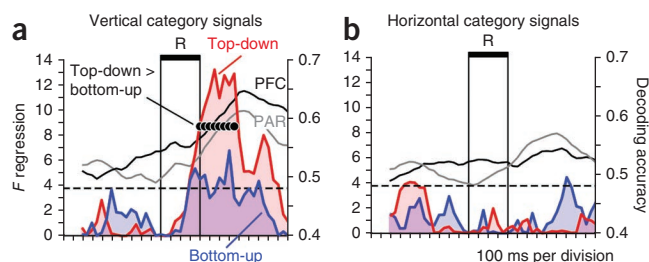
250 ms after the onset of the rule cue in the trial, when neural activity in both cortical areas started to exhibit differential activity as a function of the rule-dependent spatial category of the sample stimulus (**Supplementary Fig. 2**). Transmission of vertical category signals in the bottom-up direction, from parietal to prefrontal neurons, was significantly weaker (**Fig. 3a**). There was little evidence of significant transmission of horizontal category signals between prefrontal and parietal cortex in either direction (**Fig. 3b**), perhaps as a result of the significantly weaker neural signals encoding horizontal categories in this network (**Supplementary Fig. 3**; the proportion of correctly decoded vertical and horizontal category trials differed significantly, $P < 0.05$, z-test of proportions).

We found that our ability to detect network transmission using this method depended on the simultaneity of neural recording, as one would predict if our analysis captured physiological interaction between cortical areas in real time. We performed a bootstrap analysis in which we shuffled trials to break the simultaneity of neural activity in prefrontal and parietal cortex. In the shuffling, we kept the posterior probability time series intact, did not disrupt the relation between neural activity and spatial category, and maintained the grouping of neurons into simultaneously recorded ensembles. We found that breaking the simultaneity of activity in the two cortical areas was sufficient to destroy detected network transmission between them. Namely, the significance of top-down transmission as reflected by the time course of F-statistics obtained using the original data (**Fig. 4a**) far exceeded the 95th percentile of the bootstrap distribution

of F-statistics derived from the trial shuffled data (**Fig. 4a**; note the close approximation of the 95th percentile of the bootstrap distribution to the significance threshold for the F-statistic). This indicates that the temporal covariation in category signals that we observed did not reflect common activation of parietal and prefrontal neurons by the visual stimuli in the task, because in that case trial-shuffling would be predicted to have little effect.

To further confirm that interactions between cortical areas reflected transmission of category signals specifically and not merely coactivation of the two areas, we repeated the transmission analysis using the mean activity in each ensemble (averaged over neurons). In this analysis, we used the mean ensemble firing rate instead of the posterior probability associated with the correct spatial category in each time bin. We detected no evidence of significant transmission in this case (**Fig. 4a–d**). This provides some insight into the aspect of neuronal activation that supported signal transmission. Namely, the pattern of activity over neurons within an ensemble, and not just the overall level of activity within the ensemble, had to be evaluated to detect signal transmission between cortical areas. Taking the average ensemble rate as a proxy for more global measures of neural activity within a small volume of cortex (such as the local field potential or blood oxygen level–dependent signal), the implication is that resolving the information communicated between prefrontal cortex and connected areas requires resolving activity at the level of single neurons.

The pattern of top-down transmission of vertical category signals was robust when we relaxed our inclusion criterion to include more ensembles and neurons (**Supplementary Fig. 6**). We further investigated the potential influence of signals encoding the spatial position of the sample (rather than its spatial category) on the transmission results we obtained. First, we quantified the effect of sample position on category-selective activity. If category neurons were systematically tuned to the retinocentric locations of sample stimuli, one would predict that
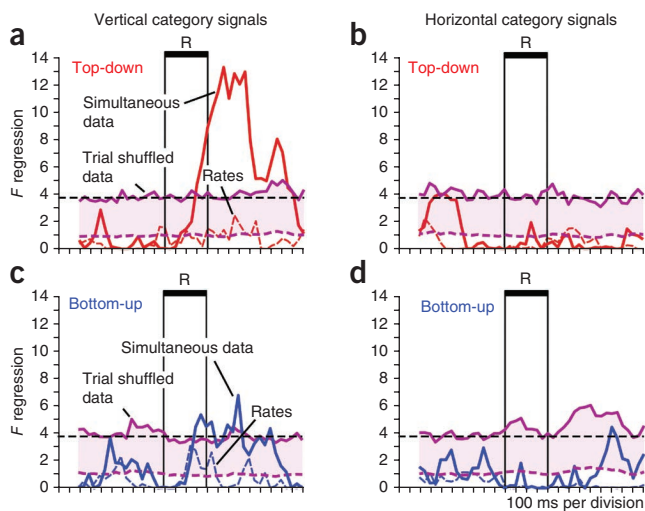
Figure 4 Dependence of lag 1 transmission on the simultaneity of neuronal activity recorded in parietal and prefrontal cortex. Dashed and solid purple lines indicate the mean and 95% confidence interval of a bootstrap distribution of F-statistics obtained in the transmission analyses after shuffling trials (100 iterations) to break the simultaneity of neural activity in parietal and prefrontal cortex. R, rule cue duration. Horizontal dashed black line indicates the threshold for significance ($P < 0.05$) of the F-statistic. (**a,b**) Top-down transmission of vertical (**a**) and horizontal (**b**) category signals based on decoded posterior probabilities (solid red lines) and mean ensemble firing rates (dashed red lines). (**c,d**) Bottom-up transmission of vertical (**c**) and horizontal (**d**) category signals based on decoded posterior probabilities (solid blue lines) and mean ensemble firing rates (dashed blue lines).

Figure 5 Sign of top-down and bottom-up interactions at lag 1. Top-down (black) and bottom-up (gray) transmission of vertical category signals between ensembles containing a minimum of two vertical category-selective neurons in both prefrontal and parietal cortex (sample sizes are as described in the legend of **Fig. 3**). R, rule cue duration. (**a,b**) F-statistics were divided into separate time courses on the basis of whether the sign of the associated regression coefficient was positive (**a**) or negative (**b**) at each time step.

