

## COGNITIVE NEUROSCIENCE

# Stable numerosity representations irrespective of magnitude context in macaque prefrontal cortex

Maria Moskaleva and Andreas Nieder

Animal Physiology, Institute of Neurobiology, Auf der Morgenstelle 28, University of Tübingen, 72076 Tübingen, Germany

*Keywords:* frontal lobe, non-human primate, numerical cognition, single-cell recordings

## Abstract

Cognitively demanding tasks require neurons of the prefrontal cortex (PFC) to encode divergent behaviorally relevant information. In discrimination tasks with arbitrary and learned categories, context-specific parameters shape and adapt the tuning functions of PFC neurons. We explored if and how selectivity of PFC neurons to visual numerosities, a 'natural' abstract category, may change depending on the magnitude context. Two monkeys discriminated visual numerosities (varying numbers of dot items) in a delayed match-to-sample (DMS) task while single-cell activity was recorded from the lateral PFC. During a given recording session, the numerosity task was either presented in isolation or randomly intermixed with DMS tasks with line lengths and colors as discriminative stimuli. We found that the context of numerosity discriminations did not influence the response properties of numerosity detectors. The numerosity tuning curves of selective neurons, i.e. the preferred numerosity and the sharpness of tuning, remained stable, irrespective of whether the numerosity task was presented in a pure numerosity block or a mixed magnitude block. Our data suggest that numerosity detectors in the PFC do not adapt their response properties to code stimuli according to changing magnitude context. Rather, numerosity representations seem to rely on a sparse and stable 'labeled line' code. In contrast to arbitrarily learned categories, numerosity as a 'natural' category may possess a privileged position and their neuronal representations could thus remain unaffected by magnitude context.

## Introduction

The prefrontal cortex (PFC) operates at the apex of the cortical hierarchy and enables primates with unprecedented cognitive flexibility (Miller & Cohen, 2001). Cognitively demanding tasks engage many neurons of the PFC, the brain's central executive, to encode divergent information in different tasks or across different task periods. Which coding strategies does the PFC employ to cope with this information-processing challenge?

The 'adaptive coding hypothesis' posits that neurons in the PFC are not inherently tuned to specific stimulus features, but rather adapt their response properties to code stimuli according to task relevance. Within this framework, context-specific parameters shape the tuning functions of PFC neurons (Duncan, 2001). Consequently, context shifts the way stimuli are encoded by single neurons within the PFC network. Adaptive coding has been demonstrated several times in monkeys learning to categorise stimuli (Freedman *et al.*, 2002; Cromer *et al.*, 2010; Roy *et al.*, 2010) and during flexible decision-making tasks (Wallis *et al.*, 2001; Bongard & Nieder, 2010; Merten & Nieder, 2012, 2013; Valentin *et al.*, 2012; Eiselt & Nieder, 2013; Stokes *et al.*, 2013).

In contrast to arbitrarily learned categories that seem to become encoded on demand, certain 'natural' categories may possess a privileged position and their neuronal representations could remain unaffected by the magnitude context. Abstract numerical quantities

represented in a dedicated fronto-parietal network might belong to this group of 'natural' categories. Neurons in macaque PFC readily encode visually (Nieder *et al.*, 2002; Nieder & Merten, 2007; Nieder, 2013) and auditorily (Nieder, 2012) presented numerosities, and maintain them online during delay periods. Recently, numerosity-tuned neurons have been described in untrained monkeys. While these numerically naïve monkeys were engaged in a color discrimination task, neurons in the lateral PFC (and the posterior parietal cortex) responded selectively to the number of the colored items presented (Viswanathan & Nieder, 2013). Together with the psychophysical finding that numerical judgments are susceptible to adaptation in the same way as visual properties (Burr & Ross, 2008) or faces (Webster & MacLeod, 2011), this suggests a 'sense of number' (Danzig, 1930; Dehaene, 1997). In other words, numerosity seems not to be a learned category, but rather a stimulus feature, which is spontaneously and naturally encoded within visual neural structures of the primate brain. If this is true, numerosity representations are expected to remain unaffected by changes of the magnitude context in which they need to be discriminated.

To address this question, we investigated the coding properties of numerosity-selective PFC neurons in different magnitude contexts. Two monkeys were trained on a visual delayed match-to-numerosity task and single-cell recordings were done from the lateral PFC. Within a given recording session, the numerosity task was either presented in isolation (pure numerosity block condition) or embedded in equivalent delayed match-to-sample (DMS) tasks with other magnitudes (line lengths and colors) as discriminative stimuli (mixed

*Correspondence:* Dr A. Nieder, as above.  
E-mail: andreas.nieder@uni-tuebingen.de

Received 5 September 2013, revised 28 October 2013, accepted 11 November 2013

magnitude block condition). By comparing the proportion and tuning properties of numerosity-selective neurons in the respective conditions, the outlined alternative coding hypotheses could be tested.

## Materials and methods

### Animals

The subjects were two adult male rhesus monkeys, *Macaca mulatta* (monkey H: 8 kg; monkey L: 7 kg). The monkeys were housed in small social groups. Both animals had experience with color and numerical stimuli, monkey H also with line stimuli, from previous experiments. The monkeys worked under a controlled fluid access protocol and received liquid rewards for correct responses. All procedures were in accordance with the guidelines for animal experimentation, approved by the authority, the Regierungspräsidium Tübingen, Germany. All experiments were carried out in accordance with the European Communities Council Directive of 24 November 1986 (86/609/EEC).

### Behavioral protocol

The monkeys were trained to perform a DMS task, with numerosities, lines of different lengths and colors as stimuli. The trial began when the monkeys grabbed a bar inside their primate chair and faced the screen. When the fixation period started, the monkeys were required to fixate a white dot superimposed on a gray circle [fixation window:  $3.5^\circ$  visual angle (VA)]. After a fixation of 500 ms, a sample, either a numerosity, a line or a colored ring (example with numerosity two in Fig. 1A), was presented for 800 ms. During a following 1000-ms delay period, the monkeys needed to maintain fixation and memorise the sample. After the delay period, a test image was presented. In half of the cases, the first test image (Test 1) matched the sample image and the monkeys were required to release the bar to receive a water reward. In 50% of the trials, Test 1 did not match the sample. In this case, the monkeys were required to keep holding the bar until after 1200 ms a second test image (Test 2) was presented, which always matched the sample. Thus, chance performance was 50% correct trials.

