Neuron, Volume *83*

Supplemental Information

Complementary Roles for Primate Frontal and Parietal Cortex in Guarding Working Memory from Distractor Stimuli

Simon Nikolas Jacob and Andreas Nieder

Figure S1, related to Figure 2

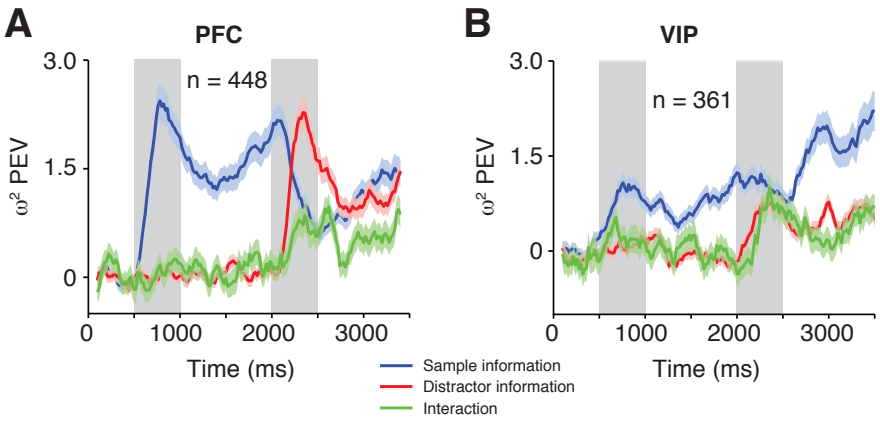

Figure S1 Explained variance analysis using two-way ANOVA. (**A, B**) Sliding window percent explained variance (ω^2 PEV) quantifying the information about the sample and distractor numerosity as well as the interaction term across task-related neurons recorded in prefrontal (A) and parietal cortex (B). Error bands, s.e.m. across neurons.