(**c**) Magnitude and sign of the regression coefficient associated with top-down (black) and bottom-up (gray) transmission of vertical category signals at each time step. Significance of regression coefficients was evaluated in a permutation test shuffling posterior probabilities within ensembles across trials to break the simultaneity of recording (500 iterations). We evaluated the significance at $P < 0.05$ (two-tailed). Asterisks indicate time bins in which the regression coefficient was significantly greater than the 97.5th percentile of the distribution of coefficients from the shuffled data. Circles indicate values lower than the 2.5th percentile.
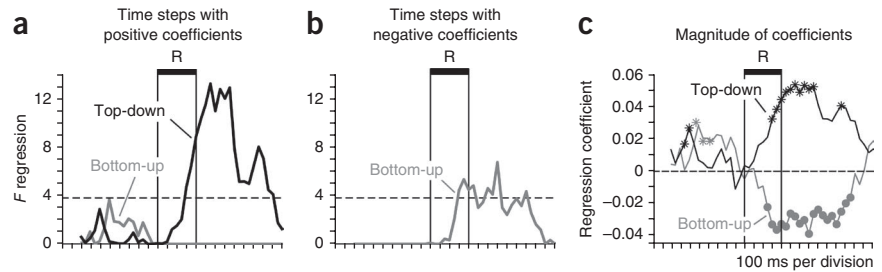
differences in their firing rate for sample positions that were next to one another would be smaller than for sample positions that were farther apart, within each neuron's preferred category. This was not the case (**Supplementary Fig. 7**). Next, we examined the influence on vertical category signal transmission of including, or excluding, category-selective neurons that also exhibited selectivity for sample position earlier in the trial. Top-down transmission of vertical category signals was not augmented by the inclusion of category-selective neurons in which activity also varied significantly as a function sample position (**Supplementary Fig. 8**). This is evidence that the transmission of category signals between prefrontal and parietal neurons was not strongly influenced by signals encoding the spatial position of the sample.

We also evaluated the transmission of neural signals encoding the categorization rule and the spatial position of the sample. We found evidence that rule signals were transmitted in a top-down direction from prefrontal to parietal neurons (**Supplementary Fig. 9b**), as based on the activity of neurons that varied as a function of the rule only and not spatial category. Evaluating neural signals encoding the position of the sample in parietal and prefrontal cortex, we found evidence that position signals covaried strongly and significantly between cortical areas at lag 0 (**Supplementary Fig. 9c**), an effect that peaked in the delay period following the offset of the sample stimulus. (Our analysis had limited ability to resolve transmission earlier in the trial, during the cue period, when feedforward transmission might be expected, because data had to be aggregated over a sequence of bins longer than the cue period before transmission could be detected.)

Transmission involves negative feedback

The transmission analysis revealed that vertical category signals in one cortical area could either drive or suppress vertical category signals in the other. When we evaluated top-down transmission of vertical category signals at a lag of 1 time bin, the resulting regression coefficients were positive (**Fig. 5**). This provided evidence that prefrontal vertical category signals drove parietal vertical category signals, in the sense that lagged changes in the reliability of category representation in the two areas were in the same direction, with prefrontal cortex leading.

When we evaluated bottom-up transmission of vertical category signals at lag 1, we found the converse. In this case, the resulting regression coefficients were predominantly negative, and stronger category signals in parietal cortex predicted weaker category signals in prefrontal cortex 1 time bin later (**Fig. 5**). This suggests that top-down signal transmission from prefrontal cortex to target cortical areas may operate under the control of negative feedback.

Transmission is temporally focused and varies with behavior

We repeated the transmission analysis evaluating interactions between cortical areas at a broader range of temporal lags, from 0 to 8 time bins (0 to 400 ms). The time course of F-statistics associated with top-down transmission far exceeded the threshold for significance at lags 1 and 3 (**Fig. 6a**), but it remained very near that threshold at other lags, providing the strongest evidence for top-down interactions occurring at 50 and 150 ms. When we evaluated the magnitude and
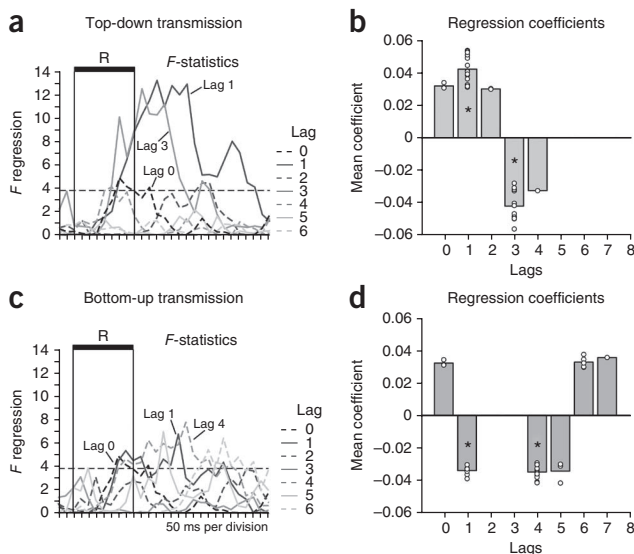

Figure 6 Significance and sign of transmission as a function of the lag between the posterior probability time series associated with vertical category representation in prefrontal and parietal cortex. (**a**) Time courses plot of the F-statistic associated with top-down transmission of vertical category signals at different lags (ranging from 0 to 6 time bins, 50 ms each), with earlier prefrontal values predicting later parietal values. R, rule cue duration. (**b**) The sign and magnitude of the corresponding mean significant regression coefficients associated with top-down transmission as a function of the lag (0–8 time bins) between prefrontal and parietal time series. Coefficients were averaged over the rule and subsequent delay periods (open circles indicate coefficients at each time point contributing to the average). Asterisks indicate mean coefficients that were significantly different from 0 (sign test, $P < 0.05$). Lags without bars indicate the failure to detect significant transmission at those lags (as reflected by the absence of significant F-statistics in the transmission time courses). (**c,d**) F-statistic time courses (**c**) and mean regression coefficients (**d**) associated with bottom-up transmission of vertical category signals. Transmission analyses were based on ensembles containing a minimum of two vertical category-selective neurons (sample characteristics are as described in the legend of **Fig. 3**).