### Behavioural protocol and stimuli

The monkeys were trained to perform a DMS task with different visual sample stimuli (Fig. 1A). Three different kinds of stimuli were used: numerosities (numbers of black dots); length (lines or different lengths); and color (colored rings; Fig. 1B). These stimuli were presented in two different trial blocks, the 'pure numerosity block' and the 'mixed magnitude block' (Fig. 1C).

In the pure numerosity block, exclusively numerosity trials were presented. This block contained 48 trials, 16 for each sample numerosity, in pseudorandom order. Numerosity stimuli were one, two or four black dots superimposed on a gray circle (Fig. 1B). The individual item's position and size were varied pseudorandomly. To control for low-level features, two different stimulus protocols, the 'standard protocol' and the 'control protocol', were used. The dots in standard stimuli had diameters between  $0.55$  and  $0.95^\circ$  of VA. In the control stimulus protocol, parameters co-varying with changes of the number of items (density and total dot area) were equalised for the different samples. The control dots had diameters between  $0.7^\circ$  VA and  $1.55^\circ$  VA. Standard and control numerosity stimuli were presented with equal probability ( $P = 0.5$ ) and in pseudorandom order in the pure numerosity block.

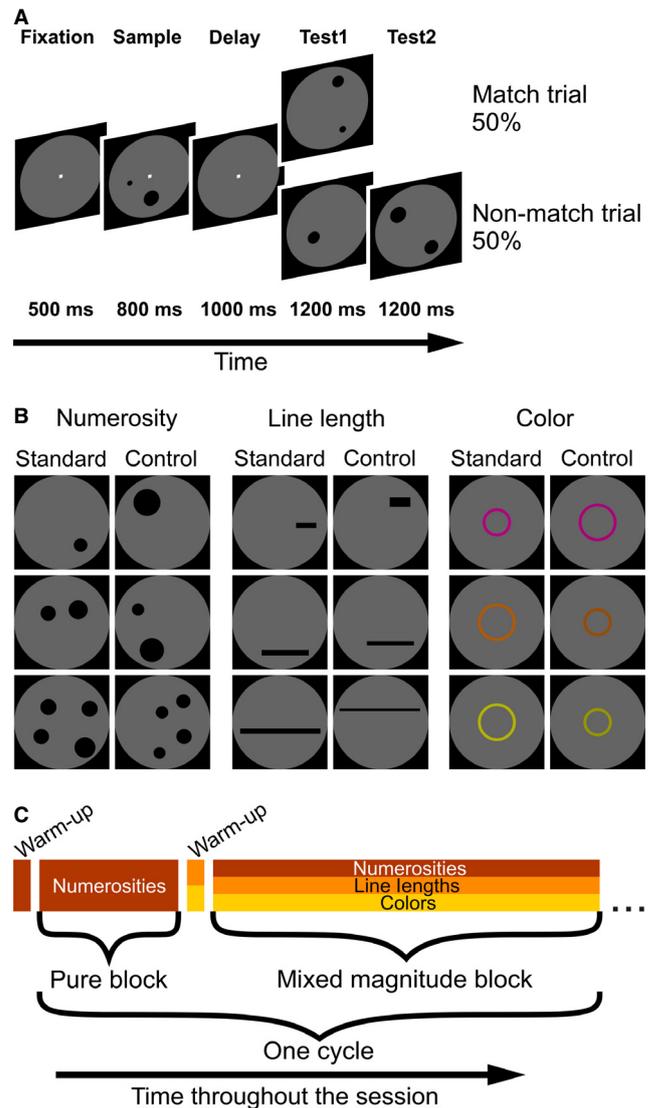


FIG. 1. Behavioral protocol. (A) DMS task. Monkeys were required to grasp a bar and to maintain fixation, after a period of 500 ms a sample appeared. After a delay of 1000 ms the monkeys viewed a test display. If the test matched the sample in the relevant feature (here: number of dots) the monkeys were required to release the bar. (B) Example stimuli: six sets of stimuli were used in every session: numerosity stimuli (one, two or four dots); line length stimuli ( $1.1$ ,  $2.6$  and  $4.5^\circ$ VA) and colored stimuli (red, orange and yellow). Control stimuli were equalised for total black area (lines and numerosities), density (numerosities) or luminance (colors). Color stimuli were presented in two sizes: small rings as sample; and large rings as test images. (C) Blocks of different trials with different magnitudes within one experimental session. To complete one session the monkeys were required to complete at least one repetition of the pure numerosity block (48 trials) and the mixed magnitude block (144 trials of numerosity, line and color trials in pseudorandom order). These blocks were interleaved by warm-up blocks, which indicated to the monkeys which block is about to start.

In the 'mixed magnitude block', numerosity trials made up only one-third of all trials and were pseudorandomly intermixed with trials in which the monkey had to discriminate line lengths (the second third of the trials) and colors (the last third of the trials) as alternative magnitudes. The mixed magnitude block contained 144 pseudorandomised trials with all three magnitudes (numerosities, lines and colors) as stimuli. Again, 16 trials per sample magnitude were presented. All stimuli were again presented in two protocols: the standard and the control. In the standard protocol, the magnitude

(numerosity, length and color) varied at the expense of some co-varying low-level visual features (see Nieder *et al.*, 2002). In the control stimulus protocol, parameters co-varying with changes of the magnitudes (density and total dot area for numerosities, total area and luminance for lines, and luminance for colors) were equalised for the different samples.

The line stimuli consisted of horizontal lines of three different lengths (1.125° VA, 2.625° VA and 4.5° VA), which were positioned pseudorandomly inside the gray circle. Standard lines had all the same thickness of 0.26° VA. Control lines had the same area irrespective of their length. Hence, they were varied in their thickness (0.525° VA, 0.225° VA and 0.1312° VA).

The color stimuli were monochromatic rings (annuli). The monkeys were required to match the color of the stimulus and to ignore the luminance. The annuli were always positioned in the middle of the gray circle. Two different sizes were used. Sample stimuli had rings with an outer diameter of 1.575° VA. To prevent adaptation effects on the retina, the rings in test stimuli were bigger and had outer diameters of 2.1° VA. The used colors were red, orange and yellow. Standard stimuli had bright colors that varied in their luminance. Control stimuli were adjusted in their color to have the same luminance of  $10.6 \pm 0.13$  cd/m<sup>2</sup>, measured with LS-100 luminance meter (Konica Minolta).