Figure S2, related to Figure 3

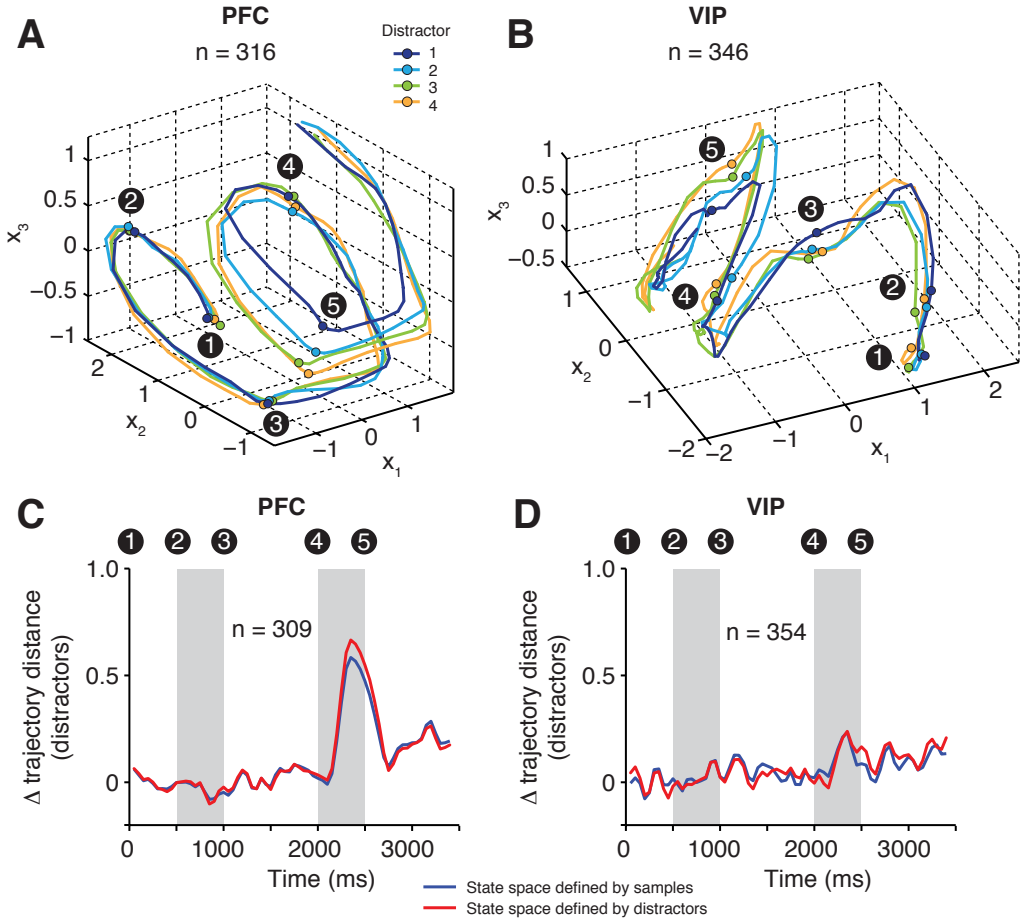

Figure S2 Neuronal population trajectories for distractor numerosities. (**A**) Factor analysis describing the state space of neuronal population activity in PFC (n = 316) for each distractor across time, plotted for the first three common factors. Time points mark the onset of the (1) fixation (pre-sample), (2) sample, (3) first memory, (4) distractor and (5) second memory period. Trajectories represent the mean across single trials. (**B**) Same analysis for the population of VIP neurons (n = 346). (**C**) Mean inter-trajectory Euclidean distance across all distractor-distractor combinations in PFC as a measure of neuronal stimulus selectivity (same population as in Fig. 3A). Blue and red curves represent the distance calculated using the state space defined by sample and distractor numerosities, respectively. (**D**) Same layout as in (C) for VIP (same population as in Fig. 3A).

Figure S3, related to Figure 3

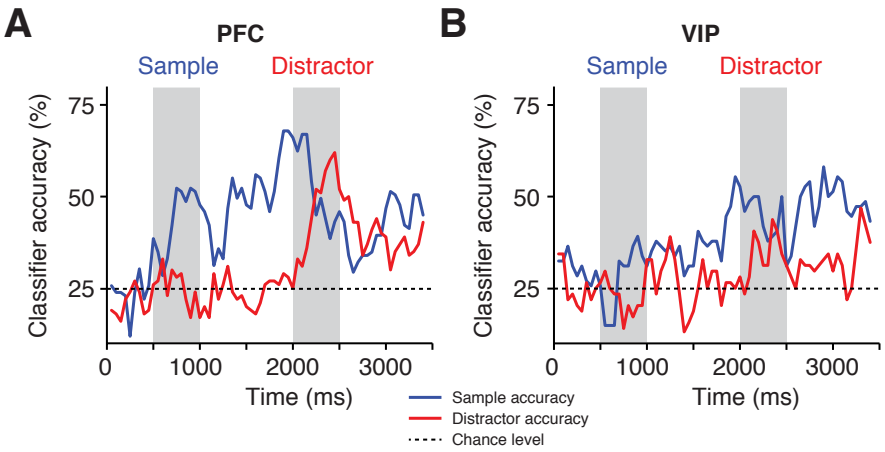

Figure S3 Single-trial decoding of sample and distractor numerosities. (**A**) Time-resolved single-trial decoding of sample and distractor numerosities from PFC population activity (n = 309 neurons). Trials were sorted according to sample or distractor numerosity and subjected to factor analysis (Fig. 3). For each single trial trajectory and time bin, the distance to the 4 mean trajectories was calculated. Classification accuracy reflects the percentage of trials that were correctly classified as being closest to the actual sample or distractor trajectory (chance level P = 0.25). (**B**) Same layout for the VIP population (n = 354 neurons).