Figure 7 Network transmission of vertical category signals is modulated by behavioral performance. Data show influence of recent reward on the strength of vertical category signal transmission. For each trial under the vertical categorization rule, we computed the proportion of the previous ten trials that had been rewarded under the vertical categorization rule. We then divided trials into high and low recent reward groups on the basis of whether the proportion of recent reward on the vertical rule was above or below the median value over all trials (0.4). R, rule cue duration. (**a**) Time courses (with shading) compare F-statistics associated with top-down transmission of vertical category signals on high recent reward trials (thick line) and low recent reward trials (dashed line). Filled black circles indicate time bins in which the difference between reward conditions was significant ($P < 0.05$) in a permutation test repeating the transmission analysis after randomly shuffling trials across reward condition (100 iterations). Black (PFC) and gray (PAR) time courses illustrate the mean posterior probability obtained from ensemble decoding in prefrontal and parietal cortex on high recent reward and low recent reward trials. (**b**) Time courses (with shading) illustrate F-statistics associated with bottom-up transmission of vertical category signals on high recent reward trials and low recent reward trials (conventions as in **a**). Transmission analyses were based on ensembles containing a minimum of two vertical category-selective neurons (sample sizes are as described in the legend of **Fig. 3**).
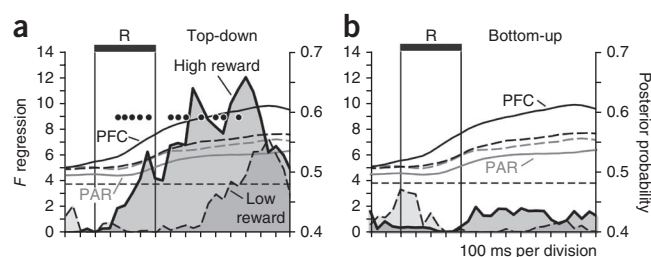
sign of the regression coefficients associated with transmission as a function of lag, we found that the sign of top-down interactions flipped from a significantly positive mean regression coefficient at lag 1 (**Fig. 6b**; sign test, $P < 10^{-5}$) to a significantly negative mean coefficient at lag 3 (**Fig. 6b**; sign test, $P < 0.001$). The inversion in the sign from earlier positive to later negative top-down effects could sharpen the interaction between cortical areas in time, effectively increasing the temporal resolution of transmission. Bottom-up interactions were comparatively weak (**Fig. 6c**), and the mean regression coefficient was negative over a broader time range—specifically, lag 1 (**Fig. 6d**; sign test, $P < 0.01$) and lag 4 (sign test, $P < 0.001$).

We investigated whether the strength of vertical category signal transmission varied as a function of behavioral performance. Rather than comparing signal transmission on correct and error trials, as relatively few error trials were available, we divided successfully performed vertical categorization trials into two groups differentiated on the basis of whether monkeys had earned a comparatively large or small proportion of reward under the vertical categorization rule in the recent past. (Trials were split into high and low recent reward groups on the basis of whether the number of rewards earned under the vertical categorization rule in the past ten trials was above or below the median value over all trials). Recent reward provided a behavioral metric analogous in some respects to local fractional income, a parameter that influences neural signals in the lateral intraparietal area during a decision-making task[22]. Top-down transmission of vertical category signals was significantly more reliable on high in comparison to low recent reward trials (**Fig. 7a**). We did not see evidence of bottom-up transmission in this analysis (**Fig. 7b**; most likely because of the reduction in statistical power associated with dividing the trials available for each analysis in half). The residual posterior probability time series obtained by fitting an ARIMA (autoregressive, integrative, moving average) model to the decoding data on high and low recent reward trials essentially overlapped (**Supplementary Fig. 10**). It was not likely, therefore, that differences in the reliability of vertical category signal transmission on high and low recent reward trials could be attributed to differences in the strength of neural signals encoding vertical categories under the two reward conditions.

DISCUSSION
We obtained evidence that neural signals encoding rule-dependent spatial categories, and therefore reflecting the executive control of a cognitive process, were selectively transmitted in a top-down direction from prefrontal to parietal neurons. Transmission was directional and selective for the nature of the transmitted information, depended

on simultaneously recorded neural activity in the two areas, occurred at a restricted time scale and was modulated as a function of behavior. These data effectively translate the information content of a physiological signal transmitted from prefrontal to parietal neurons during cognitive processing.

Several lines of evidence argue that our results captured real-time transmission of category information between cortical areas. First, transmission was abolished in trial-shuffled data, even though the shuffling procedure kept the underlying time series of category signals intact and only disrupted their simultaneity across cortical areas. Second, the transmission analysis was performed on the residual posterior probability time series after fitting an ARIMA model. This substantially reduced time trends and autocorrelation, producing a time series that essentially fluctuated around a zero mean, effectively attenuating signal transients driven by external events. Third, we did not detect transmission when the analysis was based on the mean ensemble activity rather than the information about categories encoded by the pattern of ensemble activity. Fourth, top-down and bottom-up transmission exhibited opposite signs of influence, which would not be predicted if transmission simply reflected coactivation of parietal and prefrontal neurons. Fifth, interactions between category signals in prefrontal and parietal cortex were significant at lag 1 but not lag 0, providing evidence that one cortical area drove the other, rather than both being driven by common input.

The observations that top-down transmission and bottom-up transmission of vertical category signals involved positive and negative interactions, respectively, which took place at nearly the same time during the trial, suggested to us that top-down control in prefrontal networks might operate under negative feedback control. Parietal and prefrontal cortices are linked by reciprocal excitatory projections[3]. Negative feedback from parietal cortex (possibly mediated through inhibitory interneurons) could moderate the strength of prefrontal output. This could balance top-down and bottom-up interactions in the network and prevent prefrontal output from dominating network dynamics to a point where bottom-up processing could no longer influence network activity. Additionally, we found that top-down transmission flipped from a positive influence at shorter lags to a negative influence at longer ones (**Fig. 6**). That dynamic could serve to focus the influence of transmitted signals in time, effectively increasing the temporal resolution of the interaction between cortical areas.