Both the pure numerosity block and the mixed magnitude block were preceded by short warm-up blocks, which were discarded from the analysis. The warm-up blocks had six trials with standard stimuli each. The pure numerosity warm-up block consisted of two trials with each sample numerosity. The mixed magnitude warm-up block had three line and three color trials. The warm-up blocks were used to signal to the monkey whether it had to attend to only the numerosities (pure numerosity block) or to all possible stimulus magnitudes (mixed magnitude block). To successfully complete an experimental session, the monkeys were required to complete all these four blocks (pure numerosity block and the mixed magnitude block plus one warm-up block for each) at least once. To prevent possible sequence effects, the session started with either numerosity or the mixed magnitude blocks on alternating days.

To prevent the monkeys from memorizing the visual characteristics of the displays, all stimuli with randomised features (numerosities and lines) were generated anew every day (20 images per sample and stimulus protocol), for each experimental session. In each trial, sample and test displays never showed the same image. Every magnitude was presented in a balanced way as sample and as test in control and standard conditions.

### Electrophysiological recordings

Before the experiment, the monkeys were implanted with a titanium head bar for head fixation and with a recording well, located over the right dorso-lateral PFC and centered over the principal sulcus. All surgeries were performed under general anesthesia.

Extracellular single-cell activity was recorded using arrays of 8–12 1-M $\Omega$  glass-insulated tungsten electrodes, which were lowered into the brain every day. The recorded neurons were not preselected for task-selectivity. Signals were amplified and digitised using the Multichannel Acquisition Processor (Plexon). All single units were sorted offline (Plexon).

### Data analysis

Overall the two monkeys completed 60 recording sessions (monkey H: 32 sessions; monkey L: 28 sessions). The behavior was analysed

over this entire recording period. The analysis of neural and behavioral data was performed using custom-written MatLab software (version 2011b). Significance level for all tests was  $P < 0.01$ .

The percentage of correct performance per session for a given magnitude (e.g. numerosities) was averaged over all sessions. Paired Wilcoxon tests were conducted to compare the performances under the standard and control protocols, for numerosities, line lengths and colors, respectively. For comparisons between blocks, the performance was averaged over the stimulus protocol and recording sessions. A two-way ANOVA (factors: sample numerosity and block type) was used for these comparisons.

For neural data, only single units with discharge rates above 1 Hz were analysed, if they were present for at least one complete cycle of both main blocks (Fig. 1C). If a cell was recorded for more than one complete cycle, the additional trials were truncated to the same number in all conditions. Hence, for a given cell, the same number of trials was analysed for every sample condition. The analysis included only correct trials.

The analysis of the neuronal data was conducted for two time periods, the sample and the delay phase. The sample phase began 100 ms after the sample onset and was 800 ms long. The delay phase was also 800 ms long and began 200 ms after delay onset. In these periods, the average discharge per time (discharge rate) was calculated in each trial. To determine magnitude-selective cells, these data were analysed with a two-way ANOVA for the mixed magnitude block, with the factors being stimulus protocol and sample magnitude, for numerosity, line length and color trials separately. Numerosity-selectivity and block effects were assessed in a separate three-way ANOVA with the factors sample size, block condition and stimulus protocol. Cells that showed a main effect of the protocol or interaction with it in either the sample or the delay phase were not analysed further in this phase.

Visual and selectivity latencies were determined for all cells that were numerosity-selective in the sample phase. Visual response latency was defined by the first of five consecutive 10-ms time bins (slid in 1-ms increments) that reached 3 SDs above baseline discharge rate (average activity during the period of 250 ms, starting at fixation onset). Latencies below 50 ms and above 400 ms were discarded. The latency of numerosity-selectivity was measured by a sliding Kruskal–Wallis test (kernel bin width 50 ms, slid in 1-ms increments). Numerosity-selectivity latency was defined by the first time bin, at least 50 ms after sample onset, where the test showed significant differences ( $P < 0.01$ ) in response to one of the three numerosities.

To analyse numerosity-selectivity, tuning functions were created for each cell in the two blocks and the two analysis windows by averaging the discharge rate over the trials for the different sample conditions. The sample that elicited the highest discharge rate in a cell was called the 'preferred' sample of this neuron.

A population analysis was conducted with all cells that were determined as numerosity-selective by the three-way ANOVA. A population peri-stimulus time histogram (PSTH) was created using normalised discharge rates (normalisation: difference in SDs from the average baseline discharge rate). For further analysis, numerical distance functions with normalised discharge rates were created (Fig. 4C and F). The discharge rate for the preferred numerosity during the pure numerosity block was defined as 100% and the lowest discharge rate as 0%. Discharge rates to the second preferred sample and all samples in the mixed magnitude block were normalised accordingly to these values. These normalised discharge rates were plotted against the numerical distance to the preferred quantity of this cell and averaged over all numerosity-selective cells. The

numerical distance functions for the mixed magnitude and the pure numerosity block were compared by a Wilcoxon test for each distance to preferred number (the significance level was Bonferroni-corrected).

To assess possible small, subthreshold effects of block, numerosity-selective neurons were assigned selectivity indices (SIs) for the two blocks. The index was calculated as follows:

$$SI = \frac{\text{discharge rate}_{\text{preferred}} - \text{discharge rate}_{\text{nonpreferred}}}{\text{discharge rate}_{\text{preferred}} + \text{discharge rate}_{\text{nonpreferred}}}$$

To investigate the relationship of SIs for individual neurons during the pure numerosity and mixed magnitude block, the selectivity values in the pure numerosity block were plotted as a function of the selectivity in the mixed magnitude block. The distance of the resulting dots to the bisection line was calculated. Dots with positions above the bisection line (higher selectivity in the pure numerosity block than in the mixed magnitude block) were assigned negative distance values, and dots with positions below the bisection line (higher selectivity in the mixed magnitude block than in the pure numerosity block) were assigned positive distance values. The symmetry around zero of the distance distribution was assessed by a signed-rank test. Hartigan's dip test (Hartigan & Hartigan, 1985) was used to test for bimodality.

## Results

### Behavioral performance

The monkeys were trained to perform a numerosity DMS task in two blocked conditions within one session, 'pure numerosity block' and a 'mixed magnitude block' (see Materials and methods). Figure 2A shows the comparable average performances of the monkeys for the different magnitudes and stimulus protocols in the mixed magnitude block. The monkeys were very proficient, with average performances of over 90% correct in all conditions. The performance in the different magnitude conditions was very similar, but still significantly different (Kruskal–Wallis test  $\chi^2_2 = 25.36$ ,  $P < 0.001$ ). The effect of stimulus protocol was not significant for numerosity and color stimuli (Wilcoxon test,  $Z = 2.57$ ,  $P = 0.0102$ ;  $Z = 0.92$ ,  $P = 0.36$  for numerosities and color stimuli, respectively). With a negligibly small effect size (3.7%), the only significant effect of protocol was found for line stimuli ( $Z = 5.36$ ,  $P < 0.001$ ). We therefore pooled the data over the two stimulus protocols for further analysis.

