Neurons selective to the number of visual items in the corvid songbird endbrain

Helen M. Ditz and Andreas Nieder[1]

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

It is unknown whether anatomical specializations in the endbrains of different vertebrates determine the neuronal code to represent numerical quantity. Therefore, we recorded single-neuron activity from the endbrain of crows trained to judge the number of items in displays. Many neurons were tuned for numerosities irrespective of the physical appearance of the items, and their activity correlated with performance outcome. Comparison of both behavioral and neuronal representations of numerosity revealed that the data are best described by a logarithmically compressed scaling of numerical information, as postulated by the Weber–Fechner law. The behavioral and neuronal numerosity representations in the crow reflect surprisingly well those found in the primate association cortex. This finding suggests that distantly related vertebrates with independently developed endbrains adopted similar neuronal solutions to process quantity.

single-cell recordings | crow | nidopallium caudolaterale | quantity

Birds show elaborate quantification skills (1–3) that are of adaptive value in naturalistic situations like nest parasitism (4), food caching (5), or communication (6). The neuronal correlates of numerosity representations have only been explored in humans (7–9) and primates (10–18), and they have been found to reside in the prefrontal and posterior parietal neocortices. In contrast to primates, birds lack a six-layered neocortex. The birds' lineage diverged from mammals 300 Mya (19), at a time when the neocortex had not yet developed from the pallium of the endbrain. Instead, birds developed different pallial parts as dominant endbrain structures (20, 21) based on convergent evolution, with the nidopallium caudolaterale (NCL) as a high-level association area (22–26). Where and how numerosity is encoded in vertebrates lacking a neocortex is unknown. Here, we show that neurons in the telencephalic NCL of corvid songbirds respond to numerosity and show a specific code for numerical information.

Results

Crows were trained in a delayed matching-to-sample task to match the number of (one to five) dots presented on touch-sensitive computer displays (Fig. 1 A and B). Crows watched two displays (first sample, then test) separated by a 1-s delay. They were trained to peck at the displays on the screen if the test displays contained the same number of items as the sample. We varied the exact physical appearance of the displays by randomly placing dots in arbitrary locations, and by randomly choosing dot size.

The crows performed the task proficiently (73.8 ± 0.4% and 77.5 ± 0.5% correct over all recording sessions for crow A and crow J, respectively; Fig. 1C). Average performance of both crows was significantly better than chance (50%) for all sample numerosities relative to the numerically most distant nonmatches (Binomial test, $P < 0.01$). Better performance for sample numerosities at the low (one) and high (five) numerosity range (Fig. 1D) are most likely due to "end-effects," because one and five items had to be discriminated only from higher or lower nonmatch stimuli, respectively, whereas sample numerosity three needed to be discriminated from both high and low nonmatches. Similar performance effects are seen in monkeys (11).

To determine whether the crows solved the task by truly abstracting quantity, rather than attending to low-level visual

features, we used different sets of control stimuli targeting the different covarying visual parameters (Fig. 1B and Table 1). Thus, across these stimulus sets, the exact physical appearance of each numerical quantity varied widely. The crows readily generalized to the control stimulus sets; performance was similar across them (Fig. 1D). This result suggests that crows were indeed judging numerosity.

Crows made most errors for quantities that were adjacent to the cued quantity of dots and performed progressively better as numerical distance between two displays increased ("numerical distance effect"). Also, it was harder for the crows to discriminate between two quantities of equal numerical distance as their magnitude increased. For larger sample numerosities, thus, nonmatches had to be numerically more distant to reach a similar performance level as for small sample quantities, resulting in the "numerical magnitude effect" (Fig. 1E).