Our further finding that top-down transmission was modulated in strength as a function of recent reward history suggests that reward-driven dynamics may flexibly weight distributed signals encoding cognitive strategies, in addition to signals encoding stimuli[23,24] or

planned actions[22,25,26]. That could frame executive control as an additional instance of biased competition[27], or a decision process[22], adjudicating between neural representations of different rules or cognitive strategies in addition to alternative stimuli or actions.

We found that network transmission was strongly biased, favoring vertical over horizontal category signals. There are several possible accounts. One account could relate to an observation made consistently across laboratories and experiments that parietal neurons (and, given their physiological similarity, potentially prefrontal neurons as well) often represent the various conditions of cognitive tasks in a strongly biased manner, such that a disproportionately large fraction of neurons are most strongly activated for the same single condition out of the several making up the experimental design[28]. The bias in top-down transmission favoring vertical category signals could, for example, reflect the strong bias of the underlying neural representations in favor of neurons encoding vertical categories in both cortical areas, but particularly in prefrontal cortex. Behavioral performance was significantly worse for vertical categorization than horizontal categorization in both monkeys. This suggests that vertical categorization was more difficult, potentially requiring greater top-down executive control. Both monkeys were trained on horizontal categorization first and had more experience with it, so horizontal categorization may have been more automatic or habitual. Prefrontal network dynamics may therefore vary as a function of executive demand, with top-down control being greatest when executive demand is high. That is suggested by the finding that category signals can be stronger and earlier in parietal cortex relative to prefrontal cortex in categorization tasks that do not change the categorization rule[16] (although see ref. 17). That prefrontal network dynamics are adaptive to task demands has been directly demonstrated in a study comparing parietal and prefrontal dynamics during top-down and bottom-up attention[11].

An analysis of the relative timing and strength of distributed signals shared across connected cortical areas can provide indirect evidence of how information flows between them during cognitive processing. Such differences in the timing and strength of signals distributed across parietal and prefrontal cortex are observed in various cognitive tasks[11,12,16–18]. However, differences in signal timing and strength do not demonstrate the transmission of a identified signal from one area to another, or quantify the information content of the transmitted signal. Likewise, studies of functional connectivity[29] or neural synchrony[11,30,31] measure covariation in neural activity between areas. Although these interactions can vary informatively as a function of task demand[11] and can be related to functionally distinct groups of neurons[31], covariation in neural activity does not necessarily equate to covariation in coded information. Other studies have examined how reversible inactivation of prefrontal cortex affects neural activity in parietal cortex and vice versa[32]. However, inactivating a cortical area suppresses all of the neural signals in it, and it is not clear therefore which subset of signals is transmitted to distant targets to mediate remote effects during inactivation. Finally, antidromic stimulation studies have characterized the nature of prefrontal output signals by identifying which prefrontal neurons project an axon to a target structure[33]. However, this approach does not measure the physiological effect of signal transmission in the identified pathway or quantify its impact on information processing in the target structure.

Our results identify a specific neural signal related to the executive control of cognition that is transmitted from prefrontal to parietal cortex. These results are the first, to our knowledge, to effectively decode the cognitive information that is conveyed by a signal transmitted from one cortical area to another. Rather than reflecting a monolithic computational strategy equally applied across all rule-based behaviors,

prefrontal output is strongest (**Fig. 3**) when rule conditions are most demanding (**Fig. 1e**) and is dynamically regulated as a function of recent reward history (**Fig. 7**), as monkeys flexibly adapt their cognitive processing to changing environmental conditions.

Prefrontal cortex receives and sends information along a complex network of projections that link it to other cortical association areas, via direct corticocortical projections[3,5] and transthalamic pathways[34]. The communication and exchange of neural activity along these pathways is likely to distribute the neural representation of each cognitive or task variable across multiple cortical areas, creating subpopulations of neurons in each area with activity that exhibits essentially the same relationship to behavior. That is supported by the frequent observation that activity patterns in prefrontal cortex and connected areas often differ more in degree than in kind[11,12,16–18], as we found in the current study (**Supplementary Figs. 2** and **3**), and it is often difficult to differentiate the functions of prefrontal cortex from connected cortical areas in the basis of the types of information coded by activity in each area. Analyzing how distributed neural signals are transmitted between prefrontal and parietal neurons could uncover the flow of information through prefrontal networks to discern how the final distributed pattern of activity came about. That could disambiguate the functional contributions of each area. We adopted that approach in this study to provide evidence that neural signals encoding rule-dependent spatial categories were preferentially transmitted in a top-down direction from prefrontal to parietal neurons. This provides insight into the neural mechanism by which prefrontal output could exert executive control over distributed cognitive processing in cortical networks.

METHODS
Methods and any associated references are available in the online version of the paper.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

AUTHOR CONTRIBUTIONS
D.A.C. and M.V.C. analyzed the data and wrote the manuscript. S.J.G., R.K.B., S.S., S.R.S. and A.W.M. edited the manuscript. S.J.G. and M.V.C. designed the experiment, and S.R.S. and A.W.M. contributed to experimental design. S.J.G. trained the monkeys and collected the neural data. R.K.B., M.V.C. and S.S. assisted in neuronal data collection.

COMPETING FINANCIAL INTERESTS
The authors declare no competing financial interests.

Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.

1. Medalla, M. & Barbas, H. Diversity of laminar connections linking periarcuate and lateral intraparietal areas depends on cortical structure. *Eur. J. Neurosci.* **23**, 161–179 (2006).
2. Schwartz, M.L. & Goldman-Rakic, P.S. Callosal and intrahemispheric connectivity of the prefrontal association cortex in rhesus monkey: relation between intraparietal and principal sulcal cortex. *J. Comp. Neurol.* **226**, 403–420 (1984).

