Behavioral/Systems/Cognitive

A Labeled-Line Code for Small and Large Numerosities in the Monkey Prefrontal Cortex

Andreas Nieder and Katharina Merten

Department of Cognitive Neurology, Primate Neurocognition Laboratory, Hertie-Institute for Clinical Brain Research, University of Tuebingen, 72076 Tübingen, Germany

How single neurons represent information about the magnitude of a stimulus remains controversial. Neurons encoding purely sensory magnitude typically show monotonic response functions ("summation coding"), and summation units are usually implemented in models of numerosity representation. In contrast, cells representing numerical quantity exhibit nonmonotonic tuning functions that peak at their preferred numerosity ("labeled-line code"), but the restricted range of tested quantities in these studies did not permit a definite answer. Here, we analyzed both behavioral and neuronal representations of a broad range of numerosities from 1 to 30 in the prefrontal cortex of monkeys. Numerosity-selective neurons showed a clear and behaviorally relevant labeled-line code for all numerosities. Moreover, both the behavioral and neuronal tuning functions obeyed the Weber–Fechner Law and were best represented on a nonlinearly compressed scale. Our single-cell study is in good agreement with functional imaging data reporting peaked tuning functions in humans, demonstrating neuronal precursors for human number competence in a nonhuman primate. Our findings also emphasize that the manner in which neurons encode and maintain magnitude information may depend on the precise task at hand as well as the type of magnitude to represent and memorize.

Key words: monkey; number; prefrontal cortex; psychophysics; single-cell recording; neuronal coding

Introduction

Quantitative judgments rely on magnitude representations. Sensory magnitude may be coded directly at the level of the sensory epithelium. Most abstract categories, such as the number of elements of a set, however, need to be derived at integrative processing stages of the brain.

Currently, two main hypotheses have been proposed for the neuronal coding of magnitude. The "summation coding" (also termed "monotonic coding") hypothesis assumes that quantity is encoded by the monotonically increasing and decreasing response functions of the neurons (see Fig. 1a). In technical terms, such neuronal response functions are equivalent to low-pass and high-pass filter functions. Summation coding has been demonstrated for sensory magnitude processing. For instance, prefrontal neurons encoding the magnitude (i.e., the frequency) of tactile vibration have shown exclusively monotonic functions (Romo et al., 1999; Brody et al., 2003; Romo and Salinas, 2003). Computational models of numerosity representation (Meck and Church, 1983; Zorzi and Butterworth, 1999; Zorzi et al., 2005) favor summation coding, which seems to be supported by psychophysical data (Roggeman et al., 2007).

An alternative coding mechanism is the "labeled-line code" (also termed "place code"). Here, the magnitude of a stimulus is encoded by the maximum response rate of a particular neuron and the tuning function is a peak function with the preferred magnitude represented at the maximum of discharge. Technically speaking, quantity-selective neurons act like bandpass filters. Gaussian-shaped tuning curves have been found for the representation of small numerosities in the prefrontal (PFC) (Nieder et al., 2002; Nieder and Miller, 2003) and posterior parietal cortices (PPC) of monkeys (Sawamura et al., 2002; Nieder and Miller, 2004a; Nieder et al., 2006). These findings are consistent with a functional imaging study measuring indirectly peak numerosity tuning functions in humans (Piazza et al., 2004), as well as with neural network models of numerical cognition (McCloskey and Lindemann, 1992; Dehaene and Changeux, 1993). Computationally, summation coding and labeled-line coding may not be mutually exclusive; peaked numerosity detectors may also arise after integration of preceding summation units (Dehaene and Changeux, 1993; Verguts and Fias, 2004).

Whether abstract, numerical quantity is encoded via a labeled-line code, and if so, whether the representations still obey the Weber–Fechner Law (Weber, 1850; Fechner, 1860), remains controversial because of the restricted numerosity range, from 1 to 5, tested so far (Nieder et al., 2002, 2006; Nieder and Miller, 2003, 2004a). It has been argued that tuned numerosity-selective neurons could be a special feature of the "subitizing" (object tracking) mechanism providing access to small numerosities (Mandler and Shebo, 1982; Feigenson et al., 2004) and that neu-

Figure 1. Schematic of the two different coding hypotheses. *A*, Summation coding (monotonic coding). Discharge rates vary as a monotonic function (increasing or decreasing) of the numerical quantity of the stimulus. The number of activated units (illustrated by circles) encodes the numerical magnitude, much like an accumulator (shown for numbers 5 and 10). Both decreasing (shown here) and increasing (data not shown) monotonic discharge functions may be implemented. *B*, Labeled-line code. A particular unit firing maximally at its preferred numerosity indicates the numerical magnitude of the stimulus. The tuning function is a peak function reaching its maximum at the preferred numerosity.

rons selective to the most extreme numerosities (1 and 5) that were categorized as bandpass-filtering neurons would actually be low-pass or high-pass filters (Roggeman et al., 2007).

Here, we recorded single-unit activity in the PFC of monkeys discriminating numerosities from 1 to 30. We only found evidence for a labeled-line code (i.e., numerosity-selective neurons exclusively exhibited peaked tuning profiles). Consistent with the behavioral data, the neuronal tuning functions obeyed the Weber–Fechner Law and were best described on a logarithmically compressed scale.

Materials and Methods

Subjects. Two male adult rhesus monkeys (*Macaca mulatta*) weighing 5.8 and 10.2 kg, implanted with recording hardware, were used in this study to test the coding hypotheses (Fig. 1). All surgeries were performed under sterile conditions while the animals were anesthetized with isoflurane. The animals received postoperative antibiotics and analgesics. All procedures were in accordance with the guidelines for animal experimentation approved by the Regierungspräsidium Tübingen, Germany. All presented behavioral data were collected during recordings. Both monkeys are still engaged in quantity discrimination.

