Article

Current Biology

# Neuronal Representation of Numerosity Zero in the Primate Parieto-Frontal Number Network

## Highlights

- VIP neurons encode empty sets more as a distinct, non-numerical category
- In contrast, PFC cells integrate empty sets in the numerical continuum
- Prefrontal encoding of empty sets is more abstract and behaviorally relevant
- Empty sets are translated into an abstract representation of a zero-numerosity category

## Authors

Araceli Ramirez-Cardenas, Maria Moskaleva, Andreas Nieder

## Correspondence

andreas.nieder@uni-tuebingen.de

## In Brief

In behaving monkeys, Ramirez-Cardenas et al. report neuronal empty-set representations that become steadily more abstract and gradually positioned into a numerical continuum along the parieto-frontal network. These findings elucidate how the brain transforms the absence of countable items, nothing, into an abstract numerosity zero category.

Ramirez-Cardenas et al., 2016, Current Biology *26*, 1285–1294
May 23, 2016 © 2016 Elsevier Ltd.
http://dx.doi.org/10.1016/j.cub.2016.03.052

CellPress

Current Biology
Article

# Neuronal Representation of Numerosity Zero in the Primate Parieto-Frontal Number Network

Araceli Ramirez-Cardenas,[1] Maria Moskaleva,[1] and Andreas Nieder[1,*]
[1]Animal Physiology Unit, Institute of Neurobiology, Department of Biology, University of Tübingen, Auf der Morgenstelle 28, 72076 Tübingen, Germany
*Correspondence: andreas.nieder@uni-tuebingen.de
http://dx.doi.org/10.1016/j.cub.2016.03.052

## SUMMARY

**Neurons in the primate parieto-frontal network represent the number of visual items in a collection, but it is unknown whether this system encodes empty sets as conveying null quantity. We recorded from the ventral intraparietal area (VIP) and the prefrontal cortex (PFC) of monkeys performing a matching task including empty sets and countable numerosities as stimuli. VIP neurons encoded empty sets predominantly as a distinct category from numerosities. In contrast, PFC neurons represented empty sets more similarly to numerosity one than to larger numerosities, exhibiting numerical distance and size effects. Moreover, prefrontal neurons represented empty sets abstractly and irrespective of stimulus variations. Compared to VIP, the activity of numerosity neurons in PFC correlated better with the behavioral outcome of empty-set trials. Our results suggest a hierarchy in the processing from VIP to PFC, along which empty sets are steadily detached from visual properties and gradually positioned in a numerical continuum. These findings elucidate how the brain transforms the absence of countable items, nothing, into an abstract quantitative category, zero.**

## INTRODUCTION

The number of elements in a set, its numerosity, is an abstract property that can be quickly and directly assessed [1–4]. Human mathematical abilities are thought to be, at least in part, grounded in this sense of number [5, 6]. Several lines of evidence suggest that this capacity is not unique to humans, but deeply rooted in our ancestry [7, 8]. The neural system necessary to process quantity information resides in a dedicated parieto-frontal brain network [9–13]. Studies in nonhuman primates have identified neurons selectively tuned to the number of visual items contained in a set in the ventral intraparietal area (VIP) of the intraparietal sulcus of the posterior parietal cortex and the lateral prefrontal cortex (PFC) [14–17].

Over the past 15 years, the neuronal representation of numerosities in the parieto-frontal brain network has been intensively investigated. Still, whether numerosity zero finds a place in the neuronal number line has been barely explored, probably because zero is a late achievement, both in history and individual cognitive development [18, 19]. However, recent behavioral studies have shown that young children can position empty sets in the context of other small numerosities before they comprehend the symbolic zero [20]. Even nonhuman primates can recognize and assess empty sets as numerically significant [21, 22]. In tasks involving the discrimination and ordering of visually displayed sets of dots, rhesus monkeys spontaneously treated empty sets as the void quantity they represent [23]. Such behavioral studies suggest that empty sets may be encoded in the primate parieto-frontal magnitude system as part of the numerical continuum.

Previously, PFC neurons were shown to encode the perceived presence or absence of a visual stimulus [24, 25]. In addition, results from one monkey suggested that VIP neurons signal the lack of countable items in empty sets [26]. The present study was specifically designed to explore whether single neurons and neuronal populations in the parieto-frontal network represent empty sets as conveying a null quantitative value. If that is the case, empty-set representations should reflect the cardinal relationships between numbers and therefore exhibit a numerical distance effect (an improvement in discriminability of two quantities as the numerical difference between them increases) or a numerical size effect (a worsening in the discriminability of two numbers as their magnitude increases) [10, 27, 28]. In addition, numerical representations are expected to be abstract (i.e., invariant to appearance and low-level stimulus properties) and, when numerosity is behaviorally relevant, correlate with subjects' performance.

In the current study, we simultaneously recorded single-neuron activity in VIP and PFC of two monkeys performing a numerosity discrimination task. We describe the tuning of parietal and prefrontal selective neurons to empty sets and test their behavioral relevance. In addition to single-cell coding, we investigate neuronal dynamics at the population level and clarify the respective contributions of VIP and PFC. Our results show that prefrontal representations of empty sets better meet the criteria of a primitive correlate of numerosity zero.

## RESULTS

### Behavior

Two monkeys performed a delayed match-to-numerosity task to discriminate visual numerosities from 0 (empty sets) to 4 (Figure 1A) in standard and control protocols controlling for low-level visual features (Figure 1B) (see the Experimental Procedures for

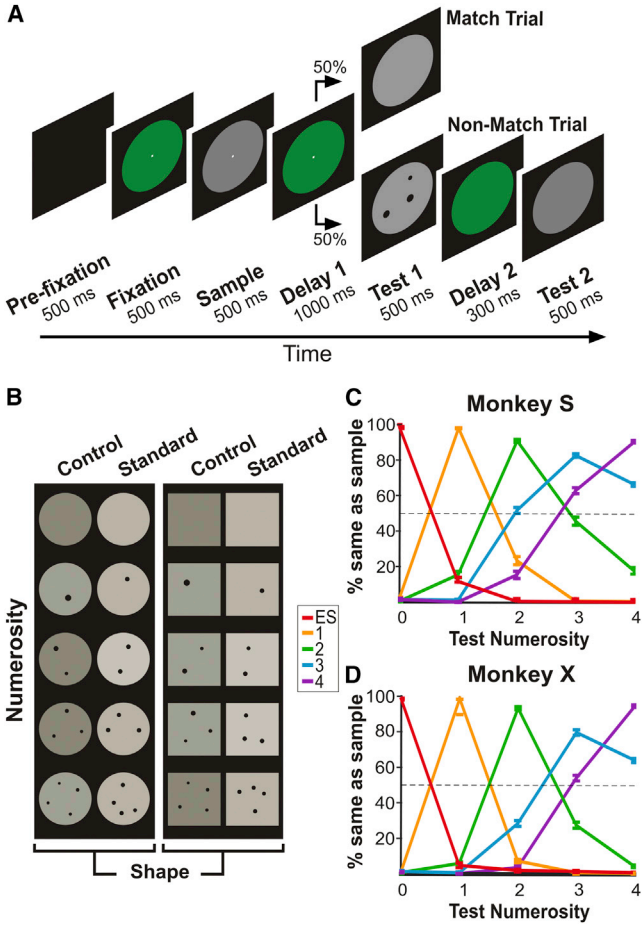


**Figure 1. Delayed Match-to-Numerosity Task and Example Stimuli**

(A) Fixating monkeys were presented with a sample numerosity ranging from 0 to 4 for 500 ms. The monkeys had to memorize the numerosity for a 1 s delay period and match it to a subsequent test stimulus (either the first or the second test stimulus was correct) by releasing a lever. For each sample numerosity, all four possible non-match numerosities were shown.

(B) Example stimuli for the different conditions. Numerosities 0 to 4 (factor numerosity) were shown in standard and control protocols (factor protocol) on a circular or square background (factor shape).