Figure S4, related to Figure 4

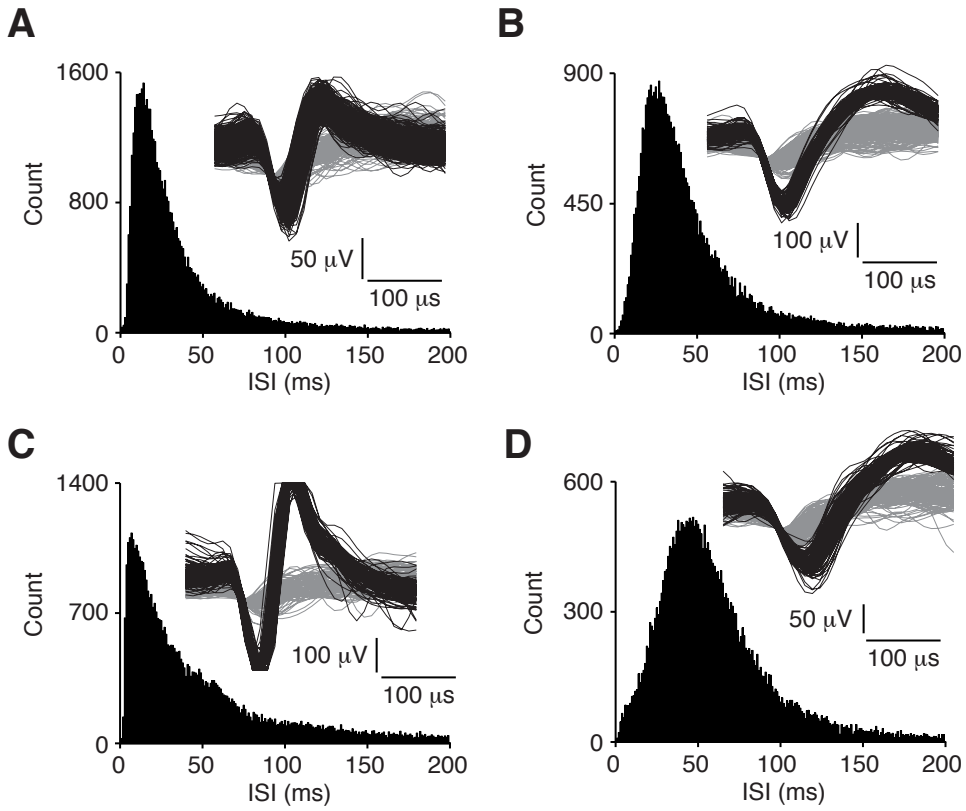

Figure S4 Extracellular single-unit recordings. (**A**) Distribution of inter-spike intervals for the single neuron in Fig. 4A together with action potential waveforms extracted from a representative 16 s extracellular voltage trace (inset; noise marked in gray). (**B**) Same layout for the single neuron in Fig. 4C. (**C**) Same layout for the single neuron in Fig. 4E. (**D**) Same layout for the single neuron in Fig. 4G.

Figure S5, related to Figure 4

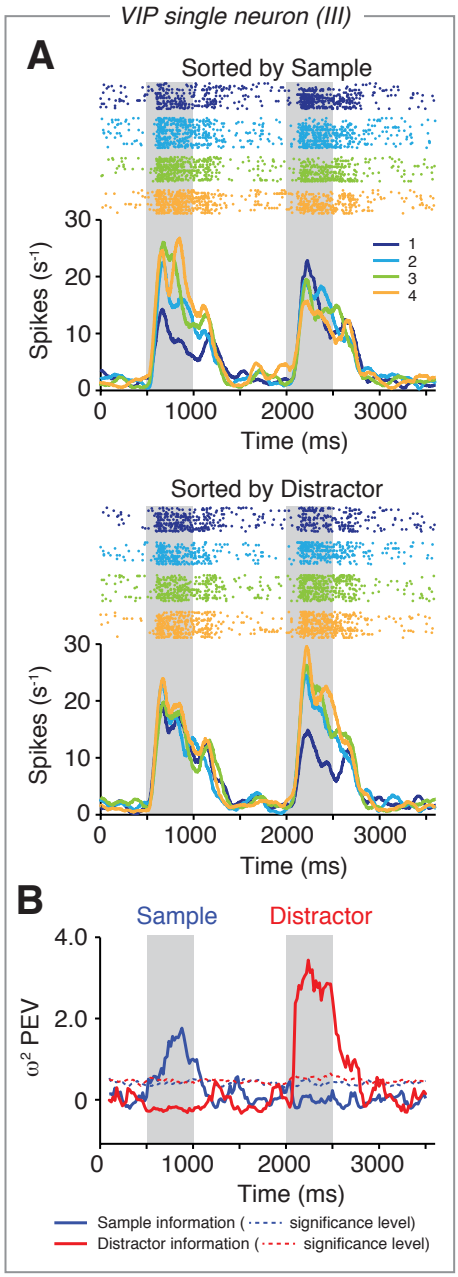

Figure S5 Rare VIP neuron encoding both sample and distractor. (**A**) Raster plots and spike density histograms for a single VIP neuron. Trials are sorted according to sample (top panel) or distractor numerosity (bottom panel). (**B**) Sliding window percent explained variance (ω^2 PEV) quantifying the information about the sample and distractor numerosity for the neuron in (A). Dashed lines mark the significance threshold (P = 0.01).

Figure S6, related to Figure 5

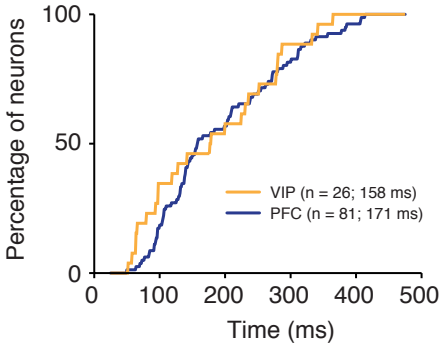

Figure S6 Time course of frontoparietal numerosity selectivity. Cumulative latency distribution of PFC and VIP neurons determined by a sliding window percent explained variance (ω^2 PEV) analysis quantifying the onset of sample information in the sample period.