Stimuli. Numerosity stimuli consisting of multiple-dot patterns were generated using a custom-written MatLab software. These routines enabled the generation of new stimuli sets for each training session. Moreover, this software provided for the control of parameters of the dot patterns. For the standard stimuli, small black filled dots (diameter of 0.17–0.28° visual angle) appeared on a gray background of a large circular area with a diameter of 7° visual angle. Each stimulus contained a defined set of dots that appeared at randomized locations within the gray background circle. The diameter of each dot was randomly varied within the given range. To prevent the monkeys from memorizing the visual patterns of the displays, each quantity was tested with many different images per session and the sample and test displays that appeared on each trial were never identical. To ensure that the numerosity-discrimination task was solved by judging the discrete quantity, low-level visual features were excluded using control stimuli in addition to standard stimuli. Two sets of control stimuli were used alternately in each session (Fig. 2*b*): area

control (total area of all items in a display equated for all stimuli in a trial) and density control (same mean density of dot patterns for all stimuli in a trial).

Behavioral protocol. Monkeys grasped a lever and fixated a central target to start a trial (see Fig. 2*a*). A sample display (500 ms) was followed by a memory delay (1000 ms). Next, a test display appeared, which was either a match (containing the same number of dots as the sample display) or a nonmatch (containing, with equal probability, more or fewer items). If the first test display was a match, monkeys released the lever to receive a juice reward. If the first test display was a nonmatch, the monkeys held the lever until a second test display (always a match) appeared that required to release the lever for a reward. Trials were randomized and balanced across all relevant features. Monkeys were required to maintain their gaze within 1.75° of visual angle of the fixation point during sample presentation and the memory delay (monitored with an infrared eye-tracking system, ISCAN, at a sample rate of 120 Hz).

Numerosities 1, 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, and 30 were presented as samples. For each sample stimulus, corresponding match and nonmatch stimuli were generated. Match stimuli were composed of the same numerosity of items like the sample, but with different arrangements of the dots and different dot sizes. The numerosities of nonmatch displays (N_{NM}) contained equidistant larger or smaller quantities of items compared with the sample numerosity (N_S). The nonmatch numerosities were calculated using the deviations $x = 0.3$ and 0.6, according to the equation $N_{NM} = \text{round}(N_S \times (1 \pm x))$. Consequently, the four nonmatch numerosities were arranged symmetrically around the sample numerosity (two smaller and two larger than the sample numerosity).

Neurophysiological recording. PFC recording chambers were placed stereotaxically over the principal sulcus and anterior to the arcuate sulcus using structural magnetic resonance imaging (MRI) scans acquired before surgery. Recordings were made from one left and one right hemisphere in two monkeys, covering the principal sulcus and the ventral bank of the PFC. Arrays of eight glass-coated tungsten microelectrodes (Alpha Omega, Nazareth, Israel) were inserted using a grid (Crist Instruments, Hagerstown, MD) with 1 mm spacing. Neurons were selected at random. Waveform separation was performed off-line (Plexon, Dallas, TX).

Data analysis. Sample activity was derived from a 500 ms interval after stimulus onset shifted by the individual response latency of a cell. To measure neuronal response latency, we generated average spike density histograms (at 1 ms resolution, smoothed by a sliding window; kernel bin width, 10 ms) for a neuron's responses to all sample stimuli. Discharges after sample onset were compared with spike rates in a 200 ms interval preceding sample onset. Response latency was defined by the first time bin that reached a value higher or lower than any value before sample onset. A default latency of 100 ms was used if no measure based on these criteria could be derived. For the delay period, activity was summed in a 800 ms interval starting 200 ms after delay onset.

To determine numerosity selectivity of the neurons, a two-factorial ANOVA was performed with numerosity (1–30) and stimulation condition (standard or control) as factors, for the sample and delay periods separately. Only cells showing a significant main effect for numerosity ($p < 0.01$), but no significant main effect for stimulus type (standard vs control) or interaction were classified as "numerosity selective," and the numerosity eliciting the largest spike rate was defined as the "preferred numerosity" of a given cell.

A cross-validation analysis was performed to estimate the reliability of preferred numerosity determination. The preferred numerosity of a cell was determined with one-half of the data (i.e., half of the randomly shuffled trials) and correlated to the firing activity obtained with the other half of the data. The preferred numerosity derived with the first half of the data were plotted against the preferred numerosity obtained with the second half of the data. This was done for the entire population of numerosity-selective neurons and the relationship between preferred numerosities in both data sets was quantified with a simple regression technique: $y = a + b \times x$, where y is the preferred numerosity of the neurons for the first half of the data, x is the preferred numerosity of the neurons for the second half of the data, a is the intercept, and b is the slope of the x–y relationship. If both data sets resulted in identical preferred

numerosities, Pearson's linear correlation coefficient is 1. The correlation analysis was repeated 100 times with newly shuffled trials and the average correlation coefficient was calculated.