(C and D) Behavioral tuning curves derived from the monkeys' performance (C, monkey S; D, monkey X) when different stimuli were presented as samples (empty sets [ES] and 1–4). The functions reflect the probability that a monkey judged displays in the test period as containing the same number of items as the sample numerosity (indicated in various colors). The peak data point of each colored curve indicates the correct performance in match trials for the different sample numerosities. Data points to the left and right of the peak reflect performance in non-match trials (i.e., when the first test numerosity was smaller or larger than the sample). Error bars indicate the SEM.

a detailed description of stimuli). A total of 49 behavioral sessions for monkey X and 54 sessions for monkey S were analyzed. Behavioral tuning curves (Figures 1C and 1D) show how often animals judged test stimuli as equal in quantity to each sample numerosity. Curve peaks depict the percentage of correct match trials, whereas non-peak values correspond to errors in non-match trials. As previously observed for countable numbers [9, 14, 27], performance declined as the difference between sample and non-match numerosities decreased (numerical distance effect). Even though the monkeys made few mistakes in judging empty sets, a distance effect was evidenced by errors made by monkeys when empty sets were presented as sample stimuli (Figures 1C and 1D, red curves). Both monkeys mistakenly matched empty sets to numerosity 1 more frequently than to numerosity 2 (monkey S: 12.5% ± 1.3% versus 1.3% ± 0.4%, $p < 0.001$; monkey X: 4.8% ± 1.3% versus 2.0% ± 0.7%, $p = 0.001$; Wilcoxon signed-rank test).

In addition, the behavioral tuning curves became wider (i.e., less selective) with increasing sample number (numerical size effect), an effect a non-linearly compressed number line could account for. We explored whether behavioral performance curves could be better described on a linear or nonlinearly compressed number line. For that purpose, we fitted the behavioral curves per session with a standard symmetric peak function (Gaussian function) when plotted on either a linear or a logarithmic ($\log_2(n + 1)$) numerical scale. The goodness-of-fit ($r^2$) values were taken as a measure of which scaling scheme describes better the data. For each monkey, the fittings were significantly better on a logarithmically compressed number line (both monkeys: $r^2 = 0.98$) than on a linear scale (versus $r^2 = 0.94$ in monkey S and $r^2 = 0.95$ in monkey X; $p < 0.001$ in both subjects; Wilcoxon signed-rank test). Thus, as predicted by the Weber-Fechner law, skewed behavioral curves became more symmetric when plotted on a logarithmic scale.

## Neuronal Numerosity Tuning in VIP and PFC

While the monkeys performed the task, we simultaneously recorded 861 neurons in VIP of the intraparietal sulcus (431 from monkey X and 430 cells from monkey S) and 476 single neurons in the dorsolateral PFC (279 neurons from monkey X and 197 cells from monkey S) (Figure 2A). Many neurons were strongly modulated by sample numerosity and discharged as a function of numerical distance between the stimuli. Figure 2 shows three example neurons from VIP (Figures 2B–2D) and PFC (Figures 2E–2G). As with countable numerosities, many neurons discharged maximally to empty sets (empty sets as preferred numerosity) (Figures 2B and 2E) or responded least to them (least-preferred numerosity) (Figures 2C, 2D, 2F, and 2G).

To verify neuronal activity differences during the sample period, we tested all neurons with a three-way ANOVA (with main factors of number, shape, and protocol; evaluated at $p < 0.01$; see the Experimental Procedures). A considerable proportion of VIP neurons (19%; 163/861) and PFC neurons (39%; 185/476) were selective for main factor number and encoded the numerosity presented to the monkeys. Table S1 shows the proportions of cells that were selective for the different main factors and interactions between main factors. Neurons that only showed a significant main effect for factor number and no significance for any other main factor or factor interactions were identified as exclusive number-selective neurons. In the whole population, 8% (70/861) of VIP neurons and 16% (78/476) of PFC neurons belonged to this most conservatively determined subpopulation. Example neurons depicted in Figures 2B–2G are exclusively number-selective neurons.

Neurons tuned to each of the sample numerosities were found among the exclusive selective population, with empty-set-preferring neurons as the most-abundant class, both in VIP (Figure 3B) and PFC (Figure 3D). Figure S1 shows the individual

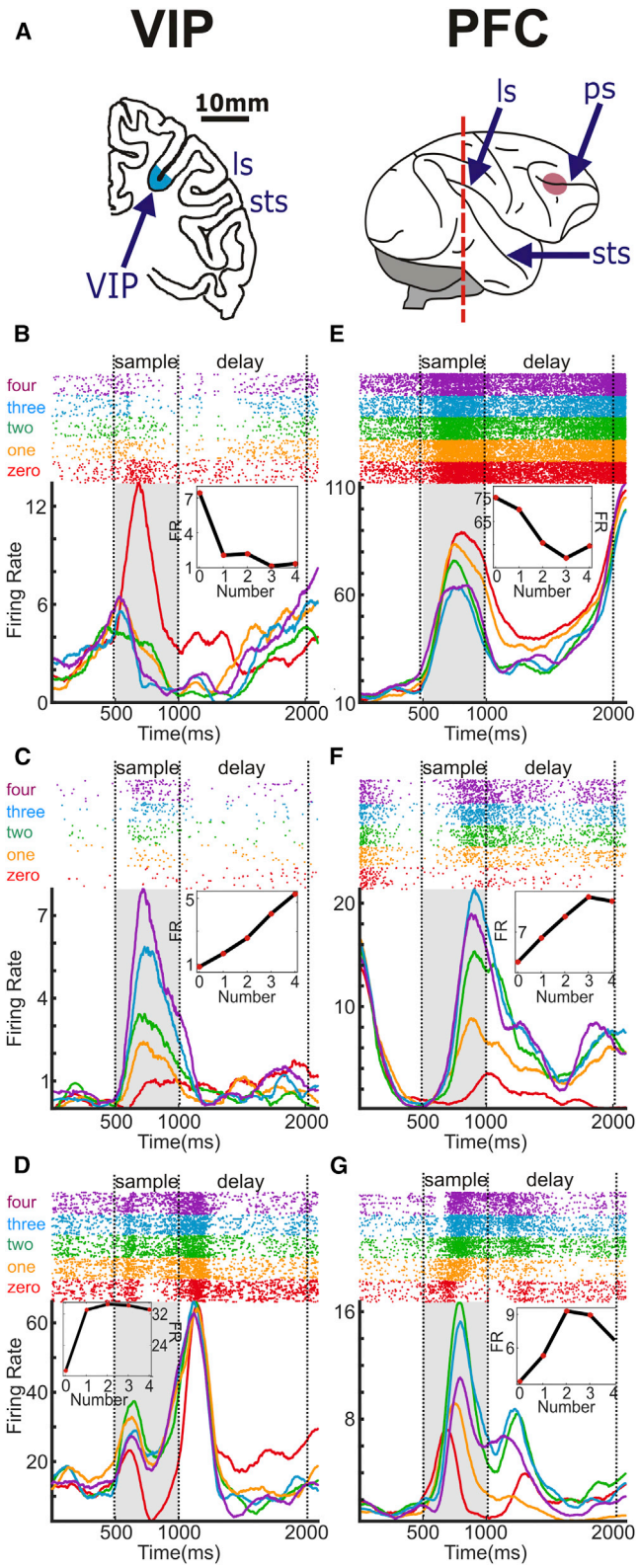


**Figure 2. Recording Sites and Neuronal Responses to Numerosity**

(A) Lateral view (right) of the right hemisphere of a monkey brain indicating the topographical relationships of cortical landmarks and coronal section (left) at the level of the dotted line in the lateral view reconstructed from a structural MRI scan. The red region on the frontal lobe and blue region in the fundus of the IPS mark the recording areas in PFC and VIP, respectively. ips, intraparietal sulcus; ls, lateral sulcus; sts, superior temporal sulcus.