3. Cavada, C. & Goldman-Rakic, P.S. Posterior parietal cortex in rhesus monkey: II. Evidence for segregated corticocortical networks linking sensory and limbic areas with the frontal lobe. *J. Comp. Neurol.* **287**, 422–445 (1989).

4. Andersen, R.A., Asanuma, C., Essick, G. & Siegel, R.M. Corticocortical connections of anatomically and physiologically defined subdivisions within the inferior parietal lobule. *J. Comp. Neurol.* **296**, 65–113 (1990).

5. Goldman-Rakic, P.S. Circuitry of primate prefrontal cortex and regulation of behavior by representational memory. in. *Handbook of Physiology: The Nervous System: Higher Functions of the Brain* (eds. Mountcastle, V.B., Plum, F. & Geiger, S.R.) 373–417 (Am. Physiol. Soc., Bethesda, Maryland, USA, 1987).

6. Andersen, R.A. & Cui, H. Intention, action planning, and decision making in parietal-frontal circuits. *Neuron* **63**, 568–583 (2009).

7. Qi, X.L. *et al.* Comparison of neural activity related to working memory in primate dorsolateral prefrontal and posterior parietal cortex. *Front Syst. Neurosci* **4**, 12 (2010).

8. Chafee, M.V. & Goldman-Rakic, P.S. Matching patterns of activity in primate prefrontal area 8a and parietal area 7ip neurons during a spatial working memory task. *J. Neurophysiol.* **79**, 2919–2940 (1998).

9. Funahashi, S., Chafee, M.V. & Goldman-Rakic, P.S. Prefrontal neuronal activity in rhesus monkeys performing a delayed anti-saccade task. *Nature* **365**, 753–756 (1993).

10. Gnadt, J.W. & Andersen, R.A. Memory related motor planning activity in posterior parietal cortex of macaque. *Exp. Brain Res.* **70**, 216–220 (1988).

11. Buschman, T.J. & Miller, E.K. Top-down versus bottom-up control of attention in the prefrontal and posterior parietal cortices. *Science* **315**, 1860–1862 (2007).

12. Nieder, A. & Miller, E.K. A parieto-frontal network for visual numerical information in the monkey. *Proc. Natl. Acad. Sci. USA* **101**, 7457–7462 (2004).

13. Nieder, A., Freedman, D.J. & Miller, E.K. Representation of the quantity of visual items in the primate prefrontal cortex. *Science* **297**, 1708–1711 (2002).

14. Vallentin, D. & Nieder, A. Representations of visual proportions in the primate posterior parietal and prefrontal cortices. *Eur. J. Neurosci.* **32**, 1380–1387 (2010).

15. Tudusciuc, O. & Nieder, A. Contributions of primate prefrontal and posterior parietal cortices to length and numerosity representation. *J. Neurophysiol.* **101**, 2984–2994 (2009).

16. Swaminathan, S.K. & Freedman, D.J. Preferential encoding of visual categories in parietal cortex compared with prefrontal cortex. *Nat. Neurosci.* **15**, 315–320 (2012).

17. Merchant, H., Crowe, D.A., Robertson, M.S., Fortes, A.F. & Georgopoulos, A.P. Top-down spatial categorization signal from prefrontal to posterior parietal cortex in the primate. *Front Syst. Neurosci* **5**, 69 (2011).

18. Goodwin, S.J., Blackman, R.K., Sakellaridi, S. & Chafee, M.V. 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 (2012).

19. Box, G.E.P., Jenkins, G.M. & Reinsel, G.C. *Time Series Analysis: Forecasting and Control* (Prentice-Hall, Upper Saddle River, New Jersey, USA, 1994).

20. Granger, C.W.J. & Newbold, P. *Forecasting Economic Time Series* (Academic, New York, 1977).

21. Christova, P., Lewis, S.M., Jerde, T.A., Lynch, J.K. & Georgopoulos, A.P. True associations between resting fMRI time series based on innovations. *J. Neural Eng.* **8**, 046025 (2011).

22. Sugrue, L.P., Corrado, G.S. & Newsome, W.T. Matching behavior and the representation of value in the parietal cortex. *Science* **304**, 1782–1787 (2004).

23. Gottlieb, J. & Balan, P. Attention as a decision in information space. *Trends Cogn. Sci.* **14**, 240–248 (2010).

24. Peck, C.J., Jangraw, D.C., Suzuki, M., Efem, R. & Gottlieb, J. Reward modulates attention independently of action value in posterior parietal cortex. *J. Neurosci.* **29**, 11182–11191 (2009).

25. Platt, M.L. & Glimcher, P.W. Neural correlates of decision variables in parietal cortex. *Nature* **400**, 233–238 (1999).

26. Barraclough, D.J., Conroy, M.L. & Lee, D. Prefrontal cortex and decision making in a mixed-strategy game. *Nat. Neurosci.* **7**, 404–410 (2004).

27. Desimone, R. & Duncan, J. Neural mechanisms of selective visual attention. *Annu. Rev. Neurosci.* **18**, 193–222 (1995).

28. Fitzgerald, J.K. *et al.* Biased associative representations in parietal cortex. *Neuron* **77**, 180–191 (2013).

29. Nee, D.E. & Brown, J.W. Dissociable frontal-striatal and frontal-parietal networks involved in updating hierarchical contexts in working memory. *Cereb. Cortex* **23**, 2146–2158 (2013).

30. Gregoriou, G.G., Gotts, S.J., Zhou, H. & Desimone, R. High-frequency, long-range coupling between prefrontal and visual cortex during attention. *Science* **324**, 1207–1210 (2009).

31. Gregoriou, G.G., Gotts, S.J. & Desimone, R. Cell-type-specific synchronization of neural activity in FEF with V4 during attention. *Neuron* **73**, 581–594 (2012).

32. Chafee, M.V. & Goldman-Rakic, P.S. Inactivation of parietal and prefrontal cortex reveals interdependence of neural activity during memory-guided saccades. *J. Neurophysiol.* **83**, 1550–1566 (2000).

33. Ferraina, S., Pare, M. & Wurtz, R.H. Comparison of cortico-cortical and cortico-collicular signals for the generation of saccadic eye movements. *J. Neurophysiol.* **87**, 845–858 (2002).