To create neural filter functions, activity rates were normalized by setting the maximum activity to the most preferred numerosity as 100% and the activity to the least preferred quantity as 0%. The normalized individual tuning curves were then averaged across all neurons that had the same preferred numerosity. Gaussian functions were fit to the pooled performance functions of both monkeys and the pooled neural filter functions of the sample and delay period (χ^2-minimization after Levenberg–Marquardt). The Gaussian was chosen because it represents the standard symmetric distribution and, thus, provided a means to compare the behavioral functions. Data were plotted along four scales: a linear scale, a power function with exponent of 0.5, a power function with exponent of 0.33, and a logarithmic scale. The scales become increasingly nonlinearly compressed along this sequence. The more symmetrical the filter functions on a particular scale, the better this fit and, therefore, the better that scale describes the data. These nonlinearly compressed scaling schemes were chosen because Stevens' power law ($S = k \times I^n$) postulates that sensation S is a power function of the stimulus magnitude I (Stevens, 1961), whereas Fechner's Law ($S = k \times log(I)$) proposes a logarithmic relationship (Fechner, 1860).

The Weber threshold is defined as the 50% correct discrimination between two stimuli. The Weber fraction for numerosity is (Weber, 1850; Van Oeffelen and Vos, 1982) as follows:

$$W_b = \frac{(\text{max} - \text{min})}{\text{min}},$$

where max and min are the upper and lower cutoff values of the performance/neural-filter functions. The Weber fractions for the behavioral and neuronal analysis were derived after plotting the data on a logarithmic scale, which best described the data, at 50% of the maximum amplitude.

Results
Behavioral data
Two monkeys were trained on a delayed match-to-numerosity task (Fig. 2a) that required them to judge whether successive visual displays contained the same number of pseudorandomly placed items (Nieder et al., 2002). Both monkeys were initially trained to discriminate numerosities 1–4. After they mastered these small numerosities reliably, the range of numerosities was extended up to 30 (presented in steps of two starting at numerosity 2).

To perform this task, monkeys needed to abstract the set size of items from multiple-dot patterns that varied widely in appearance and then retain that information in memory over a delay period. Performance on sets of control stimuli confirmed that the monkeys were relying on abstract quantity information rather than on the exact appearance of the displays or lower-level visual features (area or density of the dots). Monkeys made more errors when the numerosities were adjacent and performed progressively better as numerical distance between two displays increased (numerical distance effect) (Fig. 3a). For larger quantities, the two numerosities had to be numerically more distant for performance to reach the same discrimination level obtained with smaller quantities at closer numerical distance (numerical size effect).

As shown previously for small numerosities (Nieder and Miller, 2003), the distributions of both monkeys' performances were asymmetric when plotted on a linear scale (Fig. 3a); slopes were more moderate for numerosities larger than the sample numerosity (represented by the center of each distribution) than for numerosities smaller than the sample. However, when plotted on a logarithmic scale, the distributions became more symmetric
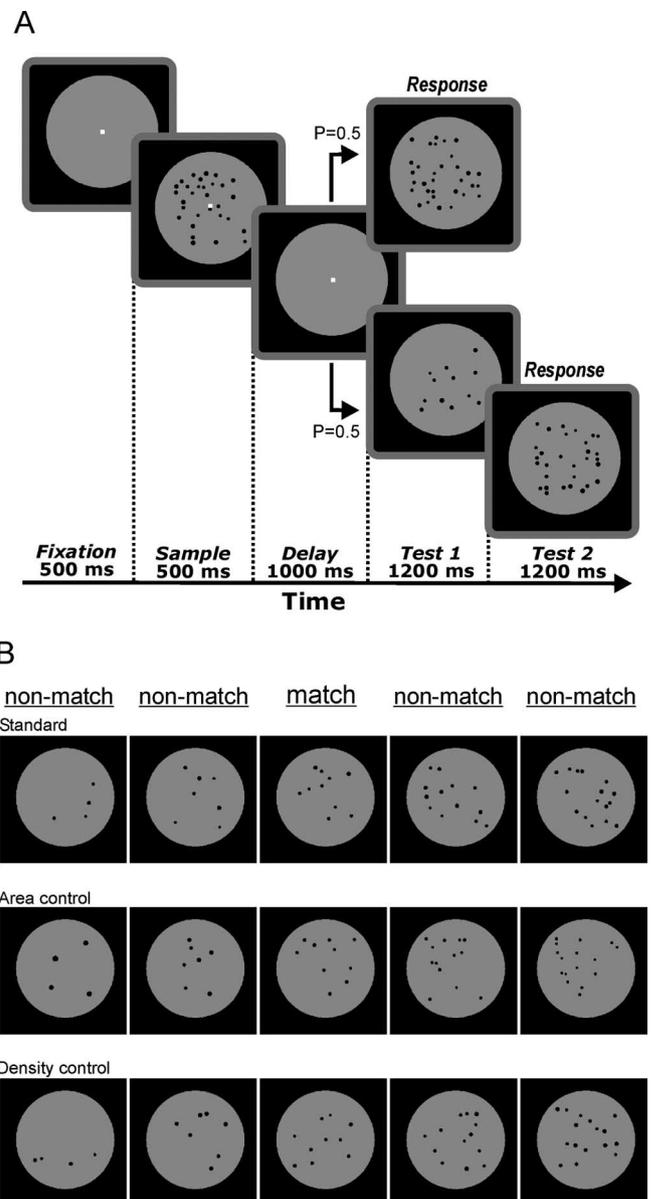
Figure 2. Delayed match-to-numerosity task and example stimuli. **A**, Fixating monkeys were cued for a given numerosity ranging from 1 to 30 by a sample display. The subjects had to memorize the numerosity in a 1 s delay period and match it to a subsequent test stimulus (either the first or the second test stimulus was correct) by releasing a lever. For each sample numerosity, four possible nonmatch numerosities were shown (two smaller than the sample by a factor of 1.3 and 1.6, and two larger by the same factors). **B**, Example stimuli (standard, area control, and density control) for sample numerosity 10, which had to be discriminated from numerosities 4, 7, 13, and 16.