(B–D) Example numerosity-selective neurons in VIP. A VIP neuron showing maximum responses to empty sets (zero) in the sample phase is shown in (B). The example neurons in (C) and (D) exhibited maximum suppression to empty sets. The top panel shows dot-raster histograms (each dot represents an action potential); the bottom panel depicts averaged spike density functions (activity averaged in a sliding 150 ms window). The first 500 ms represent the fixation period, followed by the sample and delay periods. Colors of dot histogram and spike density functions correspond to the numerosity of the sample stimulus. The inset in the spike density plot shows the neuron's tuning function (i.e., discharge rates as a function of the number of presented items) during the gray-shaded sample period.

(E–G) Example numerosity-selective neurons in PFC, showing maximum excitation (E) or strongest suppression by empty sets (F and G).

See also Table S1.

tuning curves of empty-set-preferring neurons. We constructed population tuning functions of exclusive selective neurons in VIP (Figure 3A) and PFC (Figure 3C) by normalizing the firing rates of individual neurons to the different sample stimuli and then averaging according to their respective preferred numerosity.

Next, we investigated which scaling scheme (linear of logarithmic) accounted better for the neuronal data. In VIP, a logarithmic scaling did not result in better fittings of selective neurons' tuning functions ($r^2 = 0.79$ for both scaling schemes; $p = 0.83$, Wilcoxon signed-rank test; $n = 70$). In numerosity-selective PFC neurons, however, the goodness-of-fit values were significantly higher in a logarithmic scaling ($r^2 = 0.82$) than in a linear scaling ($r^2 = 0.80$) ($p < 0.05$, Wilcoxon signed-rank test; $n = 78$). This indicates that, in agreement with the behavioral data, numerical magnitudes in PFC are best represented on a non-linearly compressed scale.

## VIP Neurons Discriminate Empty Sets More Categorically than PFC Cells

The tuning curves of neurons preferring countable numerosities in VIP (Figure 3A) and PFC (Figure 3C) showed a clear distance effect, i.e., a progressive drop-off of activity with increasing numerical distance from the preferred numerosity. However, a distance effect seemed to be absent in the average tuning curve of parietal empty-set neurons (red curve in Figure 3A), whereas PFC empty-set neurons exhibited a graded decline in firing rates with increasing numerosity (Figure 3C).

To quantify the differences in the tuning to empty sets, we used several parameters. First, we compared the responses of empty-set neurons to numerosity 1 and 2. These numerosities elicited similar levels of activity in VIP empty-set neurons (mean 0.32 for numerosity 1 versus 0.23 for numerosity 2; $p = 0.12$; $n = 28$; Wilcoxon signed-rank test). In contrast, PFC empty-set neurons responded more to numerosity 1 (mean 0.46) than to numerosity 2 (0.23) ($p < 0.001$; $n = 24$), thus coding empty sets as part of a numerical continuum. Second, we linearly fitted the firing rates for numerosity 1 to 4 and derived the slopes as a measure of firing rate decline with numerical distance. The slopes of the fits in VIP (mean –0.03) were not significantly different from 0 (one-sample t test against 0, $p = 0.07$; $n = 28$), whereas the slopes of the fits in PFC (mean –0.08) differed from 0 ($p < 0.001$; $n = 24$). More

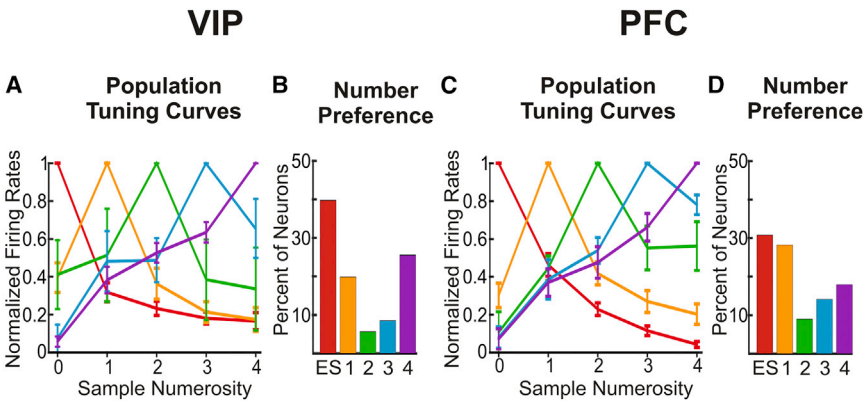


**Figure 3. Average Tuning Curves and Stimulus Preference**

(A and C) Population tuning curves obtained by averaging the normalized tuning curves of VIP (A) and PFC (C) neurons with the same preferred stimulus. The average tuning curve of empty-set-preferring neurons is shown in red. Error bars indicate the SEM.

(B and D) Proportion of neurons in VIP (B) and PFC (D) responding maximally to each of the stimulus types.

See also Figure S1.

importantly, the slopes were more negative in PFC compared to VIP (Mann-Whitney U test, $p = 0.019$). Third, we fitted the tuning curves of empty-set neurons (linear scale) with a Gaussian and derived the bell-curve width, sigma ($\sigma$). The sigma ($\sigma$) values for VIP empty-set neurons were smaller (mean 0.86) compared to PFC empty-set cells (0.94) (Mann-Whitney U test, $p = 0.028$). Finally, an exponential function was fitted to the empty-set tuning curves to measure the firing rate decay via the constant tau ($\tau$). VIP empty-set neurons showed smaller tau ($\tau$) values (mean 0.73) compared to PFC cells (mean 1.14) (Mann-Whitney U test, $p = 0.039$), indicating that the firing rate decays more steeply with number in VIP empty-set neurons. Collectively, these results confirm that parietal neurons implement a more binary (nothing versus something) representation of the different sample stimuli.

## Behavioral Relevance of Empty-Set and Numerosity-Selective Neurons

To investigate the behavioral relevance of exclusive numerosity-selective neurons in either VIP or PFC, we analyzed sample activity in error trials and compared it to responses in correct trials. First, we explored whether the firing rate of empty-set tuned neurons correlated with successful completion of the task. If the responses of these neurons to empty sets (their preferred stimulus, i.e., eliciting maximal responses) were relevant for trial outcomes, lower firing rates would be expected in failed empty-set trials. Indeed, the responses of VIP (Figure 4A) and PFC (Figure 4B) empty-set neurons were reduced in erroneous empty-set trials (VIP: $6.33 \pm 1.26$ Hz versus $5.03 \pm 0.10$ Hz, $n = 15$, $p < 0.05$; PFC: $13.04 \pm 3.94$ Hz versus $7.58 \pm 2.34$ Hz, $p < 0.05$; $n = 10$; Wilcoxon signed-rank test; note that only cells with a sufficient number of error trials were considered for analysis). This result suggests that the activity of empty-set neurons in both VIP and PFC is relevant for the outcome of trials in which an empty set was presented as sample.

Does the activity of empty-set neurons also correlate with the performance in countable numerosity trials? In correct trials, countable numerosities were encoded with low firing rates by empty-set neurons. If this low activity were relevant for performance, higher firing rates to the non-preferred numerosities of empty-set neurons might lead to errors. We compared the firing rates of empty-set neurons to their least-preferred stimulus (a countable numerosity) in correct and error trials. Empty-set neurons in VIP (Figure 4A) and PFC (Figure 4B) exhibited higher firing rates to their least-preferred numerosity in error compared to correct trials (VIP: $2.61 \pm 0.63$ Hz versus $3.01 \pm 0.63$ Hz, for correct and error trials, $p < 0.05$, $n = 24$, Wilcoxon signed-rank test; PFC: $6.54 \pm 2.5$ Hz versus $7.67 \pm 2.74$ Hz, for correct and error trials, $p < 0.05$, $n = 23$, Wilcoxon signed-rank test). Thus, the activity of empty-set neurons in both cortical areas was also correlated with the outcome of trials in which countable numerosities were presented as sample.

Finally, is the activity of countable-numerosity neurons, in turn, correlated with performance in empty-set trials? In this case, we analyzed the error-trial activity of countable-numerosity neurons which fired the least for empty sets in correct trials. The activity of countable-numerosity neurons in VIP (Figure 3C) did not differ in error and correct empty-set trials ($4.51 \pm 1.60$ versus $4.73 \pm 1.7$, $p = 0.86$; $n = 14$). In countable-numerosity PFC neurons (Figure 3D), however, the activity during the sample period increased in erroneous empty-set trials ($4.33 \pm 0.80$ versus $8.49 \pm 2.35$, $p = 0.04$; $n = 9$; Wilcoxon signed-rank test). So, only the activity of PFC countable-numerosity neurons during the sample period is correlated with the outcome of empty-set trials.