Figure S7, related to Figure 6

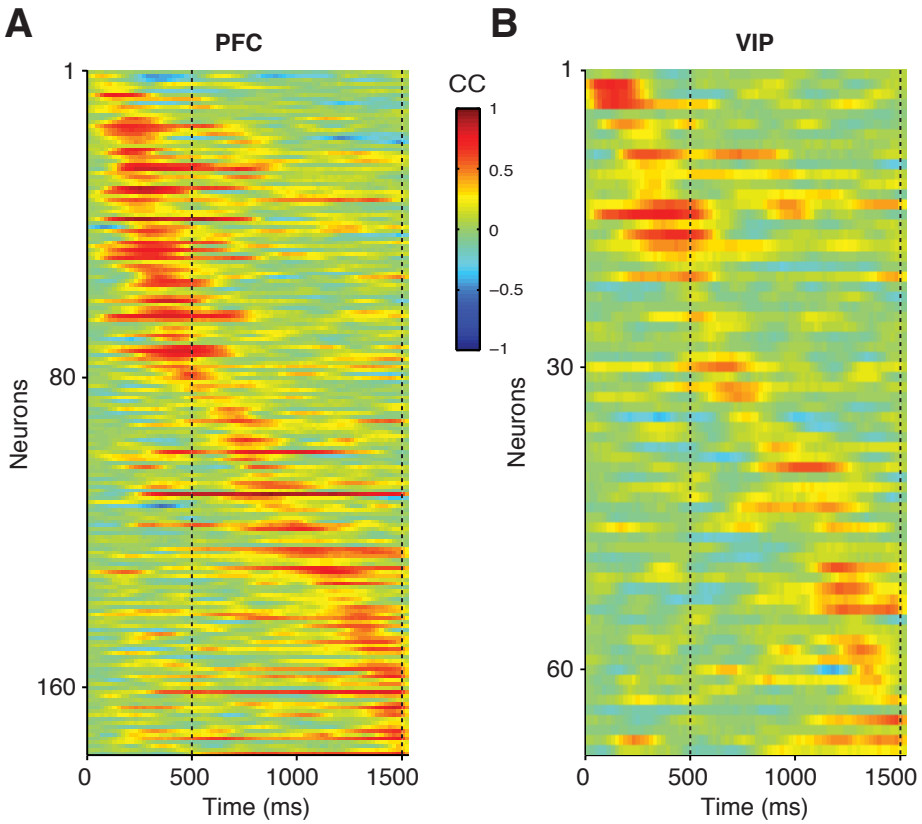

Figure S7 Sliding window cross-correlation. (**A, B**) Cross-correlation coefficients (CC) between sample and distractor tuning curves, derived as explained in Fig. 6A, for PFC (A) and VIP (B) neurons that were sample selective in the sample and/or first memory period (n = 177 and n = 68, respectively). Neurons are sorted by time of maximal correlation. Data are smoothed by a five-point running average rectangular filter. The first dashed line represents the offset of the stimulus (sample/distractor) period, the second dashed line indicates the offset of the memory (first/second) period.