(Fig. 3b), suggesting that a nonlinear coding scheme might be more appropriate. To quantify this finding, we first determined whether linear or nonlinear scaling models provided a superior Gaussian fit to the behavioral data.

We plotted the data using four scales: a linear scale, a power function with an exponent of 0.5, a power function with an exponent of 0.33, and a logarithmic scale (see Materials and Methods). All three nonlinear scales resulted in significantly better goodness-of-fit values (r^2) than the linear scale ($p < 0.05$, Wilcoxon signed ranks test, two-tailed). The mean goodness-of-fit values for the linear scale, power function with exponent of 0.5, power function with exponent of 0.33, and logarithmic scale were 0.76, 0.87, 0.89, and 0.92, respectively (Fig. 3c). (Performance
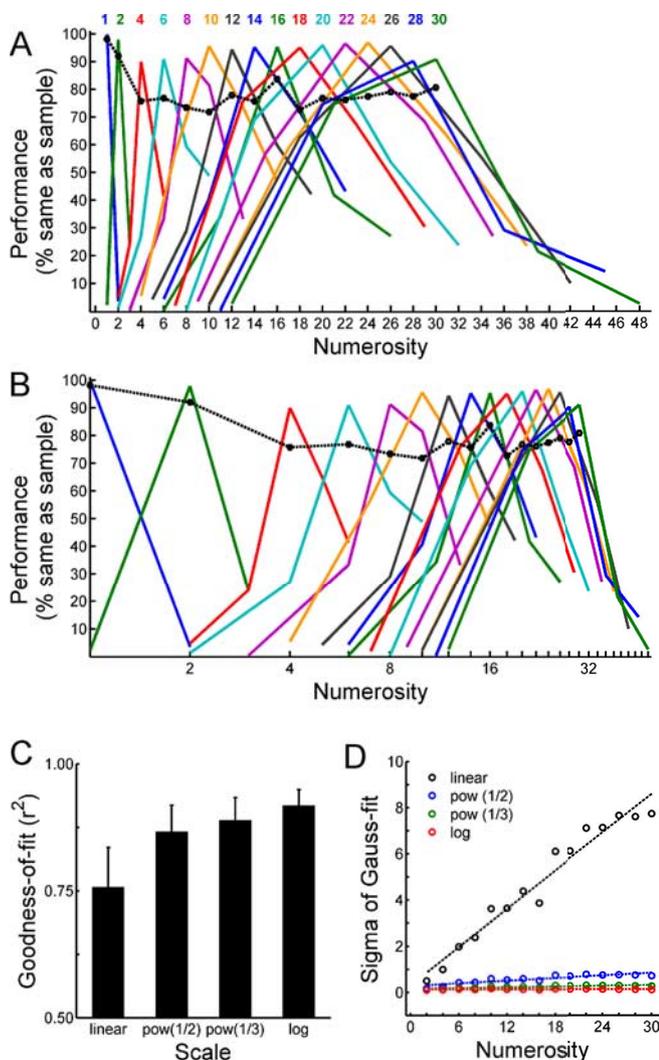
Figure 3. Behavioral performance. *A*, Behavioral tuning curves derived from the monkeys performance to the broad range of numerosities. The functions reflect the probability that a monkey judged displays in the test period as containing the same number of items as the sample numerosity. The center data point of each colored function indicates the correct performance in the match trials for the sample numerosities (shown in the same color above each curve). The data points to the left and the right of the center reflects performance in the nonmatch trials (i.e., where the first test display showed a smaller or larger number of items), presented as the percentage of errors for the respective nonmatch numerosity. *B*, Same data as in *A* plotted on a logarithmic number scale. *C*, Goodness-of-fit of Gaussian functions fitted to the performance curves plotted on different scales. The goodness-of-fit was significantly better for the three nonlinear scaling schemes (error bars ± SEM). *D*, Half-width (SD "σ") of the Gaussian fits for linear and nonlinear scalings plotted against the center of the Gaussian function (which is identical to the numerosity of the match stimulus). The dotted lines indicate linear fits. (The values of σ are related to the specific compression scheme.)

data for sample numerosities higher than 20 were excluded from statistical analysis because the monkeys may have learned that the very largest numerosities only appeared as nonmatch stimuli, made fewer errors in such cases, and distorted the performance curves.) Furthermore, the half-width of the distributions for each numerosity (i.e., σ of the Gaussian fit to the performance curves) increased proportionally with an increase of numerosity on a linear scale (slope of linear fit, 0.309), but decreased to almost constant values when the data were plotted on a power-function scale with exponents of 0.5 (slope, 0.027) and 0.33 (slope, 0.009), or a logarithmic scale (slope, 0.001) (Fig. 3d). This result is predicted by a nonlinear coding model of numerosity (Van Oeffelen

and Vos, 1982; Dehaene and Changeux, 1993; Dehaene, 2001). In other words, no differences in the widths of the distributions were observed between small and large numerosities when plotted on a logarithmic scale. For both small and large numerosity judgments, thus, performance data are better described by nonlinearly compressed scales, as opposed to a linear scale.

If the behavioral performance data obey Weber's Law, the Weber fractions for the broad range of tested numerosities should stay constant. We calculated the Weber fractions after fitting Gaussian functions to the data plotted on the best, logarithmic scale. The Weber fraction values (mean 0.42 ± 0.06 SD) were equal and constant across all small and large numerosities (see Fig. 8). Thus, all quantitative measures clearly indicate that Weber's Law holds for both small and large numerosity discriminations, and that the data are consistent with the nonlinear compression hypothesis (Nieder and Miller, 2003).