## Neuronal Population Dynamics in VIP and PFC

Next, we explored how VIP and PFC neuronal populations, irrespective of selectivity status or stimulus preference [29, 30], encoded empty sets. We analyzed the coding capacity and dynamics of population responses as a whole by performing a multidimensional state space analysis (Gaussian-process factor analysis, GPFA) [31] on similarly sized pseudo-populations of neurons in VIP and PFC (377 and 364 neurons, respectively). This approach extracts trajectories from the spiking activity of a neuronal population in individual trials. Such trajectories reflect the instantaneous firing rate of the respective neuronal population as they evolve over time. Figures 5A and 5B depict average population trajectories for the different sample stimuli in a space defined by the top-three most-meaningful dimensions.

To evaluate the population numerical tuning, we measured Euclidian distances between trial trajectories corresponding to different samples. In VIP, parietal population dynamics did not exhibit a distance effect for empty sets (Figure 5C). The inter-trajectory distances between empty sets and different countable numerosities (0–1, 0–2, 0–3, and 0–4) were comparable for all numerical distances (time-defined Kolmogorov-Smirnov permutation

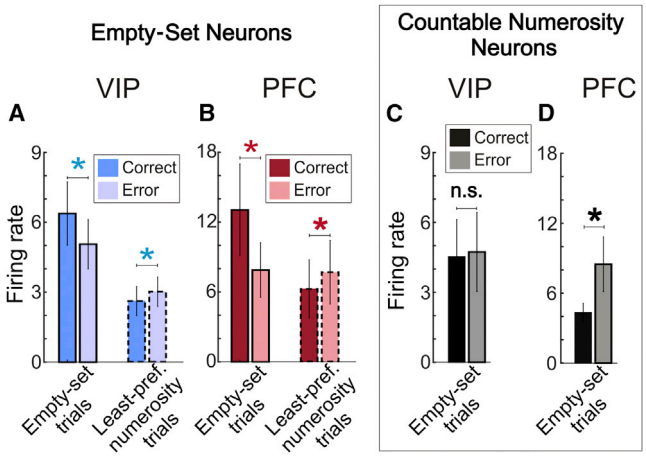


**Figure 4. Error Trial Analysis**

(A and B) Behavioral relevance of empty-set neurons in empty-set trials (preferred stimulus) (A) and countable numerosity trials (least-preferred stimulus) (B). The neuronal firing rates in the sample period are compared between correct and error trials.

(C) Firing rates of VIP and PFC countable numerosity neurons in correct and erroneous empty-set trials.

Error bars indicate the SEM.

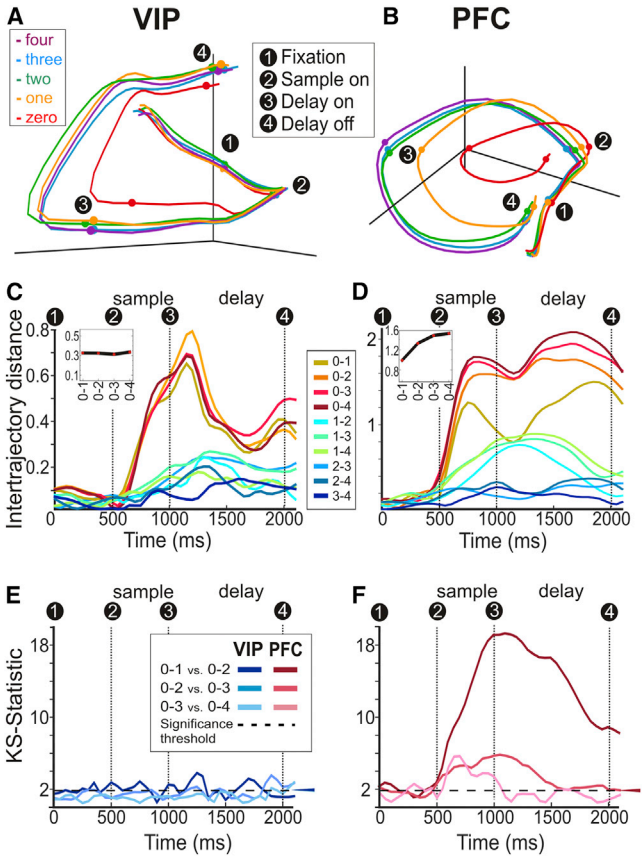


**Figure 5. Population Dynamics in VIP and PFC**

(A and B) Average whole-population state-space trajectories in VIP (A) and PFC (B) in trials with different sample stimuli.

(C and D) Average distances between trials with different sample numerosities in the whole population of VIP neurons (C) and PFC neurons (D).

(E and F) Statistical comparison of pairs of inter-trajectory distances that define a distance effect for empty sets in the whole population of VIP neurons (E) and PFC neurons (F). The distributions of trial inter-trajectory distances were compared with a Kolmogorov-Smirnov permutation test per time bin. The significance threshold for each comparison, evaluated at an alpha level of 0.05, is marked with a dotted line and a lateral colored arrow.

See also Figures S2 and S3.

test comparing the distributions of inter-trial distances for the different pairs of stimulus, 1,000 random permutations, alpha level 0.05; Figure 5E).

In contrast, in the ordered layout formed by prefrontal trajectories (Figure 5B), a population distance effect could be identified. The closer two numerosities were in the numerical continuum, the more similar were their patterns of population activity, and vice versa. This held true for empty sets. The distance between population trajectories in empty-set trials and other trials increased with the sample magnitude of the latter (Figure 5D, inset). Indeed, all inter-trajectory distance comparisons that defined a distance effect for empty sets surpassed their significant threshold in PFC (Figure 5F) during the sample period. As expected, both areas segregated the two classes when average trajectories were calculated for empty sets versus all countable numerosities (Figures S2A–S2D). Finally, the whole analysis was performed with the exclusive selective population of neurons (Figures S3A–S3F). Results exhibited the same patterns: a distance effect was not significant in VIP exclusive number-selective neurons but was clearly present in PFC exclusive number-selective neurons.

## Population Decoding

We trained a support vector machine (SVM) classifier to discriminate numerosity on the spiking activity of either VIP or PFC neurons [32] (see the Experimental Procedures). Preference-balanced pseudo-populations of 200 neurons were assembled per cortical area. Figures 6A and 6B show the temporal cross-training performance of the VIP and PFC classifiers, i.e., their accuracy to identify the correct numerosity when tested on the activity from a certain trial time period after being trained on other time bin. With a chance performance of 20% (for five classes), the classifier accuracy was higher in prefrontal than in parietal neurons throughout the sample phase (VIP: 50.1% ± 7.7%; PFC: 67.8% ± 7.1%; mean ± SD over resamples; training and testing in the same time bin). In addition, classification performance reflected the effects described in behavior: accuracy decreased along the diagonal of the confusion matrix with increasing numerosities (size effect), and the probability of misclassification of trials increased the closer two classes are in the numerical space (distance effect) (Figures 6C and 6D).

The confusion matrix in Figure 6C shows a robust accuracy for empty sets and numerosity 1 but weaker accuracy for other numerosities in VIP neurons (see also the resulting flat performance curves derived from the confusion matrix in the top panel of Figure 6C). In contrast, classification performance with PFC neurons was robust for all stimulus classes (Figure 6D), which was also reflected by sharp performance curves (Figure 6D, top panel).