Figure S8, related to Figure 7

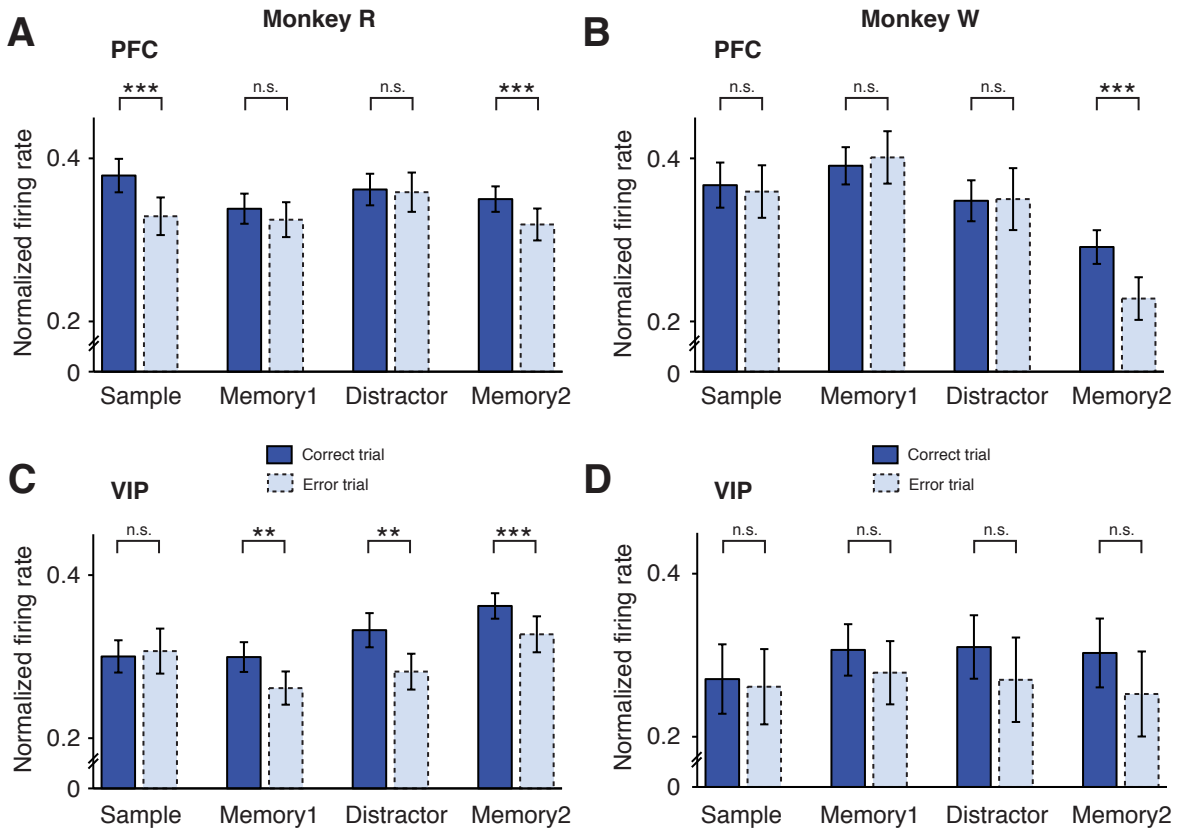

Figure S8 Error trial analysis. (**A**, **B**) Normalized firing rates for preferred numerosities in the sample, first memory, distractor and second memory periods for the same population of sample-selective PFC neurons as in Fig. 7A plotted separately for monkey R (A) and monkey W (B). Data are presented for correct trials (saturated colors, solid outlines) and for error trials (unsaturated colors, dashed outlines). (**C**, **D**) Same layout as in (A, B) for the same sample-selective VIP neurons as in Fig. 7C. Error bars, s.e.m. across neurons; **, $P < 0.01$; ***, $P < 0.001$; n.s., not significant.

Table S1, related to Figure 2

Table S1A. Number of task-related neurons showing selectivity in PFC.

	Sample	Memory1	Distractor	Memory2
Main effect of sample	88	98	43	73
Main effect of distractor			76	73
Interaction s. x d.			18	26

two-way ANOVA evaluated at $P < 0.01$

Table S1B. Number of task-related neurons showing selectivity in VIP.

	Sample	Memory1	Distractor	Memory2
Main effect of sample	27	36	24	64
Main effect of distractor			24	28
Interaction s. x d.			21	13

two-way ANOVA evaluated at $P < 0.01$

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Surgical procedures

Two adult male rhesus monkeys (*Macaca mulatta*, monkey R and monkey W) were implanted with a titanium head post and two right-hemispheric recording chambers centered over the principal sulcus of the lateral prefrontal cortex (PFC), anterior to the frontal eye fields, and over the ventral intraparietal area (VIP) in the fundus of the intraparietal sulcus guided by anatomical MRI and stereotaxic measurements. Chambers were angled to be able to penetrate the cortical surface perpendicularly. Surgery was conducted using aseptic techniques under general anesthesia. Structural magnetic resonance imaging to locate anatomical landmarks was performed before implantation. All experimental procedures were in accordance with the guidelines for animal experimentation approved by the local authority, the Regierungspräsidium Tübingen.

Task and stimuli

The monkeys were trained to match visually presented non-symbolic set sizes (numerosities) while suppressing a salient task-irrelevant, interfering numerosity (Fig. 1A). The animals grabbed a bar to initiate a trial and maintained eye fixation within 1.75 ° of visual angle of a central white dot. An infrared-based eye tracking system monitored ocular position (ISCAN, Woburn, MA). Trials were immediately aborted and excluded from further analysis if the animals broke fixation. Stimuli were presented on a centrally placed gray circular background subtending 5.4 ° of visual angle. Following a 500 ms pre-sample (pure fixation) period, a 500 ms sample stimulus containing 1 to 4 dots was shown. The monkeys had to memorize the sample numerosity for 2,500 ms and compare it to the number of dots (1 to 4) presented in a 1,000 ms test stimulus. Test stimuli were marked by a red ring surrounding the background circle. If the numerosities matched (50 % of trials), the animals released the bar (correct Match trial). If the numerosities were different (50 % of trials), the animals continued to hold the bar until the matching number was presented in the subsequent image (correct Non-match trial). Match and non-match trials were pseudo-randomly intermixed. Correct trials were rewarded with a drop of water. In 80 % of trials, a 500 ms interfering numerosity of equal numerical range was presented between the sample and test stimulus. The interfering numerosity was not systematically related to either the sample or test numerosity and therefore not

required to solve the task. In 20 % of trials, a 500 ms gray background circle without dots was presented instead of an interfering stimulus, i.e. trial length remained constant (control condition, blank). Trials with and without interfering numerosities were pseudo-randomly intermixed. Stimulus presentation was balanced. That is, across a set of trials in which a given numerosity was used as the sample, it was followed by all interfering numerosities with equal frequency. Similarly, a given interfering numerosity was preceded by all sample numerosities with equal probability. Thus, when trials are sorted by sample (or interfering numerosity), any measure of neuronal activity is directly attributable to that stimulus, because the influence of the interfering numerosity (or sample) is factored out across trials.