Single-cell data

We next analyzed the activity of 519 neurons recorded during task performance from the lateral PFCs of the two monkeys. Of these, 150 (29%) and 138 (27%) cells modulated their discharges as a function of the numerosity (1–30) during the sample or memory delay, respectively. This selectivity was found regardless of the exact appearance of the multiple-dot pattern in the standard, area-controlled, or density-controlled condition [only cells showing a significant numerosity effect, but no significant effect of stimulus type (standard vs control) or interaction, were classified as numerosity-selective neurons according to a two-factor ANOVA criterion, $p < 0.01$].

The neuronal responses showed clear numerosity tuning, with maximum discharge to a preferred numerosity and activity declining progressively with increasing numerical distance from the preferred numerosity (Fig. 4). Across the large population of single cells, we found neurons with preferred numerosities covering the entire range between numerosities 1 and 30 (Fig. 5). Occurrence of preferred numerosity 1 was most frequent, with a progressive drop-off toward higher preferred numerosities (Fig. 6). Only in the sample period, the highest preferred numerosities (26, 28, and 30) were slightly over-represented. Both small and large numerosities were neuronally represented by peak functions.

A cross-validation analysis (see Materials and Methods) was performed to estimate the reliability of preferred numerosity determination. The preferred numerosity for the population of numerosity-selective cells was highly reproducible, both for the sample (average correlation coefficient $r = 0.80$; $p < 0.0001$) and delay period ($r = 0.76$; $p < 0.0001$). This indicates that the preferred numerosity of the neurons was reliable and robust.

To analyze the tuning characteristics, mean neural tuning functions of the whole population were constructed by averaging activity across all neurons that preferred a given numerosity. Gaussian functions were fitted to the population tuning curves and quantitative measures [such as the goodness of fit (r^2), the half-width (SD "σ"), and the Weber fraction] (see Materials and Methods) were derived. The best scheme should result in most symmetric tuning functions. The goodness-of-fit (r^2) values of the Gaussian fits were taken as a quantitative measure of the symmetry of the tuning curve when plotted on a linear, power function with exponents 0.5 and 0.33, or a logarithmic scale. The (nonlinear) power function and logarithmic scales provided a better fit to the data than the linear scale ($p < 0.05$, paired-samples t test, two-tailed). No significant differences were found between the nonlinearly compressed scaling schemes (Fig. 7a).

As predicted by the numerical size effect based on Weber's Law, the half-width (σ of the Gaussian fit) of the neural distributions increased proportionally (slope for linear scale, 0.383; $p < 0.001$) with numerosity when the data were plotted on a linear scale (Fig. 7b). The widths of the distributions became constant across preferred numerosities when the data were plotted on increasingly nonlinearly compressed scales (slopes, power function with exponent 0.5, 0.029; power function with exponent 0.33, -0.004; log $= -0.009$).

As for the behavioral performance functions, we calculated the values of the Weber fractions for the neuronal population data plotted on the logarithmic scale (see Materials and Methods). Consistent with logarithmic coding, the Weber fractions (mean 1.20 ± 0.40 SD) were approximately equal and constant across the broad range of numerosities (Fig. 8). Thus, all of our analyses support the notion of a nonlinearly compressed coding scheme for both small and large numerosity information.

Behavioral relevance of neuronal responses

Next, we compared behavioral and neural Weber fractions for all numerosities. On average, the Weber fraction ratio between behavioral and neuronal data was 2.8, indicating an almost three-fold greater sensitivity on the behavioral than the neural level (Fig. 8). The Weber fraction ratio remained constant across all

Figure 4. Example numerosity-selective neuron. **A**, Dot raster display (each dot represents an action potential) showing the detailed temporal response pattern of the neuron during several trial repetitions for the complete range of numerosities tested. **B**, Tuning function derived by averaging discharges of the same neuron over the sample period resulted in a peaked function with numerosity 20 eliciting the highest spike rate. Error bars indicate SEM.

tested numerosities (slope, 0.03, linear fit), although absolute selectivity decreased with increasing numerosity, both at the behavioral and neural levels. This result suggests a direct relationship between behavioral and neuronal representations.

If the responses of numerosity-selective neurons are directly related to the monkeys' categorical judgments, then numerosity tuning should be degraded in trials in which monkeys made erroneous decisions. Therefore, we compared the responses of all selective neurons to the preferred numerosity in correct trials