Next, we assessed the ability of the classifier to discriminate each class (sample stimulus) from all others (Figures 7A and

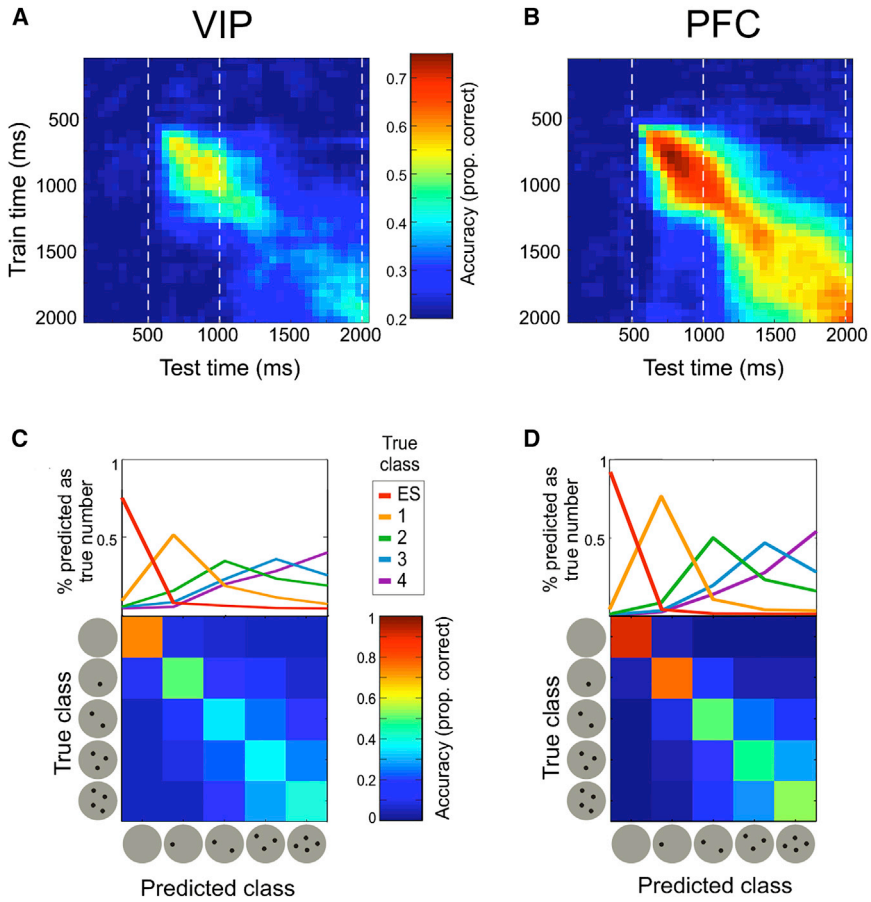


**Figure 6. Decoding Numerosity from VIP and PFC Population Activity with a SVM Classifier**

(A and B) Temporal cross-training classification accuracy in VIP (A) and PFC (B) populations.

(C and D) Confusion matrices for VIP neurons (C) and PFC neurons (D) derived from the sample period when training and testing were performed on activity from the same time bin. Performance curves for each true class are shown at the top of the confusion matrix. Each curve represents the frequency with which the activity elicited by a certain stimulus class was assigned different labels by the classifier.

See also Figures S4 and S5.

7B; evaluated with AUROC [area under the receiver operating characteristic curve], chance level 0.5). If magnitude classes were ordered along a numerical continuum, we would expect a graded decrease of discriminability with increasing numerical magnitude as a signature of the numerical size effect. In VIP, the average AUROC values during the sample period were 0.94 ± 0.04 for empty sets, 0.81 ± 0.09 for numerosity 1, and 0.68 ± 0.07 for larger numerosities (mean ± SD over resamples). Note that the classifier's capacity to discriminate numerosity 1 was similar to the discriminability of larger numerosities (Figure 7A). This indicates only a mild size effect in VIP neurons for countable numerosities. In contrast, the classifier's discriminability function for empty sets showed higher values compared to the discriminability of all countable numerosities. This decoding pattern again suggests that empty sets in VIP are treated more as a category different from other stimuli.

In PFC, however, a gradation of discriminability values from empty sets to higher numerosities was present (Figure 7B). Numerosity 1 was discriminated much better than other countable classes and slightly worse than empty sets (empty sets: 1.0 ± 0.01; numerosity 1: 0.95 ± 0.04; other numerosities: 0.80 ± 0.06; mean ± SD over resamples). Note the graded decrement in the discriminability of empty sets, numerosity 1, and larger numerosities in PFC (Figure 7B). This pattern evidences a numerical size effect at the population level and provides further evidence that prefrontal neurons integrate empty sets as part of the numerosity continuum. As expected from previous results, both classifiers did well in the binary discrimination of empty sets versus all other countable numerosities (Figure S4). A population of only tuned PFC neurons encoded numerosity better compared to a population of PFC and VIP cells (Figure S5).

### Level of Abstraction of Empty-Set Representations

To directly address the level of abstraction of empty-set representations at the whole-population level, we tested the effects of protocol and background shape with a decoding approach. We trained the SVM classifier on circle-background trials and tested it on square-background trials, and vice versa (shape generalization). We also tested generalization across protocols, with different background gray levels. Interestingly, the discriminability of empty sets by VIP neurons (0.94 ± 0.04, mean ± SD over resamples) dropped by 19.7% (0.85 ± 0.04) when training and testing were implemented in trials with different protocols. Similarly, discriminability decreased by 22% (0.84 ± 0.05) in shape generalization (Figure 7C). Parietal representations of empty sets are thus influenced to some degree by visual stimulus features.

Classification performance based on PFC neurons, however, fully generalized across sample appearance (Figure 7D). The classifier's ability to discriminate empty sets (1.0 ± 0.01) was not affected by training and testing on trials from different conditions. Discriminability decreased by only 0.2% in protocol generalization and by 0.8% in shape generalization. Thus, PFC representation of empty sets is invariant to background shape and gray level. This result points to a more abstract representation of empty sets in PFC, detached from stimulus appearance and low-level properties.

## DISCUSSION

### Monkeys Treat Empty Sets as Conveying a Null Numerical Value

The performance of both monkeys suggested that empty sets were positioned closer to numerosity 1 than to numerosity 2 on

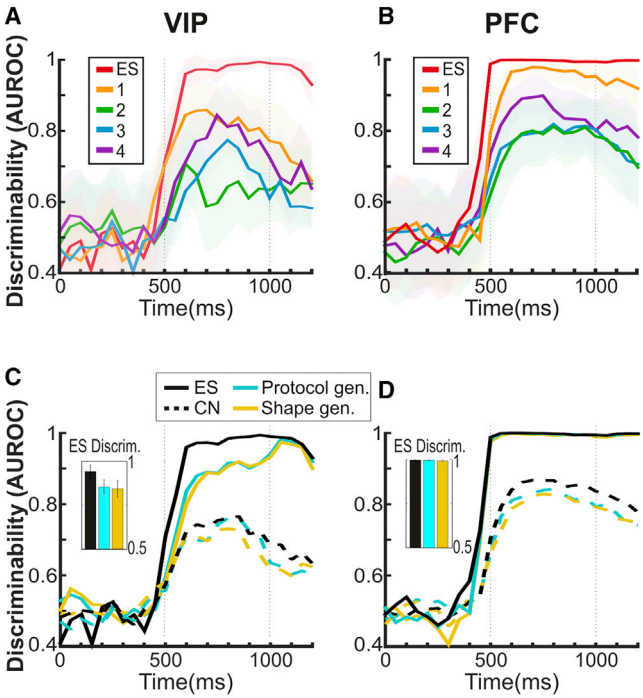


**Figure 7. Numerosity Discriminability and Generalization Performance in VIP and PFC**

(A and B) Discriminability of each sample numerosity versus others by the multi-class SVM classifier, evaluated with AUROC values. Numerosity was decoded from the firing activity of parietal (A) and prefrontal (B) general populations.