We investigated the coding scheme by plotting the performance data on different number scales. When plotted on a linear number scale, the shapes of the behavioral performance functions were asymmetric with a steeper slope toward smaller numerosities. However, when the same behavioral discrimination functions were plotted on a logarithmic axis, the shapes were roughly Gaussian, suggesting a logarithmic representation of numerosities. We verified this finding by fitting Gauss functions to the behavioral discrimination functions when plotted on a linear or three nonlinear scales with increasing compression, namely power functions with exponent 0.5 and 0.33, or a logarithmic scale (Fig. 1F). The goodness-of-fit (r^2) values of the Gauss fits, which were taken as a quantitative measure of the tuning curves' symmetry, differed between the four scaling schemes ($P < 0.001$, Friedman test, $n = 81$ sessions). The (nonlinear) power functions and logarithmic scales provided a better fit to the data than the linear scale (Fig. 1G), with better r^2 values the more compressed the scales became up to the logarithmic scale ($P < 0.001$, Wilcoxon test). Moreover, the variance of the distributions for each numerosity (i.e., sigma of the Gauss fit to the performance

Significance

Birds are known for their advanced numerical competence, although a six-layered neocortex that is thought to enable primates with the highest levels of cognition is lacking in birds. We recorded neuronal activity from an endbrain association area termed nidopallium caudolaterale (NCL) in crows that discriminated the number of items in displays. NCL neurons were tuned to preferred numerosities. Neuronal discharges were relevant for the crows' correct performance. Both the neuronal and the behavioral tuning functions were best described on a logarithmic number line, just as predicted by the psychophysical Weber–Fecher Law. Our data suggests that this way of coding numerical information has evolved based on convergent evolution as a superior solution to a common computational problem.
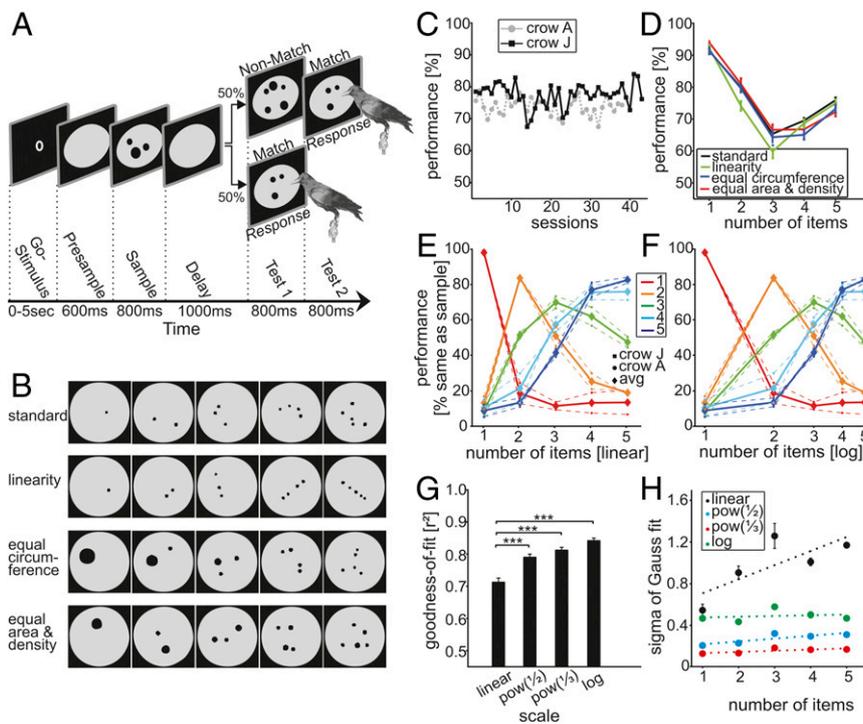
NEUROSCIENCE