34. Sherman, S.M. & Guillery, R.W. Distinct functions for direct and transthalamic corticocortical connections. *J. Neurophysiol.* **106**, 1068–1077 (2011).

ONLINE METHODS

Behavioral paradigm. We developed a DYnamic Spatial Categorization (DYSC) task that required monkeys to flexibly categorize the same set of visual stimuli according to alternative grouping criteria, or categorization rules[18] (**Fig. 1**). Trials began with the presentation of a gaze fixation target at the center of the display. Monkeys had to maintain gaze within 2.5° of this target throughout the trial. After an initial fixation period, a circular sample stimulus (0.25–0.5° diameter) was presented for 400 ms at an eccentricity of 13° (**Fig. 1a,b**; 'Sample'). The sample position was selected pseudorandomly from a circular array of positions (**Fig. 1c,d**). A delay period (**Fig. 1a,b**; 'Delay 1'; 900 ms) followed the offset of the sample, after which a rule cue was presented (**Fig. 1a,b**; 'Rule cue'; 400 ms). The rule cue was a line that bisected the sample array in either a vertical or horizontal orientation and served as a category boundary. When the rule cue was a vertical line, it instructed the horizontal categorization rule and divided the circular sample array into the spatial categories 'left' and 'right' on opposite sides of the boundary (**Fig. 1a,c**). In this case, the horizontal coordinate of the sample position was critical to computing the spatial category it belonged to, whereas the vertical coordinate was irrelevant to category membership. When the rule cue was a horizontal line, it instructed the vertical ('above' or 'below') categorization rule, requiring monkeys to reassign the same set of sample positions to a new set of spatial categories (**Fig. 1b,d**). In this case, the vertical coordinate of the sample position determined category membership whereas the horizontal coordinate was irrelevant. In the data included in this report, monkey 1 performed the DYSC task while the categorization rule was randomized over trials. Monkey 2 performed the DYSC task in blocks of random length from 7 to 12 trials. After the rule cue disappeared, a second delay period followed (**Fig. 1a,b**; 'Delay 2'; 900 ms). Following the second delay, two choice stimuli (circular, 0.25–0.5° diameter) appeared sequentially (**Fig. 1a,b**; 'Choice 1', 'Choice 2'; 700 ms each). One choice was located in the same spatial category as the sample (the match), by virtue of being on the same side of the category boundary as the sample on that trial. The other choice was located in the opposite spatial category (the nonmatch). The order of choice presentation (match and nonmatch) was randomized over trials (the boundary was not visible during the choice sequence). Monkeys responded by controlling when they depressed a single response key (a pedal with their foot) in relation to the randomized choice sequence. The outcome of the trial (correct or incorrect) was dictated by the timing of the motor response relative to the choice sequence; movement direction did not vary. If the choice stimulus that was visible at the time of the motor response matched the spatial category of the sample, defined by the sample location and the orientation of the rule cue, the monkey was rewarded with a drop of juice.

Neuronal database. We conducted neural recording in Brodmann's area 7a in posterior parietal cortex and areas 46 and 9 in dorsolateral prefrontal cortex in two male rhesus macaque monkeys (*Macaca mulatta*; 5–8 kg). We simultaneously recorded the spiking activity of small ensembles of prefrontal and posterior parietal neurons consisting of ~15–25 individually isolated cells in each cortical area using two 16-electrode Eckhorn Microdrive matrices, one positioned over each cortical area (Thomas Recording, Giessen, Germany). The electrodes controlled by each matrix could be advanced through the dura and into the brain independently under computer control to isolate the action potentials of distinct sets of simultaneously recorded neurons (ensembles). We typically recorded the electrical activity of 2–3 neural ensembles per day. The waveforms of single neurons were isolated online by two operators working in parallel. Additional details of surgery and neural recording have been reported previously[18]. The results reported in this study are based on the electrical activity of 1,360 isolated cortical neurons, including 666 neurons in parietal cortex and 694 cells in prefrontal cortex. We did not perform a priori power analyses to determine sample size. (Little prior information was available on which to base an estimate of effect sizes relating to transmission of task-related signals between cortical areas.) Rather, we obtained as large a sample of neurons as was practically feasible given technical constraints of acute single neuron recording and the necessity of isolating task-related activity in two cortical areas concurrently. All neurons encountered and isolated were recorded without preselection (i.e., blindly). By necessity, transmission analyses were performed on neurons identified as carrying specific task-related signals in a prior ANOVA (below), so experimenters were not blinded with respect to neuronal type after this stage in the analysis. All surgical and animal care procedures conformed to the Principles of Laboratory

Animal Care of the NIH and protocols approved by the Animal Care and Use Committees of the University of Minnesota and the Minneapolis Veterans Affairs Medical Center.

ANCOVA-based selection of simultaneously recorded neurons for transmission analysis. To quantify the transmission of information between cortical areas, we first applied analysis of covariance (ANCOVA) to identify the subsets of neurons within each ensemble that exhibited activity varying significantly as a function of rule-dependent spatial category. The dependent variable in the ANCOVA was the firing rate of each neuron averaged over the 'Rule cue' and following 'Delay 2' task periods (**Fig. 1a,b**). The factors in the ANCOVA were horizontal category (left or right), vertical category (above or below) and the categorization rule (horizontal or vertical), along with the interactions between rule and both horizontal and vertical category. For this analysis, we coded the horizontal and vertical categories of the sample stimulus each trial as a function of the sample's location in the stimulus array regardless of the rule in force. We included the firing rate averaged over the preceding Sample and Delay 1 periods (**Fig. 1**) for each trial as a covariate in the analysis. This allowed us to evaluate the influence of rule-dependent category on firing rate apart from any carryover influence of the position of the sample stimulus. To conduct the transmission analysis, we selected neurons that exhibited rule-dependent, category-selective activity, as indicated by a significant interaction between the categorization rule and either the horizontal or vertical category factors ($P < 0.05$).