Both animals had previously been trained to full proficiency in the standard delayed-match-to-numerosity task without task-irrelevant numerosities. Over the course of 3 to 5 months, the interfering numerosity was slowly introduced while carefully monitoring behavioral performance. The red ring surrounding the test stimuli ensured that the animals did not confuse individual stimuli and almost never responded prior to presentation of the test numerosities. During training, there were no abrupt changes in performance to suggest that the animals had switched response strategies. Neuronal recordings commenced when performance in both animals was stable over several weeks.

Low-level, non-numerical visual features could not systematically influence task performance (Nieder et al., 2002): in half of the trials, dot diameters were selected at random. In the other half, dot density and total occupied area were equated across stimuli. CORTEX software (NIMH, Bethesda, MD) was used for experimental control and behavioral data acquisition. All stimuli were produced using MATLAB (The Mathworks, Natick, MA) and generated anew before every recording session to ensure that the animals could not solve the task by memorizing stimulus sequences.

Electrophysiology

In each recording session, eight 1 MΩ glass-isolated tungsten electrodes (Alpha Omega, Israel) per chamber were acutely inserted through an intact dura with 1 mm spacing using custom-made screw microdrives. Brain tissue was allowed to settle for at least 30 minutes before recording. Stable and well-isolated neurons were recorded at random (Fig. S4); no attempt was made to preselect neurons according to particular response properties. To access area VIP at the fundus of the intraparietal sulcus (IPS), electrodes were passed along the course of the IPS to a depth of 9 to

13 mm below the cortical surface (Nieder and Miller, 2004; Nieder et al., 2006; Vallentin et al., 2012). Prior to recording neuronal activity in VIP, correct positioning of the electrodes was ensured by physiological criteria (response to moving visual stimuli and tactile stimulation). Signal acquisition, amplification, filtering and digitalization were accomplished with the MAP system (Plexon, Dallas, TX). Timestamps of trial events and action potentials were extracted for analysis. Waveform separation was performed offline using a combination of principal component analysis of waveform traces and other properties of the recorded waveforms (amplitude, peak/trough; Offline Sorter, Plexon).

Data analysis

Data analysis was performed with MATLAB. None of the reported analyses depended on the exact choice of trials to include or time windows to analyze. Repeating analyses with a different set of parameters yielded comparable results.

Unless specified otherwise, neurons were included in the analysis if the following criteria were met: first, their average firing rate across trials was at least 1 spike/s; second, they were recorded for at least 1 correct trial in all 20 conditions (4 sample numerosities x 5 interfering numerosities including the control [0] condition); and third, they modulated their firing rate in the course of the trial (task-related neurons, one-way analysis of variance (ANOVA) with average firing rates in the pre-sample (fixation), sample, first memory, interfering stimulus, second memory periods; evaluated at $P < 0.05$). A total of 448 neurons in PFC and 361 neurons in VIP fulfilled these criteria.

Behavioral data

Behavioral tuning functions were used to describe the percentage of trials (y axis) for which a test stimulus (x axis, units of numerical distance to sample numerosity) was judged as being equal in number to the sample. A numerical distance of 0 denotes match trials; the data point represents the percentage of correct trials. As the numerical distance increases, there is less confusion of the test with the sample numerosity; the data points represent the percentage of error trials. Tuning curves were calculated separately for trials without interfering numerosities (control (blank); Fig. 1C, D, H, I), for trials in which the sample and interfering numerosity were equal ('repeat sample' condition, Fig. 1C, H) and for trials in which the sample and interfering numerosity were different ('distractor' condition, Fig. 1D, I). For each condition and session, Gaussian curves were fitted to the tuning functions. The

location of the mean (0) and the amplitude (percent correct trials) were fixed, the Gaussian's width (sigma) was free. Differences in width to the control condition were averaged across sessions and tested for significance against zero median using Wilcoxon signed rank tests.

For each condition and session, reaction times (RT) were determined for match trials only (in non-match trials, the second test image following the non-match was always a match and therefore predictable). RT differences to the control condition were averaged across sessions and tested for significance against zero median using Wilcoxon signed rank tests.