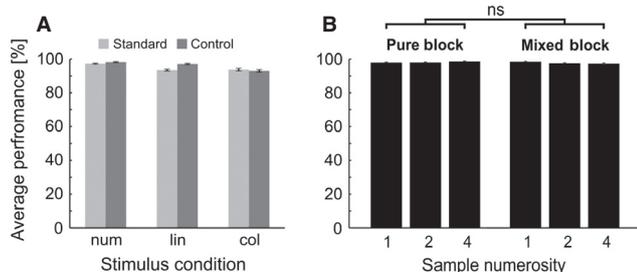


FIG. 2. Behavioral performance. (A) Average performance of both monkeys in the mixed magnitude block in the standard and control stimulus protocols during the recording sessions. (B) Average performance of both monkeys in numerosity trials in both the mixed magnitude and the pure numerosity block, pooled over standard and control stimulus protocols. Chance level was 50%. All error bars: standard error of the mean (SEM).

To test whether the addition of different magnitude trials had an impact on performance in the numerosity trials, we compared the pure numerosity block and the mixed magnitude block (numerosity conditions only; Fig. 2B). The monkeys discriminated equally well numerosities in the different block conditions, with monkey H performing at 97.2% and 98.3%, and monkey L at 98.4% and 98.0%, in the mixed magnitude and the pure numerosity block, respectively (two-way ANOVA;  $F_1 = 2.53$ ,  $P = 0.11$ ). In addition, performance for all three numerosity samples (1, 2 and 4 dots) was equal ( $F_2 = 0.82$ ,  $P = 0.44$ ).

In summary, the behavioral data show that the monkeys were highly proficient in this DMS task with all three kinds of stimuli. The performance for numerosity trials was comparable in both block types. This indicates that the addition of line and color trials in the mixed magnitude block did not change task demands for the numerosity trials, but only changed the contextual framework of numerosity discriminations.

### Neural activity of single cells

We recorded 394 single units from the PFC (monkey H: 220; monkey L: 174). The neural activity of single cells was analysed in two time windows: the sample phase, starting 100 ms after the onset of the sample stimulus; and the delay phase, starting 200 ms after the offset of the sample stimulus. Both time windows were 800 ms in duration. The discharge rates in these analysis windows were averaged over trials, and analysed separately for sample and delay.

In the mixed magnitude block, two-way ANOVAs were used separately for every magnitude to determine whether a cell was selective for this magnitude and/or for the stimulus protocol ( $P < 0.01$ ). All cells that showed a significant main effect of the stimulus protocol or an interaction with the main effect 'protocol' were excluded from further analysis, because such cells were not regarded as abstract magnitude detectors. Table 1 shows the distribution of pure magnitude-selective cells for the sample and delay phases. Overall, 14.1% (56/394) of the cells were magnitude-selective during the sample phase, and 18.4% (77/394) during the delay phase. The majority of these cells were selective only for one magnitude, only a small proportion of cells showed selectivity for two, or even all three magnitudes.

To assess the effect of magnitude context on the representation of numerosities, discharge rates of individual cells in the numerosity conditions were analysed across pure numerosity and mixed magnitude blocks, with main factors numerosity, stimulus protocol and block condition (three-way ANOVA,  $P < 0.01$ ). Cells that showed only a significant main effect of numerosity were called 'pure numerosity' cells. All numerosity-selective neurons, including the ones that, in addition to the main effect of numerosity, had a main effect of block or an interaction with it, were called 'numerosity' cells.

TABLE 1. Proportion of magnitude-selective PFC neurons in the two trial phases, sample and delay (mixed magnitude block).

Main effect of	Sample	Delay
Numerosity	5.8% (23)	7.9% (31)
Line length	3.8% (15)	4.6% (18)
Color	3% (12)	2.3% (9)
Two or more magnitudes	1.5% (6)	3.6% (14)

Selective cells were identified by a two-factor ANOVA, with the factors sample magnitude and stimulus protocol, separately for each magnitude. Total number of cells:  $n = 394$ .

Cells that showed a significant main effect of stimulus protocol or an interaction with it were excluded from further analysis. Table 2 shows the distribution of numerosity-selective cells during the sample and the delay phase. Out of the 10.7% (42/394) of task-selective cells found in the sample period, 8.9% (35/394) were purely numerosity-selective, without any other main effects or interactions. During the delay phase, the number of numerosity-selective neurons increased to 17% (67/394), with 13.2% (52/394) of cells showing pure numerosity-selectivity.

Whether the monkeys were engaged in a pure numerosity block or a mixed magnitude block seemed to have little effect on numerosity representation. During the sample phase, a main effect of block condition (in addition to the main effect of numerosity) was present only in 0.5% (2/394, 5.7% relative to all numerosity-selective cells) of the cells (neurons showing an interaction between numerosity and block condition were absent). During the delay period, a main effect of block condition could be found in only 1.8% (7/394, 13.5% relative to all numerosity-selective cells) of the neurons; in 0.5% (2/394, 3.9% relative to all numerosity-selective cells) of the cells a significant interaction between the numerosity and the block condition was present. Thus, the context of the numerosity discrimination hardly modulated the response properties of numerosity detectors.

Figure 3 shows the responses of an example cell that was purely numerosity-selective (no other main effects or interactions) during the sample phase. The PSTHs show the averaged and smoothed discharge rates (Gaussian kernel, 100-ms sliding window) plotted over time. The different colors represent different sample numerosities. Numerosity-selective cells showed discharge rates that were significantly different to one numerosity compared with the others. This numerosity that elicited the highest discharge rate was called the 'preferred' numerosity of the cell. On average, the discharge rate decreased with increasing distance to the preferred quantity, thus resulting in a tuning curve for each individual cell. The example cell in Fig. 3 showed a clear preference for the sample numerosity one. This preference was the same in the mixed magnitude block and in the pure numerosity block. The inserts show the cell's tuning curves, which were virtually identical in the pure numerosity and the mixed magnitude block. These results indicate that there was no influence of the block condition on the coding properties of single neurons in the PFC.