(C and D) Generalization performance of the SVM classifiers for VIP neurons (C) and PFC neurons (D) across different conditions. The discriminability of empty sets is depicted in solid lines, whereas dotted lines represent the discriminability of countable numerosities. Performance when training and testing were implemented in trials from different protocols (standard and control) is depicted in blue. Yellow represents generalization performance across different background shapes (circle and squares). The bars in the insets show the average discriminability of empty sets by the classifier during the sample period, when the stimulus is visually available to the subject. The performance achieved across different conditions is compared to the base performance (black lines and bars), when both training and testing were performed on mixed datasets, including trials from all conditions.

the monkeys' mental number line. This finding, a behavioral distance effect, signals a representational continuity between empty sets and countable numerosities and corroborates the conclusions of previous studies in primates [21–23]. Interestingly, behavioral findings in monkeys are reminiscent of the way pre-school children treat empty sets [20]. Moreover, a distance effect for empty sets has also been shown in adult numerate humans [20]. Humans and non-human primates treating empty sets similarly points to a common primitive and non-symbolic representation of null quantity.

## Neuronal Representations of Numerosity Zero in VIP

In addition to neurons tuned to countable numerosities [3, 9, 14–17], we identified a relatively high proportion of neurons that responded maximally to empty sets. Recently, Okuyama et al. [26] reported VIP neurons that were tuned to the absence of countable stimuli in one monkey. This monkey was trained to assess the numerosity of a target display (that could show no items) and add or subtract items in a second display to match the target numerosity. These authors classified neurons that responded maximally to empty sets into two distinct groups: exclusive/discrete types that showed no modulation to numerosities 1 to 4 (based on an ANOVA) and continuous types that exhibited a significant response to numerosity 1. Two-thirds of the empty-set neurons that they recorded in VIP were classified as a discrete type and the rest as a continuous type. No further statistical tests were applied to explore whether these cells belonged to two distinct classes. In our recordings from two monkeys, we found an even higher proportion of 93% VIP selective empty-set neurons that belonged to the discrete type class according to the definition of Okuyama et al. [26]. However, the values of tau ($\tau$) derived from exponential decay functions fitted to empty-set tuning functions did not result in bimodal distributions ($p = 0.81$, Hartigan dip test; $n = 28$). Thus, we found no evidence for two strict classes of empty-set neurons in VIP, but rather a continuum of more-discrete to more-continuous empty-set detectors. Task differences could account for this discrepancy, even though in both studies the neuronal recordings were derived from an initial target phase, when the numerosity the monkey had to match at the end of the trial was displayed. At this initial point of the trial, the task demands seem comparable. Rather, differences in stimulus protocols and data processing might account for the observed discrepancies. In particular, we controlled for the effect of visual stimulus features on neuronal numerosity responses. As discussed below, we found that VIP neurons represent empty sets primarily as separate category distinct from countable numerosities.

## Differential Encoding of Empty Sets in VIP and PFC

A differential tuning for empty sets emerged between parietal and frontal lobes. Empty-set neurons in VIP barely discriminated countable numerosities and failed to exhibit a strong neuronal distance effect. Empty sets seem to be encoded as a separate category, different from all other numerosity stimuli. The resulting binary-like tuning profile can be better described as signaling the presence or absence of countable items. Population analyses showed that this type of tuning is present in VIP neurons, irrespective of their selectivity and stimulus preference.

In contrast, empty-set neurons in PFC showed a gradual drop-off of activity with increasing numerosity. Moreover, at the whole-population level, the positioning of empty sets with respect to other numerosities was evidenced by a significant distance effect in the state space analysis. Complementing these findings, a size effect emerged in the discriminability of different samples by the classifier in PFC. These results suggest that PFC does integrate numerosity zero as the lower end of the numerical continuum.

The analysis of error trials also points to a differential integration of empty sets in VIP and PFC. Although the activity of empty-set neurons in both cortices was behaviorally relevant, only the activity of prefrontal countable-numerosity neurons during empty-set presentation affected trial outcomes. This finding suggests that prefrontal neurons integrate empty sets into the range of numerosities. In addition, as for behavioral performance functions, a logarithmic number line accounted better for the tuning curves of numerosity-selective neurons in PFC, but not in VIP. Thus, neuronal activity in PFC shows a stronger correlation with behavior than parietal activity.

### Level of Invariance of Empty-Set Representations to Stimulus Features

Some degree of abstraction is required from any neural correlate of numerosity. Therefore, we would expect a neural representation of empty sets to be invariant to image-like and low-level features of the stimuli. In our task, we varied background shape and controlled for luminance. Neurons sensitive to these non-numerical parameters were excluded from single-neuron analyses. Moreover, at the population level, a decoding approach allowed us to explore the invariance of empty-set representations in VIP and PFC. We found that the discrimination of empty sets by parietal neurons was reduced across stimulus features. This finding could be explained by a mixture of visual and numerical selectivity in VIP. In contrast, PFC represented empty sets abstractly, as evidenced by high decoding performance across stimulus properties.

Previously, the idea of abstract number representations was deemed premature based on some behavioral and human functional imaging studies [33]. In recent years, however, neurons indiscriminant to spatiotemporal and cross-modal number variations have been found, particularly in PFC [3, 9, 14–16]. Recent human imaging studies also report that the extraction of numerosity is only minimally influenced by the processing of physical stimulus features [4, 34]. These findings suggest that at least some neurons in association cortices represent numerosities abstractly. Of course, abstract number information could also be extracted from population activity [35], as evidenced by the analyses presented in the current study. Whether PFC neurons encode empty sets in different formats (across modalities and spatiotemporal presentation) requires further investigation.

### Hierarchical Processing of Empty Sets from VIP to PFC

In the context of previous studies, our results suggest a progressive transformation of empty-set representations from VIP to PFC. Empty sets become detached from visual properties and gradually positioned in a numerical continuum. Supporting this interpretation, simultaneous recordings have repeatedly reported that parietal neurons respond earlier to number than PFC cells [3, 9]. Moreover, PFC is known to host higher-level representations of magnitude. For example, it has been shown that PFC neurons, but not VIP neurons, respond supramodally to numerosity [16]. In addition, PFC neurons signify symbol-numerosity associations, whereas IPS neurons do not [36]. Finally, PFC sorts relevant from distracting information [37, 38] and processes magnitudes according to quantitative rules [39–42].

### Numerosity Zero in a Labeled-Line Code for Number

Several computational models of numerosity detection operate with intermediate-stage summation units that show monotonically increasing or decreasing discharges as a function of number (also found in area LIP [43]) before giving rise to peak-tuned numerosity detectors at the output stage [44, 45]. Due to the truncation of the number line, empty-set cells show decreasing rate functions reminiscent of decreasing summation units. On average, however, their tuning curves were too selective (i.e., narrow) to render them suitable graded summation units over a range of numerosities. This suggests that empty-set-preferring neurons are better considered as detectors tuned to numerical value 0. Conversely, the class of frequent neurons tuned to numerosity 4 may mirror increasing summation units. However, this class could include neurons preferring higher numerosities whose tuning curves has not been completely sampled. When broader ranges of numerosities (1 to 30) are tested, numerosity tuning preference becomes evenly distributed [46], supporting the notion that numerosity-selective cells in VIP and PFC are essentially tuned to specific numerical values (see also [26]).

### Numerosity Zero in a Non-linearly Compressed Number Line

We have previously reported that behavioral and neuronal representations of numerosity in monkeys [27, 28, 46] and crows [47, 48] are best described on a non-linearly compressed, logarithmic number scale. This finding is confirmed in the current study with a new set of data. The logarithmic scheme accounts for the decrease in the discrimination of two stimuli when their magnitude increases (as predicted by Weber-Fechner psychophysical law). A non-linearly compressed scaling of numerosity has the advantage of providing scale-invariance and preference-independent neuronal variability. Even though the logarithm of 0 is not defined, the differences between numerical values can still be represented on a log scale. Note that the Weber-Fechner law is concerned with the perception of differences, rather than absolute magnitudes. Starting with the interval between numerosity 0 (n) and numerosity 1 (n + 1), all differences between higher numbers can be represented on a log scale. Representations of cardinality 0 would therefore not dispute the notion of a nonlinearly compressed scaling.