Fig. 1. Task protocol, stimuli, and behavioral performance. (*A*) DMS task. The crow initiated a trial by keeping its head still in front of the monitor (automatically detected) to activate a go stimulus. After a 600-ms presample period, a sample stimulus was presented for 800 ms, followed by a 1,000-ms delay. The crow had to peck the test1 display if it contained the same number of items as the sample and had to refrain from pecking if it did not (probability = 0.5). All numerosities were used as nonmatch stimuli (probability = 0.125) for the respective sample numerosities. (*B*) A small subset of the stimulus displays are shown as examples. The physical appearance of the displays varied widely for the same quantities (see details in Table 1). (*C*) Behavioral performance (percent correct) for both crows in the DMS task over all recording sessions (chance level = 50%). (*D*) Average performance (error bars represent SEM) of both crows to the standard and the control conditions during the recording sessions. (*E* and *F*) The behavioral performance functions (both crows and average) indicated whether they judged the first test stimulus (after the delay) as containing the same number of items as the sample display ("% same as sample"). Colors represent performance curves for a given sample numerosity. Behavioral performance functions are plotted on a linear (*E*) and logarithmic number scale (*F*); the latter resulted in more symmetric functions. (*G* and *H*) Quantification of Gauss fits to the behavioral data. (*G*) Goodness-of-fit of Gauss 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). (*H*) The SD (sigma) of the Gauss fits for nonlinear scaling plotted against the center of the Gauss function (which is identical to the numerosity of the match stimulus). Dotted lines indicate linear fits (error bars ± SEM). (The values of sigma are related to the specific compression scheme.)

curves) was constant when the data were plotted on a power function scale with 0.33 exponent (slope of linear fit = 0.135) and the logarithmic scale (slope of linear fit = 0.007) (Fig. 1*H*), which is predicted by a nonlinear coding model of numerosity (27–29). Thus, performance data for numerosity judgments is better described by using a power function compressed or above all a logarithmically compressed scale, as opposed to a linear scale. The data follow the Weber–Fechner law ($S = k \times \log(I)$), which states that linear increments in sensation S are proportional to the logarithm of stimulus magnitude I. This postulated logarithmic compression is present in the data because the behavioral representations (i.e., the performance curves) are better described by a symmetric Gaussian distribution on a log scale.

We recorded from 499 randomly selected neurons from the NCL (Fig. 2*A*) of two crows while they performed the numerosity

discrimination task. Of these neurons, 20% (98/499) modulated their discharges as a function of the numerosity (from 1 to 5) during sample presentation. This selectivity was found irrespective of the exact appearance of the multiple-dot patterns (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 ($P < 0.01$).

Five such neurons that generalized across changes in the physical appearance of the sample displays are shown in Fig. 2 *B–F*. The example neuron shown in Fig. 2*B* was tuned to "one" and showed remarkably similar activity to the standard versus linear stimuli. Other neurons were tuned to "two" (Fig. 2*C*) or "five" (Fig. 2*F*), respectively, and responded equally well to standard dots and those that equated the total circumference, whereas example

Table 1. Stimulus protocols

Stimulus type protocol	Spatial arrangement	Surface area	Circumference	Density
Standard	Randomized†	Increasing with quantity	Increasing with quantity	Increasing with quantity
Equal area and equal density	Randomized†	Equal across quantity	Increasing with quantity	Equal across quantity‡
Equal circumference	Randomized†	Decreasing with quantity	Equal across quantity	Increasing with quantity
Linear	One-dimensional, linear	Increasing with quantity	Increasing with quantity	Increasing with quantity