### Population responses

To assess whether the magnitude context caused changes at the neuronal population level, we analysed the temporal and the tuning characteristics of numerosity-selective neurons. To determine whether the time course of numerosity representation was altered as a function of the magnitude context, we calculated the visual and

TABLE 2. Proportion of numerosity-selective PFC neurons in the two trial phases, sample and delay

Numerosity-selectivity	Sample	Delay
Pure	8.9% (35)	13.2% (52)
With main effect of block	0.5% (2)	1.8% (7)
With interaction with block	–	0.5% (2)
With main effect and/or interaction with protocol	1.3% (5)	1.5% (6)

Selective cells were identified by a three-factor ANOVA, with the factors sample numerosity, presentation block and stimulus protocol. Total number of cells:  $n = 394$ .

the numerosity-selectivity latencies for all cells that were numerosity-selective in the sample phase. The median latency of visual response was 200 ms in the pure numerosity block and 185 ms in the mixed magnitude block. This difference between the blocks was not significant (Mann–Whitney  $U$ -test,  $Z = 1.07$ ,  $P = 0.29$ ). This indicates a similar onset in visual response in the two block types.

The latency of numerosity-selectivity was measured by a sliding Kruskal–Wallis test for each numerosity-selective neuron in the sample phase. The median selectivity latencies were 299.5 ms for the pure numerosity and 260 ms for the mixed magnitude block conditions. There was no significant difference in the onset of selectivity between the two block conditions (Mann–Whitney  $U$ -test,  $Z = 0.76$ ,  $P = 0.491$ ). Hence, the context of numerosity discrimination did not affect the time course of the numerical representations in the PFC.

Next, we analysed the tuning properties of numerosity-selective cells in the two block conditions. Figure 4 shows average PSTHs of the numerosity-selective neurons. Figure 4A and B shows the population of sample-selective neurons in the mixed magnitude and the pure numerosity block, respectively. The population responses were very similar in the two blocks, with a clear differentiation between the responses to the preferred numerosity from the responses to the second preferred numerosity about 200 ms after sample onset. Delay-selective numerosity cells (Fig. 4D and E) showed the same discharge properties in the mixed magnitude and the pure numerosity block, differentiating between the numerosities from the beginning of the delay phase.

To assess the sharpness of tuning (e.g. the width of the tuning curve) and thus how well the neurons discriminated between the numerosities in the two block conditions, the discharge rates of numerosity neurons were normalised and plotted against the numerical distance to the preferred numerosity of the cell. The highest discharge rate of each cell, in the pure numerosity block, was defined as 100%, the lowest as 0%. All other discharge rates were normalised relative to these values. These normalised discharge rates were averaged over all cells for the two different blocks (Fig. 4C and F). For both blocking conditions, the discharge rates dropped monotonously with increasing numerical distance to the preferred quantity. The normalised discharge rates in the two blocks were compared separately for each distance to the preferred numerosity. There were no significant differences between the two blocks, neither in the sample nor in the delay phase, indicating the same sharpness of tuning irrespective of the magnitude context (Wilcoxon test with Bonferroni correction, for all comparisons  $P > 0.01$ ).

To compare the strengths of numerosity-selectivity, a SI was calculated for individual neurons in the mixed magnitude block and the pure numerosity block (Fig. 5A). No difference in SI values was detected in the sample phase (mean pure numerosity block SI = 0.39; mean mixed magnitude block SI = 0.35; Wilcoxon test,  $Z = 1.24$ ,  $P = 0.21$ ;  $n = 37$ ). Similarly, SI values were equal during the delay period (mean pure numerosity block SI = 0.42; mean mixed magnitude block SI = 0.42; Wilcoxon test,  $Z = 0.23$ ,  $P = 0.82$ ;  $n = 61$ ). Even if SIs are equal on average, it might be possible that two separate neuron populations react differently in the two blocks, leading on average to indiscriminable differences between blocks. To address this question, the SI in the pure numerosity block was plotted as a function of selectivity in the mixed magnitude block (Fig. 5B and C). The distance of the resulting points to the bisection line was calculated. The distribution of distances is depicted in the inserts. A skewed or a bimodal distribution would suggest two different populations of neurons. Hence, we tested the distribution of distances for symmetry around zero and bimodality. The distribution was not significantly asymmetrical,

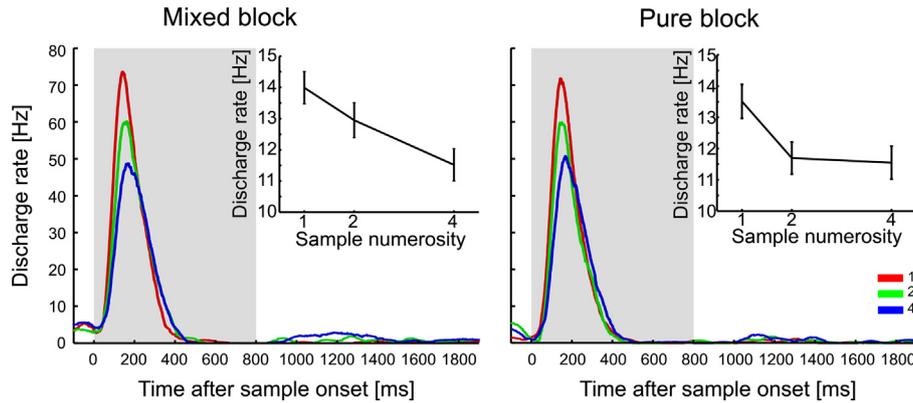


FIG. 3. Single-cell responses. Example neuron exhibiting numerosity-selectivity in both the mixed magnitude (left panel) and the pure numerosity block (right panel). Samples are color-coded; each curve is an average over 42 trials (smoothed with Gaussian kernel, 100 ms wide). Shaded area: presentation of the sample. Inset: neuronal filter function (tuning curve) over the sample phase.