Figure 5. **A–H**, Example tuning functions of sample numerosity-selective neurons with different preferred numerosities 1 (**A**), 2 (**B**), 4 (**C**), 6 (**D**), 20 (**E**), 20 (**F**), 28 (**G**), and 30 (**H**). Error bars indicate SEM.
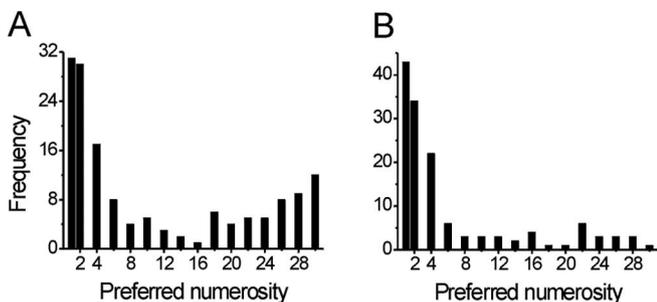
Figure 6. *A*, *B*, Frequency histograms showing the proportions of neurons tuned to different preferred numerosities in the sample (*A*) and delay (*B*) period, respectively. Each of the displayed sample numerosities were preferred by individual neurons.
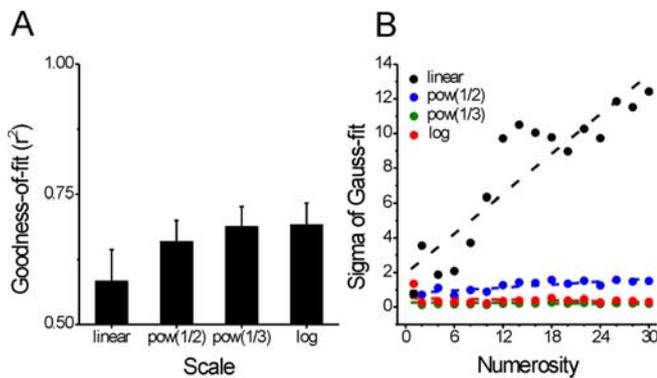
Figure 7. Characterization of neuronal tuning functions. *A*, Goodness-of-fit of Gaussian functions fitted to the tuning curves plotted on different scales (error bars ± SEM). *B*, Half-width (SD "σ") of the Gaussian fits for linear and nonlinear scalings plotted against the center of the Gaussian function. Dotted lines indicate linear fits.
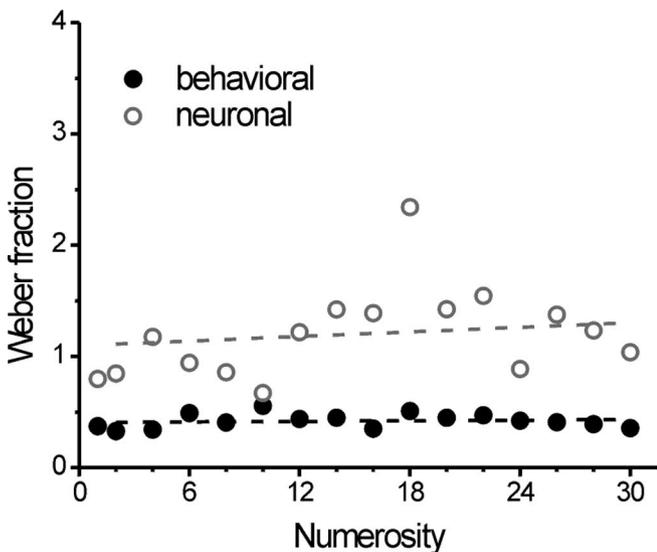
Figure 8. Weber fractions for the behavioral and neuronal tuning functions. The dotted lines indicate linear fits.

Figure 9. Comparison of neuronal tuning on correct and error trials. *A*, *B*, Normalized and averaged discharges during sample (*A*) and delay (*B*) epochs of all tested cells plotted against the numerical distance from the preferred number of items. Discharges to the preferred numerosity (numerical distance 0) were significantly reduced on error trials.

with error trials. Neural activity for the preferred numerosity was significantly reduced to 71.4 and 69.5% of that observed on correct trials (normalized to 100%) in the sample and delay epochs, respectively ($p < 0.001$, paired-samples t test). This finding is illustrated by a "clipping" of the averaged tuning peak on error trials (Fig. 9). This suggests that the monkeys tended to make

judgment errors whenever the neurons were not able to properly encode their preferred numerosity.

Discussion

We trained monkeys on a delayed match-to-sample task to discriminate 1–30 items on a visual display. Numerosity-selective neurons in the PFC showed clear peaked tuning functions, arguing for a pure labeled-line code implemented for explicit numerosity judgments. Both the behavioral and neuronal tuning functions obeyed the Weber–Fechner Law and were best represented on a compressed scale. An examination of error trials suggested that the activity of PFC neurons is directly related to the monkeys' performance. Our single-cell data are in good agreement with functional imaging data reporting peaked tuning functions in humans.

Behavior

The average Weber fraction of 0.42 in the current study with large numerosities was concordant with a previously reported value of 0.35 for small numerosities using the same task design (Nieder and Miller, 2003). Interestingly, humans seem to have a lower

Weber fraction of 0.17 when discriminating multiple-dot displays (Piazza et al., 2004). It needs to be emphasized, however, that only sample numerosities 16 and 32 were used in the study by Piazza et al. (2004), whereas in our approach the monkeys had to discriminate almost any given numerosity from the other, causing maximum task difficulty. A direct comparison of the performance of humans and monkeys using the very same task design should clarify whether this difference persists under identical conditions. Interestingly, humans' and monkeys' performances on an ordinal numerosity comparison task were found previously to be qualitatively and quantitatively similar (Cantlon and Brannon, 2006).

Analog magnitude system
Both the behavioral and neurophysiological data clearly demonstrate the presence of a pure analog magnitude system. The analog magnitude system can represent an unlimited amount of set sizes and shows all characteristics postulated by the Weber–Fechner Law, both for small and large numerosities. Thus, whenever animals need explicit access to numerical quantity, the analog magnitude system is activated (Nieder, 2005). In agreement with previous findings in explicitly discriminating monkeys (Brannon and Terrace, 1998; Nieder and Miller, 2004b) and other animals, evidence for an object tracking system (also termed subitizing) that is thought to allow precise representations, but only for up to four items (set-size limit; in particular when items are presented serially) (Feigenson et al., 2004), was absent. Thus, object tracking mechanisms remain restricted to studies exploiting spontaneous, unconditioned behavioral discriminations by accessing implicit representations of the number of objects (Hauser et al., 1996, 2000).