†Three dots tend to be arranged as triangle, four dots as quadrangle, five dots as pentagon.
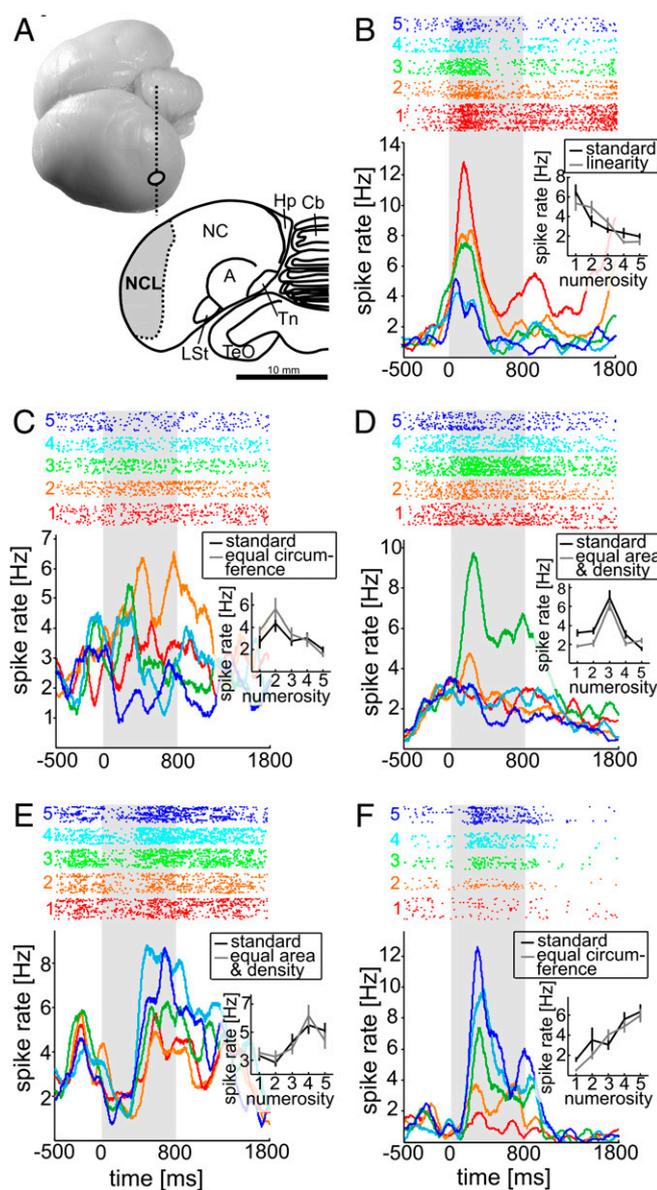‡Density determined by calculating the average distance between all dots in a given display.

density across numerosities with comparable firing rates. Each neuron showed peak activity for one of the visual quantities and a systematic dropoff of activity as the number of sample items varied from the preferred value. The quantities "one" and "five" were preferred by the neurons, but neural preference was distributed across all five numerosities (numerosity 1: 29%; 2: 13%; 3: 8%; 4: 19%; 5: 31%). Few cells (not included in the group of numerosity-selective neurons) were responsive to both numerosity and stimulus type (8% or 39/499; two-way ANOVA, effect of numerosity and stimulus protocol or interaction between stimulus protocol and numerosity, $P < 0.01$). Thus, the numerosity of sample items was the dominant factor encoded by these neurons, and not the physical appearance of the displays.

Neural activity in the NCL seemed to underlie a systematic, orderly representation of numerosity; neurons showed peak activity to a specific number of items and a progressive dropoff as the numerosity progressively varied (e.g., Fig. 2 *B–F*). This activity was evaluated across the population. Population neural tuning functions were calculated by averaging the normalized activity for all neurons that preferred a given quantity. Neural activity formed band-pass filters with increasingly attenuated activity as distance from the preferred quantity increased (Fig. 3*A*). The neuronal data mirrored the behavioral distance and magnitude effects by the fact that the neural filters were also peak functions that became less selective (wider) with increasing preferred numerosity.

Much like the behavioral data (Fig. 1*E*), the neural filter functions were asymmetric when plotted on a linear scale (Fig. 3*A*), but more symmetric when plotted on a logarithmic scale (Fig. 3*B*). We applied the same goodness-of-fit tests that were applied to the behavioral data. Once again, the four different scaling schemes resulted in significantly different goodness-of-fit values across all numerosity functions ($P < 0.01$, Friedman test, $n = 98$) (Fig. 3*C*). The (nonlinear) power functions and logarithmic scale provided a

Fig. 2. Recording site and neuronal responses. (*A, Top*) Dorso-lateral view of a carrion crow brain (ellipse indicates electrode penetration site). (*A, Bottom*) Coronal section (indicated by dashed vertical line in *Top*) through the brain of a carrion crow illustrating the borders of the NCL in the caudal telencephalon based on immunohistochemistry for tyrosine-hydroxylase. A, Arcopallium; Cb, Cerebellum; Hp, Hippocampal formation; LSt, striatum laterale; NC, Nidopallium caudale; NCL, Nidopallium caudolaterale; TeO, Tectum opticum; Tn, Nucleus taeniae amygdalae. (*B–F*) Responses of five example NCL neuron selective to numerosity 1 (*B*), 2 (*C*), 3 (*D*), 4 (*E*), and 5 (*F*). *Top* shows dot-raster histograms (each dot represents an action potential); *Bottom* depicts averaged spike density functions (activity averaged and smoothed by a 150-ms Gauss kernel). Each colored line shows the time course of activity for the five tested quantities. The spike density functions represent the rate of action potentials, i.e., the number of action potentials per time interval divided by the number of trials, thus accounting for slightly different trial numbers. The first 500 ms represent the presample period (baseline). Gray shading represents the analyzed sample period (800 ms). Colors of dot histogram and spike density functions correspond to the number of items in the sample displays. The tuning function insets indicate the mean activity of the neurons to each of the two stimulus protocols (error bars represent SEM) in the sample period.