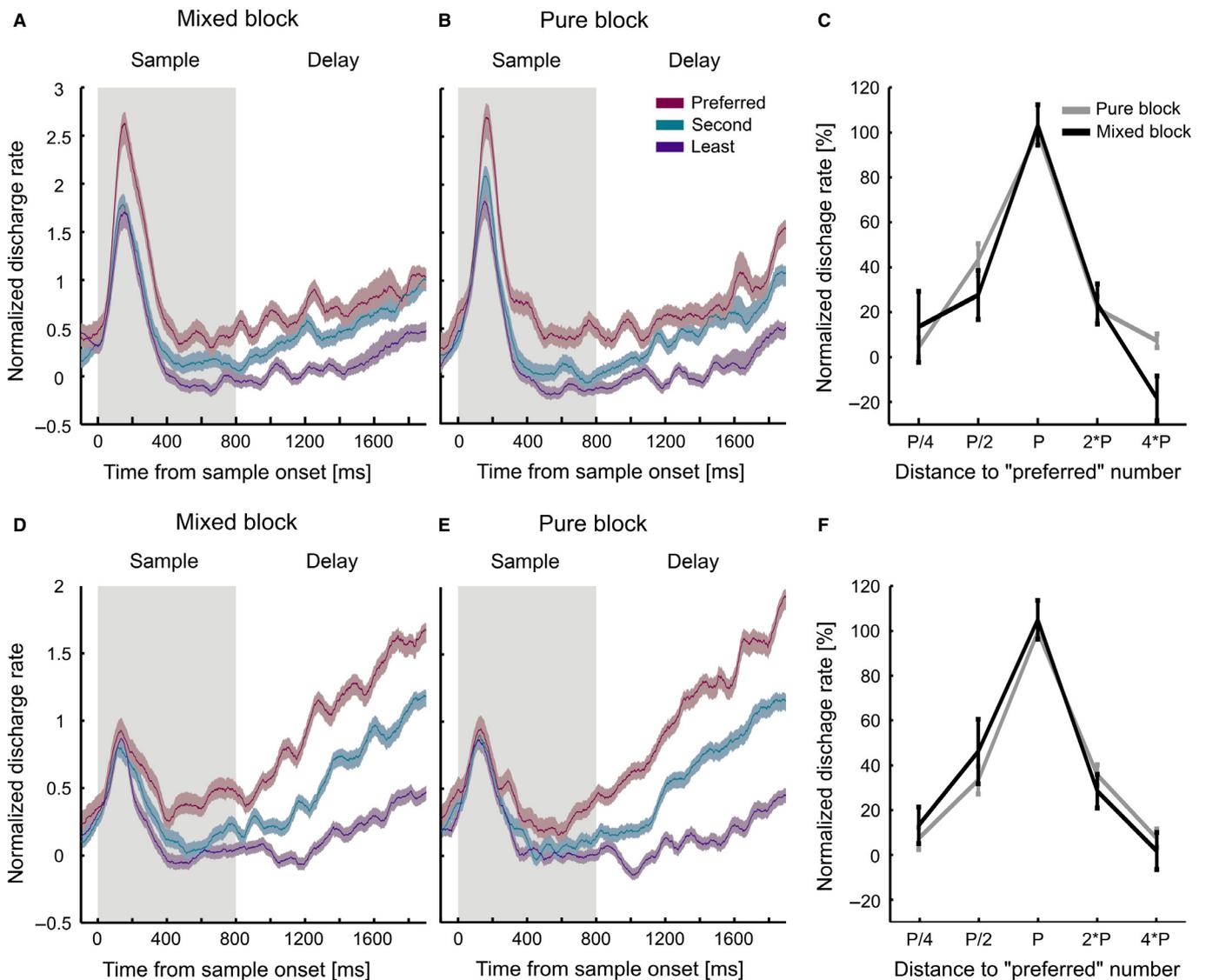


FIG. 4. Average PSTH for preferred, second and least preferred numerosities for sample-selective (A) and (B), and delay-selective cells (D) and (E). (C and F) Normalised response rates of numerosity-selective neurons as a function of distance to the preferred numerosity in the pure numerosity and the mixed magnitude block, for sample- and delay-selective neurons, respectively.

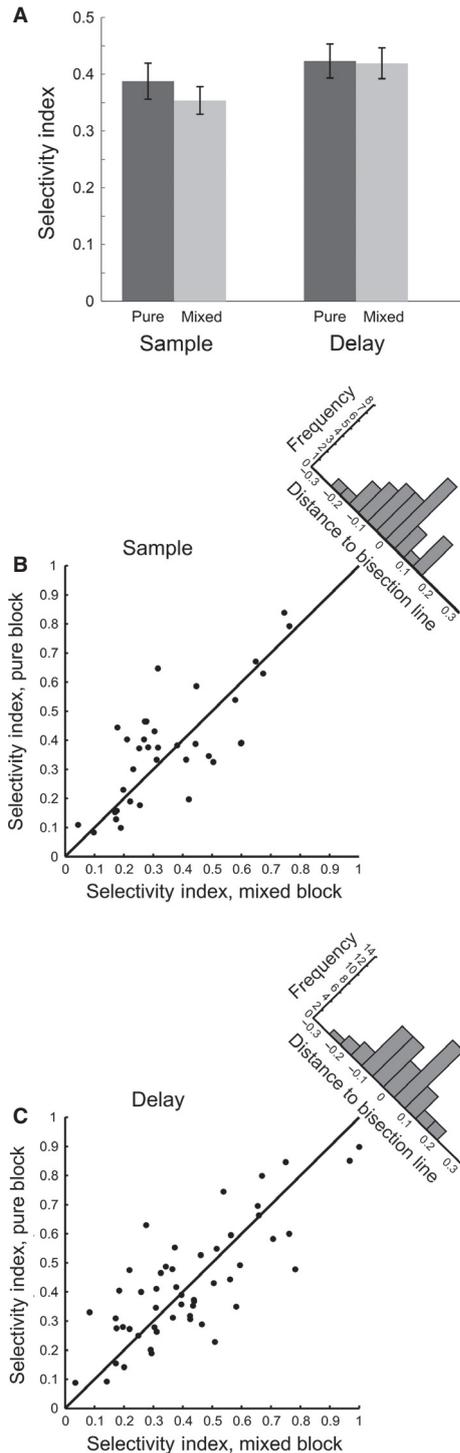


FIG. 5. (A) Average SIs of all numerosity-selective neurons in the pure numerosity and the mixed magnitude blocks. Scatter plot of SIs in the pure numerosity block as a function of selectivity in the mixed magnitude block (B) sample- and (C) delay-selective cells. Each dot represents one cell. Inserts: distances of the dots to the bisection line, positive values were assigned to dots below the line (higher selectivity in the mixed than in the pure numerosity block) and negative to the dots above the line (higher selectivity in the pure numerosity than in the mixed magnitude block).

neither in the sample (signed-rank test,  $Z = 1.24$ ,  $P = 0.21$ ) nor in the delay phase (signed-rank test,  $Z = 0.23$ ,  $P = 0.89$ ). A potential bimodal distribution was tested with the Hartigan's dip test

(Hartigan & Hartigan, 1985) and was also not significant (sample:  $P = 0.78$ ; delay:  $P = 0.82$ ).