Supplementary Figure 4a illustrates two groups of simultaneously recorded neurons in parietal and prefrontal cortex, subsets of which have activity relating significantly to vertical categories (colored circles). We defined a neural ensemble to be a group of simultaneously recorded neurons within a cortical area with category-selective activity under the same categorization rule. For example, an ensemble might consist of 3 prefrontal neurons each of which was preferentially activated to represent the vertical categories 'above' and 'below', or 2 parietal neurons both of which were preferentially activated to represent the horizontal categories 'left' or 'right'. We evaluated transmission between neural ensembles in prefrontal and parietal cortex that (i) were recorded simultaneously in the two cortical areas and (ii) contained a minimum number neurons with activity significantly related to vertical categories in both cortical areas or to horizontal categories in both cortical areas. Restricting the analysis to ensembles containing a minimum number of cells with activity encoding spatial categories along the same spatial axis (horizontal or vertical) involved a tradeoff between the minimum numbers of category-selective neurons each ensemble had to contain to be included and the number of ensembles available for analysis. We performed the analysis requiring that simultaneously recorded ensembles in prefrontal and parietal cortex both contain a minimum of either 1 or 2 significantly category-selective neurons with activity relating to spatial categories along the same spatial axis (horizontal or vertical). Due to the constraints of random sampling, the number of ensembles containing larger numbers of simultaneously recorded neurons encoding the same sets of spatial categories was limited.

The main results of this paper are based on application of the more stringent criterion (minimum of 2 significant neurons simultaneously recorded in each of the two areas), which yielded a total of 22 simultaneously recorded ensembles containing 63 task-related neurons. Transmission analysis of vertical category signals was based on decoding obtained from 32 vertical category-selective neurons. These neurons were simultaneously recorded in 5 parietal ensembles (containing 18 significant neurons) and 5 prefrontal ensembles (containing 14 significant neurons). Transmission analysis of horizontal category signals was based on the activity of 31 neurons, segregated into 6 parietal ensembles (15 neurons) and 6 prefrontal ensembles (16 neurons). However, we relaxed the inclusion criteria and repeated the transmission analyses requiring that ensembles include a minimum 1 significant category-selective neuron along the same spatial axis (horizontal or vertical) in both cortical areas. Using this criterion, results of the transmission analysis were based on 52 neuronal ensembles containing a total of 108 neurons. For vertical categories, the analysis was based on 15 parietal ensembles (35 neurons) and 15 prefrontal ensembles (27 neurons). For horizontal categories, the analysis was based on 11 parietal ensembles (23 neurons) and 11 prefrontal ensembles (23 neurons).

Decoding rule-dependent categories from ensemble activity in prefrontal and parietal cortex. To characterize the transmission of cognitive signals between

prefrontal and parietal neurons, we first applied a pattern classification analysis to decode spatial category from the pattern of firing rates observed in neural ensembles (**Fig. 2** and **Supplementary Fig. 4**). To capture temporal variation in category signals over the course of the trial, we applied the decoding analysis to ensemble activity patterns measured in a sequence of 50-ms time bins. At each time step, the output of the decoding algorithm provided the posterior probabilities that the sample stimulus belonged to each of the two alternative spatial categories possible under the current rule on the basis of the pattern of firing rates over neurons observed. The posterior probability quantifies the confidence of each category assignment on the basis of neural activity and therefore indicates the strength with which neural activity patterns encode categories. For the transmission analysis, we used the posterior probabilities associated with the correct spatial category each trial, as defined by the sample and rule cue shown. In addition, we restricted the analysis to the half of the trials in the set for which the category preference of the neurons being analyzed was relevant to the behavioral choice the monkeys had to make. For example, when decoding the vertical categories 'above' and 'below', we selected neurons with activity significantly influenced by the interaction between rule and vertical category and then used their activity to decode the vertical category of the sample on trials when the vertical categorization rule was in force.

To perform the decoding analysis, we constructed an ensemble rate vector capturing a snapshot of activity within a short, 150-ms time window. The ensemble rate vector within the time window was constructed by concatenating the firing rate of each significant category-selective neuron in the ensemble measured within 3 consecutive 50-ms time bins. We then advanced this 150-ms window through the trial in 50-ms steps and decoded category at each time step to produce a decoding time course of posterior probabilities. This time course measured rapid variation in signal strength over time (**Figs. 1f–i** and **2** and **Supplementary Fig. 4b**).

We applied a linear discriminant pattern classification algorithm (the 'classify' function in the Matlab Statistical Toolbox; MathWorks, Natick, MA) to compute the posterior probabilities associated with each spatial category. We performed this analysis using fivefold cross-validation. At each time point, we trained the classifier using 4/5 of trials (the training set). Using the training data, we computed the mean and covariance of the ensemble activity vectors associated with each of the two alternative categories under the current categorization rule. These values provided the free parameters of classification functions modeling the distribution of activity patterns associated with each category as a multivariate Gaussian distribution. We then used the classification functions to decode spatial categories from the ensemble rate vectors observed on the remaining 1/5 of trials (the test set). For each time bin in each test trial, the classification functions computed the distance between the ensemble rate vector and the means of the two Gaussian distributions in the rate space that were associated with the two spatial categories in the training data (**Fig. 2a,b**). The functions then converted these distances to posterior probabilities that the sample belonged to each category based on the activity pattern observed (**Fig. 2c,d**). We used the classification functions defined by neuronal activity within a single time bin to classify trials on the basis of activity in that same time bin. We retrained the classifier to perform the classification in each subsequent time bin. We classified successive one-fifths of trials, using the remaining trials as training data, until all trials were classified. The time series of posterior probabilities captured variation in the strength of signals encoding rule-dependent categories during individual trials in each cortical area (**Fig. 2c,d** and **Supplementary Fig. 4b**), enabling us to evaluate correlation in these signals over time (below).

We evaluated whether differences in decoding accuracy in parietal and prefrontal cortex were significant using the z-test of proportions applied to counts of accurate and inaccurate decoding of spatial category over trials in individual time bins, or accumulated over the time bins between the onset of the rule cue and first choice stimulus.