### From Nothing to Zero

Sense organs have evolved to encode the intensity of a stimulus. Then, how can the absence of stimulation be detected? In order to make use of this information, the nervous system needs to encode it actively. Indeed, it has been shown that neurons in the frontal lobe increase their discharge rate to the categorical absence of a stimulus [24, 25]. Zero is an example of information conveyed by the lack of a signal. In this case, the brain generates a quantitative representation (zero) from the absence of a behaviorally relevant sensory signal (nothing), a process that would require a high level of cognitive control. In this context, it may not be surprising that neurons in PFC are particularly engaged in the representation of null quantity.

It has been argued that the conceptual demands imposed by representing nothing as a numerical category may explain the delayed discovery of zero in human history [49]. Zero first appeared as a placeholder symbol in notational systems. Only later, Indians used zero also as a numeral signifying null quantity in mathematics [50]. This cultural delay is mirrored in ontogeny: children seem to master the cardinal and ordinal properties of small numbers before they can deal with zero [18]. Still, it has been suggested that pre-school children understand the numerical value of numerosity 0 and position empty sets in the context of other small numerosities before they have developed a concept of symbolic zero [20]. These results suggest that the representation of empty sets as non-symbolic carriers of null quantity can be grasped by children and some animals. Our results suggest that a humble precursor of the non-symbolic zero can be identified in the primate PFC.

## EXPERIMENTAL PROCEDURES

### Subjects and Surgery

Two adult rhesus monkeys (*Macaca mulatta*) were implanted with two recording chambers each, centered over the principal sulcus in the dorsolateral PFC and the VIP in the posterior parietal cortex. All procedures were performed in accordance with the guidelines for animal experimentation approved by authorities (Regierungspräsidium Tübingen, Germany) (see the Supplemental Experimental Procedures for details).

### Stimuli

Numerosity stimuli were presented on an LCD screen and consisted of multiple-dot patterns against a gray background. To ensure that the monkeys solved the task by judging discrete quantity, we controlled low-level visual features in two stimulus protocols. In the standard protocol, black solid dots appeared at randomized locations and their diameter was pseudo-randomly varied. In the control protocol, overall dot area, dot density, and total stimulus luminance were kept constant across countable numerosities (1–4). Background luminance was varied across and between protocols to control for luminance differences that may occur for the empty set and to detect their effect on neuronal responses. For testing of how invariant the neuronal representation of empty sets is to image-like features, both stimulus protocols (standard and control) were shown either with a circular background (Figure 1B, left) or a square background (Figure 1B, right) (see the Supplemental Experimental Procedures for details).

### Behavioral Protocol

Monkeys were required to grab a bar and keep fixation in order to start a trial. Then, a green square or circle background appeared on the screen during a 500 ms fixation period. Subsequently, a sample stimulus consisting of a gray background containing zero to four dots was shown for 500 ms. After a 1 s delay, during which the green background was again shown, a test stimulus appeared and the monkeys were expected to release the bar if it matched the sample stimulus in quantity. That was the case in 50% of the trials, referred to as match trials. Otherwise, in non-match-trials, a 300 ms second delay was followed by a second test stimulus (500 ms) that always matched the sample stimulus in number. The green fixation and delay background displays framed the sample epoch. This background was chosen to match in luminance with the gray level displayed in the trial sample stimulus. Background shape was kept constant across the different epochs of a single trial, i.e., all displays in the circle-shape trials presented a circle as background, whereas all displays in the square-shape trials showed a square as background. Correct responses were rewarded with water (see the Supplemental Experimental Procedures for details).

### Behavioral Data Analysis

For each session, behavioral performance functions were derived from the percent of correct responses to all possible stimulus combinations. Behavioral curves per session were fitted with a Gaussian function when plotted either on a linear or a logarithmic numerical scaling ($\log_2(n + 1)$), and goodness-of-fit values ($r^2$) were derived. Overall performance tuning functions were obtained by averaging behavioral tuning functions over sessions (see the Supplemental Experimental Procedures for details).

### Neurophysiological Recording

In each session, arrays of up to eight glass-coated tungsten microelectrodes were inserted in each recording chamber using a grid with 1 mm spacing. Neurons were selected at random, as no attempt was made to preselect neurons according to response properties. Waveform sorting was performed off-line (see the Supplemental Experimental Procedures for details).

### Neuronal Data Analysis

#### *Selectivity and Tuning Analysis*

Neurons that had a minimum average firing rate of 1 Hz and at least three stimulus repetitions per specific condition (20 specific conditions from five sample numerosities × two types of protocol × two background shapes) were analyzed. For determination of numerosity selectivity, activity during the sample phase was derived from a 500 ms interval after stimulus onset. To account for differences in the response latencies between brain areas, we shifted the analysis window by 50 ms after sample onset for VIP and 100 ms for PFC neurons. This differential shift captured the intervals after physical stimulus onset when neurons are responsive to the sample stimulus [3, 9]. To determine numerosity selectivity of individual neurons, we ran a three-way ANOVA with factors number (five sample numerosities), protocol (standard and control), and shape (circle and square). Significance was evaluated for each factor at $p < 0.01$. For creation of neuronal filter functions, activity rates were normalized by setting, for each neuron, the firing rate to the most-preferred numerosity as 1 and to the least-preferred numerosity as 0. The normalized individual tuning curves were then averaged across neurons with the same preferred numerosity. The responses of individual empty-set neurons to countable numerosities (1 to 4) were fitted with a linear function to derive a slope. The tuning curves of empty-set neurons were also fitted with an exponential decay function to derive the time constant tau ($\tau$) as a measure of the decrease in firing rates as a function of number. Finally, the tuning curves of all selective neurons were fitted by a Gaussian function to evaluate tuning width (sigma, $\sigma$) and derive goodness-of-fit values ($r^2$) in different numerical scalings (see the Supplemental Experimental Procedures for details).

#### *Error Trial Analysis*

For comparison of firing rates in error and correct trials, we included selective neurons with at least three non-correct trials per involved stimulus. Because mistakes were rare, especially in trials involving small sample numerosities, a reduced number of neurons was included in the comparisons of correct and error trials.

#### *Population Analyses*

A GPFA [31] was used to extract low-dimensional neuronal trajectories in state space from the spiking activity of similarly sized pseudo-populations of neurons in both cortical areas (see the Supplemental Experimental Procedures for more information). In addition, a multi-class linear SVM classifier [32] was trained and tested on trial firing rates to discriminate sample numerosity. We evaluated the decoding performance of the classifier in balanced VIP and PFC pseudo-populations of neurons (see the Supplemental Experimental Procedures for details).

## SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, five figures, and one table and can be found with this article online at http://dx.doi.org/10.1016/j.cub.2016.03.052.

## AUTHOR CONTRIBUTIONS

A.N. and A.R.-C. designed the research; A.R.-C. performed the research; M.M. assisted with experiments; A.R.-C. analyzed the data; A.N. provided analytical guidance; and A.N. and A.R.-C wrote the paper.

## ACKNOWLEDGMENTS

This work was supported by DAAD Research Scholarship 91540420 to A.R.-C. and by DFG grant NI 618/8-1 to A.N. We thank J.F. Ramirez-Villegas for insightful comments on decoding methods and P. Viswanathan for proofreading the manuscript.

Received: December 13, 2015
Revised: March 10, 2016
Accepted: March 24, 2016
Published: April 21, 2016

## REFERENCES

1. Burr, D., and Ross, J. (2008). A visual sense of number. Curr. Biol. *18*, 425–428.
2. Dehaene, S. (1997). The Number Sense: How the Mind Creates Mathematics (Oxford University Press).
3. Viswanathan, P., and 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.