## Discussion

In this study, we report that the context of numerosity discriminations did not influence the response properties of numerosity detectors. Whether monkeys discriminated visual numerosity in an isolated delayed match-to-numerosity task or embedded in equivalent DMS tasks with other magnitudes as discriminative stimuli had little, if any, effect on the tuning curves of selective neurons (i.e. the preferred numerosity and the sharpness of tuning), the time course of selectivity or the average population responses.

These results were not expected based on the 'adaptive coding hypothesis', which posits that neurons in the PFC are not inherently tuned to specific stimulus features, but rather adapt their response properties to code stimuli according to task relevance. Within this framework, context-specific parameters would shape the tuning functions of PFC neurons (Duncan, 2001). Support for this hypothesis comes from studies with complex behavioral protocols (Stokes *et al.*, 2013), showing dynamic response properties of PFC neurons that often are not specialised for a single function but highly adaptive. Selectivity for arbitrary visual categories often emerges after explicit training to distinguish those categories. For example, Freedman *et al.* (2002) showed that monkeys trained to discriminate computer-generated stimuli into 'cats' and 'dogs' categories had PFC neurons selective for both categories. Subsequently, the monkeys were retrained to assign the same stimuli into three new categories (with the two new category boundaries orthogonal to the original two-category boundary). After this learning process, tuning to the previously learned, now-irrelevant, cat and dog categories was lost. Instead, information about the three-category scheme was evident in the population of PFC neurons. Accordingly, it might be expected that PFC neurons change, at least to some extent, their tuning to numerosity and split or adapt their coding capacities according to the different magnitude contexts at hand. After all, encoding and memorizing three magnitudes (numerosity, length and color) in one task requires three times more coding capacities than representing only one quantity (numerosity). One way to deal with this increased coding demand might have been for a single neuron to represent more than one magnitude simultaneously and become multitasking (Cromer *et al.*, 2010). Alternatively, cells might have switched from coding line lengths or colors to coding numerosities in the pure numerosity block, leading to an increased number of selective cells. Instead, the vast majority of PFC neurons showed stable selectivity (Roy *et al.*, 2010) for a specific, preferred numerosity category, irrespective of the magnitude context. Alternatively, neurons could have decreased their strength of numerosity coding as a function of the increased stimulus space in the three-magnitude block condition. This was observed by Meyer *et al.* (2011) who examined the spatial and shape selectivity of neurons in the PFC after training in various working memory tasks. Neurons were sampled on a spatial working memory task, a feature working memory task, and a spatial-feature conjunction working memory task. Relative to the selectivity found in the feature working memory task alone, the average neuronal selectivity decreased in the conjunction task (requiring both feature and spatial working memory). This observed shift in neuronal selectivity was not due to an increase of difficulty in the conjunction task, because the monkeys performed equally well in both the feature working memory task and a spatial-feature conjunction working memory task. In our study, however, SI values of individual numerosity-tuned neurons remained unchanged between the pure

numerosity block and the mixed magnitude block. Thus, in contrast to simple spatial or feature discrimination task, the strength of numerosity representations remained stable irrespective of task demands or context.

While learning effect requires plasticity of neurons and their response properties, ubiquitously changing selectivities of PFC neurons may not be the best computational strategy for all types of abstract information. The current data suggest that numerosity representations in the PFC rely on a sparse code (Olshausen & Field, 2004) with dedicated and stable 'labeled lines' (Nieder & Merten, 2007). Perhaps numerosity, like faces (Gross *et al.*, 1972; Moeller *et al.*, 2008), constitutes an exceptionally relevant type of information essential for survival, which is best represented by strictly specialised neurons.

Our recent finding of numerosity-selective neurons in numerically naïve monkeys (Viswanathan & Nieder, 2013) supports the idea of a visual 'number sense', the faculty to perceive visual collections intuitively (Danzig, 1930; Dehaene, 1997). Neurons in areas 46/45 of the lateral PFC reliably encode the number of visual items in numerically naïve monkeys, i.e. monkeys that have never been trained to discriminate numerosities and demonstrably ignored them during the color discrimination task. Thus, visual numerosity-selective neurons may develop spontaneously and naturally within visual neural structures of the primate brain, prior to learning how to use this information (in an arbitrary task). Based on psychophysical findings, Burr & Ross (2008) suggested visual numerosity as a sensory attribute because perceived numerosity is susceptible to adaptation just like color, contrast or speed. However, it is difficult to imagine numerosity to be represented at the level of the early visual cortex. Moreover, adaptation is not restricted to primary visual attributes, but also is observed for high-level visual categories, such as faces (Webster & MacLeod, 2011). Adaptation processes indicate a specialised neural pathway with a limited number of units, which get recruited by the adaptive stimulus and are biased by it when the stimulus changes. Such adaptation of complex visual categories is also found in other domains, particularly within the ventral visual stream leading to the anterior inferotemporal (IT) cortex (Kovács *et al.*, 2006). In the IT cortex, neurons are specifically responsive to faces, places and body parts, and seem to be located in dedicated neural substrates (Tsao *et al.*, 2008; Pinsk *et al.*, 2009; Bell *et al.*, 2011). The category 'set size' could therefore emerge as a special perceptual category represented spontaneously in a dedicated parieto-frontal network. Neuronal selectivity for numerosity, body parts or faces may thus constitute 'natural' categories, and be present at birth in both humans (Fried *et al.*, 1997; Kreiman *et al.*, 2000) and monkeys (Rodman *et al.*, 1991, 1993).

To further test the notion of stable numerosity coding in the PFC, it will be necessary to also investigate the tuning properties of numerosity detectors in more radically changing contexts, such as genuine task-switching protocols. For instance, it would be interesting to see whether or not switching from a delayed match-to-numerosity task (Nieder *et al.*, 2002; Nieder & Merten, 2007) to a rule-switching task based on numerosities (Vallentin *et al.*, 2012; Eiselt & Nieder, 2013) would modify the coding properties. In addition, it remains to be investigated whether long-term learning is suited to modify the proportion or tuning functions of numerosity-selective neurons.

## Abbreviations

DMS, delayed match-to-sample; IT, inferotemporal; PFC, prefrontal cortex; PSTH, peri-stimulus time histogram; SD, standard deviation; SEM, standard error of the mean; SI, selectivity index; VA, visual angle.