Labeled-line code in the primate cortex
The fact that we found neurons tuned to any presented numerosity from 1 to 30 unequivocally argues for the presence of a labeled-line code in explicit numerosity discriminations, regardless of the amount of magnitudes to represent. Our current data, together with previous reports about peak-tuned numerosity detectors in the parietal lobe (Sawamura et al., 2002; Nieder and Miller, 2004a; Nieder et al., 2006), suggest the presence of the same labeled-line code for explicit numerosity representations in the parietal lobe. Because we have shown that monkeys understand the concept of numerosity and can immediately apply their knowledge in transfer tests showing novel numerosities (Nieder et al., 2006), it is plausible to assume that peak tuned numerosity functions are spontaneously present and not forced by behavioral conditioning.

Our single-cell data in the nonhuman primate are in good agreement with functional imaging studies in humans. Using a functional MRI adaptation method, Piazza et al. (2004) investigated tuning to numerosity in humans. Functional MRI adaptation to numerosity was present in the bilateral intraparietal sulcus (IPS). Most notably, functional MRI adaptation revealed peaked tuning profiles in the human IPS, which seemed to obey the Weber–Fechner Law (Piazza et al., 2004). A previous study by the same group (Piazza et al., 2007) has shown that numerosity-related information does not only exist in the IPS, but also in areas of the human frontal lobe. Consistent with these studies, we now can show that Gaussian-shaped single-cell tuning curves in the monkey exist for a broad range of numerosities. This commonality suggests that humans and macaques have a comparable labeled-line coding mechanism for the representation of numerosities.

Absence of summation units in explicit numerosity representation
Many computational models of numerosity detection operate with summation units on a level preceding the emergence of peak-tuned numerosity detectors (Dehaene and Changeux, 1993; Verguts and Fias, 2004). Our new data set, however, favors a pure labeled-line code for numerosity. First, although neurons preferring numerosity 1 were over-represented, their response profiles were too selective (i.e., their slopes were too steep) to make them suitable as low-pass filters from which (in combination with high-pass filter functions) peak functions with different preferred numerosities could emerge. For instance, if the intersection between (presynaptic) high- and low-pass filters determined the center of (postsynaptic) peaked tuning functions, monotonic functions of variable slope (i.e., "cut-off" numerosities) would be necessary to cover the broad range of preferred numerosities. In keeping with this hypothesis, response profiles of neurons with preferred numerosity 1 should vary considerably with respect to shape and slope. As a consequence, the average Weber fraction of neurons with preferred numerosity 1 should be much larger compared with those of peaked functions. However, we observed the opposite effect (Fig. 8). The neuronal Weber fraction of 0.80 for neurons tuned to numerosity 1 was even smaller than the average Weber fraction across all tested numerosities (mean, 1.22). Thus, neurons tuned to numerosity 1 were even more selective compared with the tuning functions of cells preferring larger numerosities. This result strongly suggests that such neurons do not act as low-pass summation units, but rather need to be considered as detectors tuned to cardinal value 1.

The same finding held true for neurons preferring numerosity 30. Even neurons with the highest preferred numerosity showed an average Weber fraction of 1.04 that was lower than the mean. This observation again demonstrates steeper slopes for such neurons and not shallower ones, as expected for high-pass filters. In addition, cells discharging maximally for numerosity 30 were hardly more frequent than neurons preferring intermediate numerosities. (Note that a slightly increased frequency of preferred numerosity 30 is even to be expected because few neurons assigned to this class may, in fact, have been tuned to numerosities larger than 30.) If the emergence of peaked functions relied on a presynaptic integration of low- and high-pass response functions, a clear and balanced over-representation of monotonically decreasing and monotonically increasing response functions would be expected. In conclusion, both the selectivity and the frequency of neurons tuned to numerosities 1 and 30 support the notion of a pure labeled-line code subserving explicit numerosity judgments.

Different coding strategies
Our results on a labeled line-based population code contrast with studies of working-memory encoding of somatosensory stimuli (Romo et al., 1999; Brody et al., 2003; Romo and Salinas, 2003) in which purely monotonic response profiles characteristic for summation units were found in the PFC. Behavioral training effects are unlikely to account for the observed differences because the monkeys were highly trained to perform the tasks in both the numerosity and the tactile discrimination projects. The difference in the nature of quantities, however, may play an important role. In our study, numerosity was chosen as a discrete and highly abstract category devoid of sensory particularities, whereas Romo et al. (1999) explored the coding of a continuous and dedicated sensory stimulus (vibrotactile frequency). In addition, the precise behavioral requirements may entail different

coding schemes. Based on network simulations, Verguts (2007) proposed previously that a comparison task (larger than/smaller than) as applied by Romo et al. (1999) may favor summation units, whereas the match-to-sample task (same/different) applied in our numerosity studies could give rise to a labeled-line code. This interesting hypothesis has yet to be tested empirically in the number domain. Yet another modeling study suggests that optimal tuning curves are shaped according to their impact on downstream neural circuits and on behavior (Salinas, 2006). In sum, these contrasting findings emphasize that most likely there is not a single type of working-memory code. The manner in which neurons encode quantity information may depend heavily on the precise task at hand as well as on the stimulus dimension to represent and memorize.

Behavioral relevance of tuned numerosity detectors

Both the behavioral and neural data obeyed Weber's law and were closely related. The monkeys' behavioral precision in numerosity discrimination was found to be superior by a factor of 2.8 compared with their neuronal filters' selectivity. A comparatively better behavioral performance is consistent with the "lower envelope principle," which posits that behavioral detection thresholds are supported by the most sensitive individual neurons of a population, not the population average (Parker and Newsome, 1998).