neurons tuned to "three" (Fig. 2*D*) or "four" (Fig. 2*E*), respectively, encoded standard dots and those equating both total area and

Fig. 3. Neuronal representation of numerosities in the NCL during the sample period. (*A* and *B*) Normalized tuning functions averaged for neurons preferring the same numerosity (indicated by same color) when plotted on a linear number scale (*A*) or on a logarithmic number scale (*B*). (*C*) Goodness-of-fit for the four different scaling schemes. (*D*) SD values for the scaling schemes across preferred numerosities (error bars ± SEM).

NEUROSCIENCE

better fit to the data than the linear scale ($P < 0.05$, Wilcoxon test, $n = 98$). The mean goodness-of-fit values for the linear scale, the power function with exponent of 0.5, the power function with exponent of 0.33, and the logarithmic scale were 0.75, 0.77, 0.78, and 0.78. Also similar to the behavioral data, and as predicted by a nonlinear coding model (27–29), the variance of neural distributions was more or less constant with increasing preferred numerosity when the data were plotted on a logarithmic scale (slope of linear fit = −0.022), but increased with numerosity when the data were plotted on a linear scale (slope = 0.068) (Fig. 3D). In terms of the scaling scheme, the neural data mirrored the behavioral findings.

Further evidence that NCL neural activity was linked with behavior came from an examination of error trials. When crows made judgment errors, neural activity for the preferred quantity was significantly reduced, as can be seen for two numerosity-selective example neurons (Fig. 4 A and B). This effect was significant at the population level and resulted in a reduction of discharge rate to 71% than observed on correct trials (100%) ($P < 0.05$, Wilcoxon test, $n = 81$). As a result of this reduced activity (and the orderly representation of numerosity), on error trials the activity elicited by a sample of a given numerosity was more similar to that elicited by adjacent numerosities on correct trials (Fig. 4C). This decrease in activity may reflect the sample being mistakenly encoded as an adjacent numerosity.

Discussion

These results indicate that neurons in the corvid NCL can participate in high-level, abstract visual representations that are contributive to judgments of numerical quantity. Our behavioral and neuronal data show an impressive correspondence of neuronal mechanisms found in the avian brain with those reported earlier in the nonhuman and human primate brain: First, NCL neurons were tuned to individual preferred numerosities characteristic of a "labeled-line code," enabling an unequivocal representation of numerosity by a neuronal population. Complementary findings have been made in the monkey prefrontal and posterior parietal cortex; both in the highly trained (10–18) and in the numerically naïve monkey (30), numerosity is encoded by a labeled-line code. Second, neuronal discharges of selective neurons proved to be relevant for the crows' correct performance; if the neurons did not properly encode their preferred numerosity, the crow was prone to make mistakes. Similar results have been reported repeatedly for numerosity-selective neurons in monkeys (10–18). Third, both the neuronal and the behavioral tuning functions were best described on a logarithmic number line, arguing for a nonlinearly compressed coding of numerical information. Because logarithmic coding is postulated by the psychophysical Weber–Fechner Law and has been established for human brain signals (8, 9) and primate neurons (28), this finding suggests that abstract numerical and sensory representations share the same fundamental mechanisms and neural coding schemes both in songbirds and primates.