4. Park, J., DeWind, N.K., Woldorff, M.G., and Brannon, E.M. (2016). Rapid and direct encoding of numerosity in the visual stream. Cereb. Cortex *26*, 748–763.
5. Halberda, J., Mazzocco, M.M., and Feigenson, L. (2008). Individual differences in non-verbal number acuity correlate with maths achievement. Nature *455*, 665–668.
6. Starr, A., Libertus, M.E., and Brannon, E.M. (2013). Number sense in infancy predicts mathematical abilities in childhood. Proc. Natl. Acad. Sci. USA *110*, 18116–18120.
7. Brannon, E.M., and Terrace, H.S. (1998). Ordering of the numerosities 1 to 9 by monkeys. Science *282*, 746–749.
8. Nieder, A. (2005). Counting on neurons: the neurobiology of numerical competence. Nat. Rev. Neurosci. *6*, 177–190.
9. Nieder, A., and Miller, E.K. (2004). A parieto-frontal network for visual numerical information in the monkey. Proc. Natl. Acad. Sci. USA *101*, 7457–7462.
10. Piazza, M., Izard, V., Pinel, P., Le Bihan, D., and Dehaene, S. (2004). Tuning curves for approximate numerosity in the human intraparietal sulcus. Neuron *44*, 547–555.
11. Nieder, A., and Dehaene, S. (2009). Representation of number in the brain. Annu. Rev. Neurosci. *32*, 185–208.
12. Harvey, B.M., Klein, B.P., Petridou, N., and Dumoulin, S.O. (2013). Topographic representation of numerosity in the human parietal cortex. Science *341*, 1123–1126.
13. Eiselt, A.K., and Nieder, A. (2016). Single-cell coding of sensory, spatial and numerical magnitudes in primate prefrontal, premotor and cingulate motor cortices. Exp. Brain Res. *234*, 241–254.
14. Nieder, A., Freedman, D.J., and Miller, E.K. (2002). Representation of the quantity of visual items in the primate prefrontal cortex. Science *297*, 1708–1711.
15. Nieder, A., Diester, I., and Tudusciuc, O. (2006). Temporal and spatial enumeration processes in the primate parietal cortex. Science *313*, 1431–1435.
16. Nieder, A. (2012). Supramodal numerosity selectivity of neurons in primate prefrontal and posterior parietal cortices. Proc. Natl. Acad. Sci. USA *109*, 11860–11865.
17. Viswanathan, P., and Nieder, A. (2015). Differential impact of behavioral relevance on quantity coding in primate frontal and parietal neurons. Curr. Biol. *25*, 1259–1269.
18. Wellman, H.M., and Miller, K.F. (1986). Thinking about nothing: development of concepts of zero. Br. J. Dev. Psychol. *4*, 31–42.
19. Menninger, K. (1969). Number Words and Number Symbols (MIT Press).
20. Merritt, D.J., and Brannon, E.M. (2013). Nothing to it: precursors to a zero concept in preschoolers. Behav. Processes *93*, 91–97.
21. Beran, M.J. (2012). Quantity judgments of auditory and visual stimuli by chimpanzees (Pan troglodytes). J. Exp. Psychol. Anim. Behav. Process. *38*, 23–29.
22. Biro, D., and Matsuzawa, T. (2001). Use of numerical symbols by the chimpanzee (Pan troglodytes): Cardinals, ordinals, and the introduction of zero. Anim. Cogn. *4*, 193–199.
23. Merritt, D.J., Rugani, R., and Brannon, E.M. (2009). Empty sets as part of the numerical continuum: conceptual precursors to the zero concept in rhesus monkeys. J. Exp. Psychol. Gen. *138*, 258–269.
24. Merten, K., and Nieder, A. (2012). Active encoding of decisions about stimulus absence in primate prefrontal cortex neurons. Proc. Natl. Acad. Sci. USA *109*, 6289–6294.
25. Merten, K., and Nieder, A. (2013). Comparison of abstract decision encoding in the monkey prefrontal cortex, the presupplementary, and cingulate motor areas. J. Neurophysiol. *110*, 19–32.
26. Okuyama, S., Kuki, T., and Mushiake, H. (2015). Representation of the Numerosity 'zero' in the Parietal Cortex of the Monkey. Sci. Rep. *5*, 10059.
27. Nieder, A., and Miller, E.K. (2003). Coding of cognitive magnitude: compressed scaling of numerical information in the primate prefrontal cortex. Neuron *37*, 149–157.
28. Merten, K., and Nieder, A. (2009). Compressed scaling of abstract numerosity representations in adult humans and monkeys. J. Cogn. Neurosci. *21*, 333–346.
29. Mante, V., Sussillo, D., Shenoy, K.V., and Newsome, W.T. (2013). Context-dependent computation by recurrent dynamics in prefrontal cortex. Nature *503*, 78–84.
30. Raposo, D., Kaufman, M.T., and Churchland, A.K. (2014). A category-free neural population supports evolving demands during decision-making. Nat. Neurosci. *17*, 1784–1792.
31. Yu, B.M., Cunningham, J.P., Santhanam, G., Ryu, S.I., Shenoy, K.V., and Sahani, M. (2009). Gaussian-process factor analysis for low-dimensional single-trial analysis of neural population activity. J. Neurophysiol. *102*, 614–635.
32. Chang, C.-C., and Lin, C.-J. (2011). LIBSVM: a library for support vector machines. ACM Trans. Intell. Syst. Technol. *2*, 1–27.
33. Cohen Kadosh, R., and Walsh, V. (2009). Numerical representation in the parietal lobes: abstract or not abstract? Behav. Brain Sci. *32*, 313–328, discussion 328–373.
34. Leibovich, T., Vogel, S.E., Henik, A., and Ansari, D. (2016). Asymmetric Processing of Numerical and Nonnumerical Magnitudes in the Brain: An fMRI Study. J. Cogn. Neurosci. *28*, 166–176.
35. Tudusciuc, O., and Nieder, A. (2007). Neuronal population coding of continuous and discrete quantity in the primate posterior parietal cortex. Proc. Natl. Acad. Sci. USA *104*, 14513–14518.
36. Diester, I., and Nieder, A. (2007). Semantic associations between signs and numerical categories in the prefrontal cortex. PLoS Biol. *5*, e294.
37. Lennert, T., and Martinez-Trujillo, J. (2011). Strength of response suppression to distracter stimuli determines attentional-filtering performance in primate prefrontal neurons. Neuron *70*, 141–152.
38. Jacob, S.N., and Nieder, A. (2014). Complementary roles for primate frontal and parietal cortex in guarding working memory from distractor stimuli. Neuron *83*, 226–237.
39. Bongard, S., and Nieder, A. (2010). Basic mathematical rules are encoded by primate prefrontal cortex neurons. Proc. Natl. Acad. Sci. USA *107*, 2277–2282.
40. Vallentin, D., Bongard, S., and Nieder, A. (2012). Numerical rule coding in the prefrontal, premotor, and posterior parietal cortices of macaques. J. Neurosci. *32*, 6621–6630.
41. Eiselt, A.K., and Nieder, A. (2013). Representation of abstract quantitative rules applied to spatial and numerical magnitudes in primate prefrontal cortex. J. Neurosci. *33*, 7526–7534.
42. Ott, T., Jacob, S.N., and Nieder, A. (2014). Dopamine receptors differentially enhance rule coding in primate prefrontal cortex neurons. Neuron *84*, 1317–1328.
43. Roitman, J.D., Brannon, E.M., and Platt, M.L. (2007). Monotonic coding of numerosity in macaque lateral intraparietal area. PLoS Biol. *5*, e208.
44. Dehaene, S., and Changeux, J.P. (1989). A simple model of prefrontal cortex function in delayed-response tasks. J. Cogn. Neurosci. *1*, 244–261.
45. Verguts, T., and Fias, W. (2004). Representation of number in animals and humans: a neural model. J. Cogn. Neurosci. *16*, 1493–1504.
46. Nieder, A., and Merten, K. (2007). A labeled-line code for small and large numerosities in the monkey prefrontal cortex. J. Neurosci. *27*, 5986–5993.
47. Ditz, H.M., and Nieder, A. (2015). Neurons selective to the number of visual items in the corvid songbird endbrain. Proc. Natl. Acad. Sci. USA *112*, 7827–7832.
48. Ditz, H.M., and Nieder, A. (2016). Numerosity representations in crows obey the Weber-Fechner law. Proc. Biol. Sci. *283*, 20160083.
49. Butterworth, B. (1999). The Mathematical Brain (Macmillan Press).
50. Ifrah, G. (1998). Universal History of Numbers: From Prehistory to the Invention of the Computer (The Harvill Press).