Neuronal information

To quantify the information about the sample or interfering numerosity that was carried by a neuron's firing rate, we used the percent explained variance (ω^2 PEV) measure (Buschman et al., 2011; Hentschke and Stüttgen, 2011; Puig and Miller, 2012). ω^2 reflects how much of the variance in a neuron's firing rate can be explained by the numerosity of a particular stimulus. It was calculated using

$$\omega^2 = \frac{SS_{Groups} - df * MSE}{SS_{Total} + MSE},$$

where the individual terms are derived from a one-way categorical ANOVA (Buschman et al., 2011; Siegel et al., 2009): SS_{Groups} denotes the sum-of-squares between groups (numerosities), SS_{Total} the total sum-of-squares, *df* the degrees of freedom, and *MSE* the mean squared error. The number of trials in each group was balanced. Balancing was accomplished by stratifying the number of trials in each group to a common value: A random subset of trials was drawn (equal to the minimum trial number across groups) and the statistic was calculated. This process was repeated 25 times, and the overall statistic was taken to be the mean of the stratified values. ω^2 is an unbiased, zero-mean statistic when there is no information: when we recomputed the analysis after subtracting the shuffled PEV (Siegel et al., 2009; Puig and Miller, 2012), we obtained identical results (not shown). For every neuron and analysis window (200 ms duration, 20 ms step), ω^2 shuffle was calculated after randomly shuffling the association between firing rate and numerosity. This process was repeated 1,000 times, and the shuffled PEV was taken to be the mean of the individual values. Neurons were separately tested for sample and interfering numerosity PEV. Bin-wise Wilcoxon paired signed rank tests

(evaluated at $P < 0.05$ and $P < 0.01$) were used to compare sample and interfering numerosity PEV within the population of PFC and VIP neurons. Repeating the analysis using a two-way categorical ANOVA with factors sample and interfering numerosity yielded identical results (Fig. S1).

To determine at which point in the trial a neuron carried significant information about either the sample or interfering stimulus (Fig. 5), we used a permutation test in a sliding window analysis. For every analysis window (200 ms duration, 20 ms step), we created a null distribution of PEV values by randomly shuffling the association between firing rates and numerosities and calculating ω^2. This process was repeated 1,000 times. The significance threshold for the amount of information in any given window was set to $P < 0.01$ (one-sided), i.e. the actual PEV value was required to be larger than 99 % of values in the null distribution. To control for multiple comparisons, a neuron was said to significantly encode the sample and/or interfering stimulus if it crossed the respective thresholds for 5 consecutive windows. The onset of the first of these windows was taken to be the neuron's response latency. For the high-resolution comparison of sample selectivity latency in PFC and VIP (Fig. S6), we restricted the analysis to the sample phase (50 ms window, 1 ms step, 25 consecutive windows).

Because the animals performed significantly above chance level, we obtained considerably less error trials than correct trials. Therefore, we could not record any neurons that were present for at least 1 error trial in all possible trial conditions. For the analysis of ω^2 PEV in error trials, we therefore relaxed the selection criteria and included neurons that were recorded during at least 4 error trials across each sample and interfering numerosity. 76 % (n = 182/239) of sample- or distractor-selective neurons in PFC and 75 % (n = 114/152) of sample- or distractor-selective neurons in VIP fulfilled these criteria. For every neuron included, correct and error trial ω^2 PEV values were averaged for the sample, first memory, interfering stimulus and second memory periods and compared using a Wilcoxon paired signed rank test.

In the same population of neurons, we performed an error trial analysis with firing rates to preferred numerosities, defined as the numerosity that elicited maximal firing in a 500 ms time window aligned to an individual neuron's sample selectivity latency. Firing rates in correct and error trials were normalized (maximum and minimum firing rate set to 1 and 0, respectively), averaged for the sample, first memory, interfering

stimulus and second memory periods and compared using Wilcoxon paired signed rank tests. The analysis was performed separately in both monkeys.

Stimulus-selectivity index

We calculated a time-resolved stimulus-selectivity index (SSI) to determine whether a neuron's firing rate carried more information about the sample or interfering numerosity. The SSI was determined for all bins where a neuron significantly encoded the sample and/or interfering numerosity (see Neuronal Information). It was calculated by

$$SSI = \frac{SS_{Groups,\ Sample} - SS_{Groups,\ Distractor}}{SS_{Total}},$$

where the terms are derived from one-way ANOVAs: $SS_{Groups,\ Sample}$ and $SS_{Groups,\ Distractor}$ denote the sum-of-squares between groups (numerosities) when trials are sorted by sample or interfering numerosity, respectively, and SS_{Total} is the total sum-of-squares of either ANOVA (identical values for sorting by sample or interfering stimulus). Positive values indicate that the neuron's firing rate carries more information about the sample numerosity; negative values indicate that discharge rates vary more strongly with the interfering numerosity. For illustration purposes, the SSI was normalized by the maximum of the absolute values of each neuron.