Because the labeled-line code is found not only in the primate neocortex, but also in the avian endbrain, this code may be computationally superior compared with alternative neuronal representations such as summation coding in which numerosity is encoded via monotonic response functions of neurons (27). Neurophysiological constraints may therefore have favored a labeled-line code for numerical information across vertebrate species. Thus, this code has been implemented (at least) twice during the course of evolution, irrespective of the precise origin and anatomical structures found in intelligent vertebrate brains, based on convergent evolution (31). Recently, information-processing principles that define the canonical cortical microcircuit in the mammalian neocortex have been described in the avian auditory pallium (32). This finding could indicate that canonical endbrain microcircuits evolved in a common ancestor of mammals and birds. Perhaps this result might also provide a physiological explanation for the evolution of neuronal computations that give rise to numerical competence in both vertebrate groups. More comparative approaches in neuroscience will therefore be indispensable for deciphering these evolutionary stable neuronal mechanisms (33).

Materials & Methods

Subjects. We used two hand-raised crows (*Corvus corone corone*), one male and one female, in these experiments. The crows were trained on a delayed match-to-sample (DMS) task with the number of items in dot displays as discriminative stimuli. All crows were obtained from the institute's breeding facilities. The birds were housed in social groups in spacious indoor aviaries (34). They were maintained on a controlled feeding protocol during the sessions and earned food during and after the daily tests. All animal preparations and procedures fully complied with the NIH *Guide for Care and Use of Laboratory Animals* (35) and were approved by the local ethical committee and authorized by the national authority (Regierungspraesidium).

Apparatus. The crows were attached to a wooden perch by a leather jess and placed in an operant conditioning chamber in front of a touchscreen monitor (3M Microtouch; 15 inches', 60 Hz refresh rate). All stimuli were displayed on this monitor. Reward [birdseed pellets or mealworms (*Tenebrio molitor* larvae)] was delivered by a custom-built automated feeder below the screen. The CORTEX program (National Institute of Mental Health) was used for experimental control and behavioral data acquisition. An infrared light barrier, in combination with a reflector attached to the bird's head, registered when the bird was positioned in front and facing the screen.

Behavioral Protocol. In the DMS task (Fig. 1A), the crows initiated a trial by moving their head into the light barrier when a go stimulus ("O"; 3×4 mm) was shown on the screen. The crows had to keep their head still throughout the trial; if they moved their head before the response period (as detected by the light barrier), the trial was aborted. As soon as the crow kept its head at the defined location, the go stimulus turned off, followed by a 600-ms presample with a gray background circle on the screen. The sample stimulus was presented in the center of the screen for 800 ms and was pseudo-randomly chosen from a set of 80 different images (16 different dot images for each of the numerosities from 1 to 5). For each daily session, all images were generated anew with random dot layouts by using Matlab routines and exchanged every day. The screen showed a gray background circle during
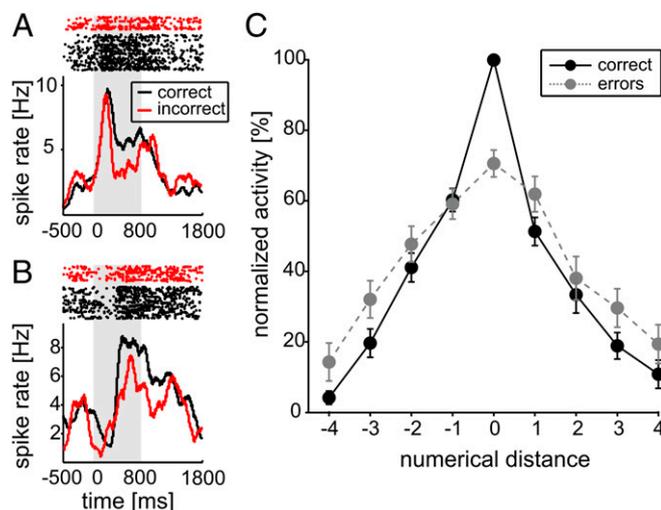
Fig. 4. Behavioral relevance of selective neurons. (*A*) Responses of a "three"-neuron (same as in Fig. 2D) to its preferred numerosity during correct and error trials. *Top* shows dot-raster histograms; *Bottom* depicts averaged spike density functions. (*B*) Responses of a "four"-neuron (same as in Fig. 2E) to its preferred numerosity during correct and error trials. Layout as in A. (*C*) Normalized average tuning function across all preferred numerosities and selective neurons. Functions for correct (solid lines) and error trials (dotted lines) are shown. Error bars indicate SEs across cells.

the subsequent 1,000-ms delay, in which the bird had to remember the sample numerosity to solve the task. All analyses focus on these sample periods.