## References

- Bell, A.H., Malecek, N.J., Morin, E.L., Hadj-Bouziane, F., Tootell, R.B. & Ungerleider, L.G. (2011) Relationship between functional magnetic resonance imaging-identified regions and neuronal category selectivity. *J. Neurosci.*, **31**, 12229–12240.
- Bongard, S. & Nieder, A. (2010) Basic mathematical rules are encoded by primate prefrontal cortex neurons. *Proc. Natl. Acad. Sci. USA*, **107**, 2277–2282.
- Burr, D. & Ross, J. (2008) A visual sense of number. *Curr. Biol.*, **18**, 425–428.
- Cromer, J.A., Roy, J.E. & Miller, E.K. (2010) Representation of multiple, independent categories in the primate prefrontal cortex. *Neuron*, **66**, 796–807.
- Danzig, T. (1930) *Number - The Language of Science*. The Macmillan company, New York.
- Dehaene, S. (1997) *The Number Sense*. Oxford University Press, Oxford.
- Duncan, J. (2001) An adaptive coding model of neural function in prefrontal cortex. *Nat. Rev. Neurosci.*, **2**, 820–829.
- Eiselt, A.K. & Nieder, A. (2013) Representation of abstract quantitative rules applied to spatial and numerical magnitudes in primate prefrontal cortex. *J. Neurosci.*, **33**, 7526–7534.
- Freedman, D.J., Riesenhuber, M., Poggio, T. & Miller, E.K. (2002) Visual categorization and the primate prefrontal cortex: neurophysiology and behavior. *J. Neurophysiol.*, **88**, 929–941.
- Fried, I., MacDonald, K.A. & Wilson, C.L. (1997) Single neuron activity in human hippocampus and amygdala during recognition of faces and objects. *Neuron*, **18**, 753–765.
- Gross, C.G., Rocha-Miranda, C.E. & Bender, D.B. (1972) Visual properties of neurons in inferotemporal cortex of the Macaque. *J. Neurophysiol.*, **35**, 96–111.
- Hartigan, J.A. & Hartigan, P.M. (1985) The dip test of unimodality. *Ann. Stat.*, **13**, 70–84.
- Kovács, G., Zimmer, M., Bankó, E., Harza, I., Antal, A. & Vidnyánszky, Z. (2006) Electrophysiological correlates of visual adaptation to faces and body Parts in humans. *Cereb. Cortex*, **16**, 742–753.
- Kreiman, G., Koch, C. & Fried, I. (2000) Category-specific visual responses of single neurons in the human medial temporal lobe. *Nat. Neurosci.*, **3**, 946–953.
- Merten, K. & Nieder, A. (2012) Active encoding of decisions about stimulus absence in primate prefrontal cortex neurons. *Proc. Natl. Acad. Sci. USA*, **109**, 6289–6294.
- Merten, K. & Nieder, A. (2013) Comparison of abstract decision encoding in the monkey prefrontal cortex, the presupplementary, and cingulate motor areas. *J. Neurophysiol.*, **110**, 19–32.
- Meyer, T., Qi, X.L., Stanford, T.R. & Constantinidis, C. (2011) Stimulus selectivity in dorsal and ventral prefrontal cortex after training in working memory tasks. *J. Neurosci.*, **31**, 6266–6276.
- Miller, E.K. & Cohen, J.D. (2001) An integrative theory of prefrontal cortex function. *Annu. Rev. Neurosci.*, **24**, 167–202.
- Moeller, S., Freiwald, W.A. & Tsao, D.Y. (2008) Patches with links: a unified system for processing faces in the macaque temporal lobe. *Science*, **320**, 1355–1359.
- Nieder, A. (2012) Supramodal numerosity selectivity of neurons in primate prefrontal and posterior parietal cortices. *Proc. Natl. Acad. Sci. USA*, **109**, 11860–11865.
- Nieder, A. (2013) Coding of abstract quantity by 'number neurons' of the primate brain. *J. Comp. Physiol. A*, **199**, 1–16.
- Nieder, A. & Merten, K. (2007) A labeled-line code for small and large numerosities in the monkey prefrontal cortex. *J. Neurosci.*, **27**, 5986–5993.
- Nieder, A., Freedman, D.J. & Miller, E.K. (2002) Representation of the quantity of visual items in the primate prefrontal cortex. *Science*, **297**, 1708–1711.
- Olshausen, B.A. & Field, D.J. (2004) Sparse coding of sensory inputs. *Curr. Opin. Neurobiol.*, **14**, 481–487.
- Pinsk, M.A., Arcaro, M., Weiner, K.S., Kalkus, J.F., Inati, S.J., Gross, C.G. & Kastner, S. (2009) Neural representations of faces and body parts in macaque and human cortex: a comparative fMRI study. *J. Neurophysiol.*, **101**, 2581–2600.
- Rodman, H.R., Skelly, J.P. & Gross, C.G. (1991) Stimulus selectivity and state dependence of activity in inferior temporal cortex of infant monkeys. *Proc. Natl. Acad. Sci. USA*, **88**, 7572–7575.
- Rodman, H.R., Scialidhe, S.P. & Gross, C.G. (1993) Response properties of neurons in temporal cortical visual areas of infant monkeys. *J. Neurophysiol.*, **70**, 1115–1136.

- Roy, J.E., Riesenhuber, M., Poggio, T. & Miller, E.K. (2010) Prefrontal cortex activity during flexible categorization. *J. Neurosci.*, **30**, 8519–8528.
- Stokes, M.G., Kusunoki, M., Sigala, N., Nili, H., Gaffan, D. & Duncan, J. (2013) Dynamic coding for cognitive control in prefrontal cortex. *Neuron*, **78**, 364–375.
- Tsao, D.Y., Schweers, N., Moeller, S. & Freiwald, W.A. (2008) Patches of face-selective cortex in the macaque frontal lobe. *Nat. Neurosci.*, **11**, 877–879.
- Vallentin, D., Bongard, S. & Nieder, A. (2012) Numerical rule coding in the prefrontal, premotor, and posterior parietal cortices of macaques. *J. Neurosci.*, **32**, 6621–6630.
- Viswanathan, P. & Nieder, A. (2013) Neuronal correlates of a visual “sense of number” in primate parietal and prefrontal cortices. *Proc. Natl. Acad. Sci. USA*, **110**, 11187–11192.
- Wallis, J.D., Anderson, K.C. & Miller, E.K. (2001) Single neurons in prefrontal cortex encode abstract rules. *Nature*, **411**, 953–956.
- Webster, M.A. & MacLeod, D.I. (2011) Visual adaptation and face perception. *Philos. T. Roy. Soc. B*, **366**, 1702–1725.