The analysis of trials in which the monkeys made judgment errors further emphasizes the significance of numerosity tuned neurons for behavioral responses. The spike rates at the preferred numerosity of the neurons were significantly reduced whenever the animals made a wrong decision. In other words, whenever the numerosity detectors did not properly encode the preferred numerosity by maximum discharges, the animals failed. This observation agues for a direct relationship between the neurons peaked numerosity selectivity and the task performance of the neurons.

References

Brannon EM, Terrace HS (1998) Ordering of the numerosities 1 to 9 by monkeys. Science 282:746–749.

Brody CD, Hernandez A, Zainos A, Romo R (2003) Timing and neural encoding of somatosensory parametric working memory in macaque prefrontal cortex. Cereb Cortex 13:1196–2007.

Cantlon JF, Brannon EM (2006) Shared system for ordering small and large numbers in monkeys and humans. Psychol Sci 17:401–406.

Dehaene S (2001) Subtracting pigeons: logarithmic or linear? Psychol Sci 12:244–246.

Dehaene S, Changeux JP (1993) Development of elementary numerical abilities: a neural model. J Cogn Neurosci 5:390–407.

Fechner GT (1860) Elemente der Psychophysik, Vol II. Leipzig: Breitkopf and Härtel.

Feigenson L, Dehaene S, Spelke E (2004) Core systems of number. Trends Cogn Sci 8:307–314.

Hauser MD, MacNeilage P, Ware M (1996) Numerical representations in primates. Proc Natl Acad Sci USA 93:1514–1517.

Hauser MD, Carey S, Hauser LB (2000) Spontaneous number representation in semi-free-ranging rhesus monkeys. Proc R Soc Lond B Biol Sci 267:829–833.

Mandler G, Shebo BJ (1982) Subitizing: an analysis of its component processes. J Exp Psychol Gen 111:1–22.

McCloskey M, Lindemann A (1992) Mathnet: preliminary results from a distributed model of arithmetic fact retrieval. In: The nature and origins of mathematical skill (Campbell JID, ed), pp 365–409. Amsterdam: Elsevier.

Meck WH, Church RM (1983) A mode control model of counting and timing processes. J Exp Psychol Anim Behav Process 9:320–334.

Nieder A (2005) Counting on neurons: the neurobiology of numerical competence. Nat Rev Neurosci 6:177–190.

Nieder A, Miller EK (2003) Coding of cognitive magnitude: compressed scaling of numerical information in the primate prefrontal cortex. Neuron 37:149–157.

Nieder A, Miller EK (2004a) A parieto-frontal network for visual numerical information in the monkey. Proc Natl Acad Sci USA 101:7457–7462.

Nieder A, Miller EK (2004b) Analog numerical representations in rhesus monkeys: evidence for parallel processing. J Cogn Neurosci 16:889–901.

Nieder A, Freedman DJ, Miller EK (2002) Representation of the quantity of visual items in the primate prefrontal cortex. Science 297:1708–1711.

Nieder A, Diester I, Tudusciuc O (2006) Temporal and spatial enumeration processes in the primate parietal cortex. Science 313:1431–1435.

Parker AJ, Newsome WT (1998) Sense and the single neuron: probing the physiology of perception. Annu Rev Neurosci 21:227–277.

Piazza M, Izard V, Pinel P, Le Bihan D, Dehaene S (2004) Tuning curves for approximate numerosity in the human intraparietal sulcus. Neuron 44:547–555.

Piazza M, Pinel P, Le Bihan D, Dehaene S (2007) A magnitude code common to numerosities and number symbols in human intraparietal cortex. Neuron 53:293–305.

Roggeman C, Verguts T, Fias W (2007) Priming reveals differential coding of symbolic and non-symbolic quantities. Cognition, in press.

Romo R, Salinas E (2003) Flutter discrimination: neural codes, perception, memory and decision making. Nat Rev Neurosci 4:203–218.

Romo R, Brody CD, Hernandez A, Lemus L (1999) Neuronal correlates of parametric working memory in the prefrontal cortex. Nature 399:470–473.

Salinas E (2006) How behavioral constraints may determine optimal sensory representations. PloS Biol 4:2383–2392.

Sawamura H, Shima K, Tanji J (2002) Numerical representation for action in the parietal cortex of the monkey. Nature 415:918–922.

Stevens SS (1961) To honor Fechner and repeal his law. Science 133:80–86.

Van Oeffelen MP, Vos PG (1982) A probabilistic model for the discrimination of visual number. Percept Psychophys 32:163–170.

Verguts T (2007) How to compare two quantities? A computational model of flutter discrimination. J Cogn Neurosci 19:409–419.

Verguts T, Fias W (2004) Representation of number in animals and humans: a neural model. J Cogn Neurosci 16:1493–1504.

Weber EH (1850) Der Tastsinn und das Gemeingefühl. In: Handwörterbuch der Physiologie, Vol III, part 2 (Wagner R, ed), pp 481–588. Braunschweig: Vieweg.

Zorzi M, Butterworth B (1999) A computational model of number comparison. In: Proceedings of the Twenty First Annual Conference of the Cognitive Science Society (Hahn M, Stoness SC, eds), pp 778–783. Mahwah, NJ: Erlbaum.

Zorzi M, Stoianov I, Umiltà C (2005) Computational modeling of numerical cognition. In: Handbook of mathematical cognition (Campbell J, ed), pp 67–84. New York: Psychology.