ARIMA preprocessing of decoding time series. We then sought to determine whether short-term fluctuations in the time series of posterior probabilities in one cortical area preceded and predicted lagged fluctuations in the time series in the other cortical area, providing evidence that a neural signal specifically encoding rule-dependent categories had been transmitted (directly or indirectly) from parietal to prefrontal cortex. Estimates of cross-correlation between any two time series are biased and inflated in the case that the times series are themselves

autocorrelated or nonstationary[19–21]. In this case, the internal regularity of the time series increases the degree to which they appear to be interdependent. To obtain estimates of interaction between cortical areas that were relatively free of these concerns, we applied a transformation developed by Box, Jenkins, Granger and others[19,20] to the posterior probability time series, in order to remove their autocorrelation and make then stationary. For that purpose, we fit an ARIMA (autoregressive, integrative, moving average) model (of order 10, 2, 2) to the time series of posterior probabilities in each cortical area, using the 'armax' function of the System Identification Toolbox in Matlab. ARIMA models are linear models that predict each value in a time series as a weighted sum of prior values and also prior errors in prediction, after differencing the time series to improve stationarity. Models of order 10, 2, 2 first difference the time series twice, then fit regression coefficients to the 10 prior values of the series and the 2 prior errors in estimation that provide the best estimate of each value in the series. The order of the ARIMA model employed was selected on the basis of preliminary testing because this order effectively eliminated autocorrelation in the data (**Supplementary Fig. 5**).

Linear regression analysis of preprocessed decoding time series to detect network transmission. We then applied a regression analysis to evaluate whether the residual posterior probability time series capturing rapid fluctuations in the strength of category signals significantly covaried in prefrontal and parietal cortex (**Supplementary Fig. 4d**). To evaluate top-down transmission, we regressed the posterior probability in each time bin of the parietal series onto the posterior probability bin in the prefrontal series shifted earlier by a variable lag of from 0 to 8 time bins (**Supplementary Fig. 4d**, left; the analysis at lag 1 is shown predicting each value of the parietal time series at time t as a function of the value of the prefrontal time series at time $t – 1$). To evaluate bottom-up transmission, we did the converse, regressing the posterior probability in each time bin in the prefrontal series onto the probability in a previous time bin in the parietal series shifted earlier by a variable lag of from 0 to 8 time bins (**Supplementary Fig. 4d**; right shows an example at lag 1). To evaluate changes in the strength of transmission over time, we advanced a sliding 500-ms window (consisting of 10 consecutive 50-ms bins) in 50-ms steps through the posterior probability time series in the two cortical areas, accumulated the posterior probabilities within the window over trials and ensembles and then performed the regression analysis on these accumulated data at each time step. The resulting time series of the F-statistic associated with the sliding window regression provided a time-resolved measure of signal transmission between the two cortical areas. We plotted the F-statistic at each time step at the leading edge of this sliding window. In order to compare transmission and decoding results directly, the decoding time courses in **Figure 3** were likewise averaged within a sliding window of 10 successive bins.

Bootstrap test of significance of transmission results. Transmission of category signals reflecting real-time physiological interaction between prefrontal and parietal cortex should only be detected in the case that the neural activity used for the analysis was recorded simultaneously in the two cortical areas. It was therefore important to show that transmission was not detected between prefrontal and parietal cortex when the analysis was based on neural signals recorded at different times in the two areas. To make that determination, we repeated the transmission analysis after shuffling trials in a bootstrap procedure in which we kept the residual posterior probability time series produced by the decoding and ARIMA steps intact. In the shuffling, we also kept neurons grouped into the same simultaneously recorded ensembles within each area that we had in the original analysis. Finally, we constrained the shuffling so that the spatial category associated with each posterior probability time series did not change as a consequence of the shuffling. Thus, trial shuffling broke the simultaneity of neural activity in prefrontal and parietal cortex but did not otherwise modify the posterior probability time series entered in the transmission analysis. The results of the transmission analysis after the trial-shuffling procedure estimated the degree of detected transmission that could be attributed to spurious factors, such as parallel entrainment of population activity in prefrontal and parietal cortex to behavioral events. We performed the regression analysis after shuffling posterior probabilities over trials in each time bin 500 times, generating a distribution of F-statistics and regression coefficients corresponding to the null hypothesis that the detected transmission reflected spurious factors, rather than actual signal transmission between simultaneously recorded neurons.

We considered transmission between prefrontal and parietal cortex to be significant at a given time bin if the F-statistic or regression coefficient derived from the original neural data recorded simultaneously in the two cortical areas exceeded the 95th percentile of the bootstrap distribution of the same statistic derived from the trial shuffled data in that bin.

In addition, we tested the significance of the difference between top-down and bottom-up transmission of category signals by shuffling posterior probability time series across cortical areas (100 iterations) and repeating the transmission analysis. Differences between the strength of top-down and bottom-up signal transmission were considered significant if the difference in the associated F-statistics derived from the original data exceeded the 95th percentile of the differences in F-statistics associated with top-down and bottom-up transmission in the bootstrap distribution.

Modulation of transmission by reward history. We investigated whether transmission of category signals varied as a function of behavioral performance. We did not undertake a comparison of category signal transmission between cortical areas on error and correct trials because few error trials were available, and the difference in statistical power to detect transmission between areas would confound interpretation of the outcome of the analysis. Rather, we compared the strength of category signal transmission between two groups of trials of equal size, differentiated on the basis on whether a given categorization rule had been rewarded with relatively high or low frequency within the last 10 trials. For each trial, we counted the number of trials within the last 10 trials that were rewarded under a given categorization rule. We divided trials into high recent reward and low recent reward groups on the basis of whether the number of rewards earned under a given categorization rule within the previous 10 trials was above or below the median value of rewards for this rule, over all trials. We then performed the lagged regression separately using high and low recent reward trial subsets (**Fig. 7**) and a larger sliding window of 1,000 ms (20 bins). We evaluated whether differences between transmission on high and low recent reward trials was significant by determining whether the difference in F-statistics associated with the two reward conditions in each time bin exceeded the 95th percentile of differences obtained in a bootstrap distribution after shuffling trials across reward status.