Tuning curve cross-correlation

Cross-correlation of neuronal tuning curves (Diester and Nieder, 2007) provided a measure to quantify the extent to which single neurons switched from encoding the sample to representing interfering numerosities. We sorted trials by sample numerosity in the sample and first memory period and by interfering numerosity in the interfering numerosity and second memory period (Fig. 6A). Tuning curves were derived from time windows (200 ms duration, 20 ms step) at equivalent positions in the sample-sorted and interfering-numerosity-sorted part of the trial and cross-correlated by calculating

$$CC = \frac{\sum_{n=1}^{4}\left(FR_{Sample}(n) - \overline{FR}_{Sample}\right) * \left(FR_{Distractor}(n) - \overline{FR}_{Distractor}\right)}{\sqrt{\sum_{n=1}^{4}\left(FR_{Sample}(n) - \overline{FR}_{Sample}\right)^2} * \sqrt{\sum_{n=1}^{4}\left(FR_{Distractor}(n) - \overline{FR}_{Distractor}\right)^2}},$$

where $FR_{Sample}(n)$ is the average firing rate for numerosity $n \in \{1,2,3,4\}$ when trials were sorted by sample, $\overline{FR}_{Sample}$ the average firing rate across all numerosities when

trials were sorted by sample, $FR_{Distractor}$ (n) the average firing rate for numerosity n ∈ {1,2,3,4} when trials were sorted by interfering numerosity, and $\overline{FR}_{Distractor}$ the average firing rate across all numerosities when trials were sorted by interfering numerosity. Thus, the higher the cross-correlation coefficients (CCs), the more a neuron switched from encoding the sample to representing the interfering numerosity.

CCs were calculated for neurons that were selective for the sample numerosity in the sample and/or first memory period as determined by a one-way ANOVA ($P < 0.01$; n = 177 in PFC, n = 68 in VIP). A random subset of trials (equal to the minimum trial number across numerosities, 8 trials at most) was drawn for each numerical value and time window (sorted as described above). Thus, trial numbers were equated for all neurons, both regions, and sample and interfering numerosities. Tuning functions were built with the averaged firing rates of these trials, and the CC was calculated. This process was repeated 25 times, and the overall statistic was taken to be the mean of the stratified values. CCs for PFC and VIP neurons were compared using Wilcoxon rank sum tests. To control for differences in cell counts or coding strength, we calculated CCs between tuning curves derived for the sample numerosity in all epochs using the same time windows as for the sample/interfering-numerosity CCs. Shuffle predictors for each region were calculated by creating a null distribution of CCs by randomly shuffling the association between firing rates and numerosities (1,000 repetitions). Shuffled values were centered on zero (mean < 10^{-3}). Fig. 6B, C shows mean values plus 3 standard deviations (across neurons and time windows).

Factor analysis

Factor analysis was used to describe the temporal evolution of activity in a large ensemble of recorded neurons, i.e. the population dynamics. Factor analysis extracts low-dimensional neuronal trajectories from noisy spiking activity and represents these in state space, where each data point corresponds to the instantaneous firing rate of the population of neurons at a given point in time. The trajectory is the path that links the sequence of activation states at each time point (Stokes et al., 2013). Compared to PCA, factor analysis better captures shared fluctuations in activity across the population and more effectively discards independent variability specific to individual neurons (Harvey et al., 2012). Analysis was performed using MATLAB toolboxes (Yu et al., 2009). Firing rates in correct trials were processed using 5 latent dimensions, a bin width of 50 ms and a Gaussian smoothing kernel of 50 ms.

Optimal parameters were determined using cross-validation routines from the toolbox (Yu et al., 2009). To obtain a comparable number of PFC and VIP neurons for the analysis of sample trajectories, we included PFC neurons with at least 33 trials per sample numerosity (n = 309) and VIP neurons with at least 19 trials per sample (n = 354). For the analysis of interfering numerosity trajectories, we included PFC neurons with at least 24 trials per interfering numerosity (n = 316) and VIP neurons with at least 15 trials per interfering numerosity (n = 346). These neurons formed populations of pseudo-simultaneously recorded neurons. For each numerosity, we picked the minimum number of trials shared across neurons from one recording area and calculated single-trial population trajectories. Trajectories were then averaged across trials to represent a mean population response for the given numerosity. The analysis was performed separately for trials sorted by sample and by interfering numerosity. To determine how well the state space defined by sample numerosities captured variability in the neuronal responses to interfering numerosities, we repeated the analysis using the population of neurons used in the sample trajectory analysis (PFC: n = 309; VIP: n = 354) and calculated interfering numerosity trajectories in sample space.

The Euclidean distance between two trajectories at corresponding time points was used as a measure of the difference in the population's activation state between individual numerosities. For each pair, the distance in the pre-sample (fixation) period was used as a baseline and subtracted from distances calculated at subsequent time points.

Classification of single trials was performed based on distances to the individual numerosities' mean trajectories. Trials were sorted either according to sample or to interfering numerosity. For each trial and time point, the distance between the test and each of the 4 reference trajectories was calculated. A trial was classified correctly if the test trajectory's numerosity was equal to the numerosity of the closest reference trajectory. Classification performance is expressed as the mean accuracy across trials (chance level 25 %). Classification was performed using leave-one-out cross validation, i.e. mean reference trajectories were calculated excluding the test trajectory (Harvey et al., 2012).