After the delay, a test1 stimulus appeared (800 ms). In 50% of the cases, the test1 was a match, i.e., it showed the same number of dots (but always in a different layout than the sample numerosity). The birds indicated a match by pecking the dot display. If their choice was correct, the automated feeder delivered feedback via light, and a reward sound was played. In the other 50% of the trials, test1 displayed a nonmatch, i.e., more or less dots than the sample display; in this case, the crow was not allowed to peck but had to wait until a second, test2 stimulus was shown. The test2 stimulus was always a match and had to be pecked to receive a reward.

The crows were rewarded with food for each correct trial. If the birds chose incorrectly, the trial was aborted and a short timeout (3 s) was presented before the start of the next trial. If no response occurred within 1,600 ms, the trial was dismissed. All relevant task parameters (match/nonmatch, numerosity, standard vs. controls) were balanced.

Stimuli. Numerosity stimuli consisting of multiple-dot patterns were generated by 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. Small black filled dots (diameter of 2.8°–0.4° visual angle) appeared on a gray background of a large circular area with a diameter of 10° 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 crows from memorizing the visual patterns of the displays, each quantity was tested with many different images per session that were newly generated for each session. The sample and test displays that appeared on each trial were never identical. The standard stimuli consisted of dots of pseudorandom size that were pseudorandomly located on the background circle to form shape-like arrangements (triangle, quadrangle, pentagon). To ensure that the numerosity-discrimination task was solved by judging the discrete quantity, low-level visual features were excluded by using control stimuli in addition to standard stimuli (Table 1). Three sets of control stimuli were alternatively used in each session (Fig. 1B): area and density control (total area of all items and mean density of dot patterns in a display equated for all stimuli in a trial), circumference control (total circumference of all items equated for all stimuli in a trial), and linear control (two or more dots formed a line).

Surgery and Recordings. All surgeries were performed under sterile conditions while the animals were under general anesthesia. Crows were anesthetized with a ketamine/rompun mixture (50 mg ketamine, 5 mg/kg xylazine initially and supplemented by hourly 17 mg of ketamine, 1.7 mg/kg xylazine i.m.). After the surgery, the crows received analgesics [Butorphanol (Morphasol), 1 mg/kg i.m.]. The head was placed in the stereotaxic holder that was customized for crows with the anterior fixation point (i.e., beak bar position) 45° below the horizontal axis of the instrument (36). Using stereotaxic coordinates (center of craniotomy: anterior-posterior 5 mm; medial-lateral 13 mm), we chronically implanted two to four microdrives with four electrodes each, a connector for the headstage, and a small headpost to hold the reflector for the light barrier.

We recorded 8 to 16 chronically implanted microelectrodes on two to four custom-built microdrives. We used glass-coated tungsten microelectrodes with 2 MΩ impedance (Alpha Omega LTD). The electrodes targeted the NCL. Tracing electrode tracks of an identically implanted crow used for a different study (25, 26) confirmed that recording locations were within NCL. Cryostat sections were immunohistochemically stained for tyrosine-hydroxylase to identify dopaminergic cells, which characterize the NCL (25). Both crows used in this study are still alive and are participating in related experiments.

At the start of each recording session, the electrodes were advanced manually until a good neuronal signal was detected on at least one of the channels of each microdrive. Neurons were not preselected for involvement in the task. Each microdrive had a range of ~6 mm, which was exploited to record from the NCL across different depths over a period of several weeks. Signal amplification, filtering, and digitizing of spike waveforms was accomplished by using the Plexon system.

For each recording session, the birds were placed in the recording setup, and a headstage containing an amplifier was plugged into the connector implanted on the bird's head and connected to a second amplifier/filter and the Plexon MAP box outside the setup by a cable above and behind the bird's head (all components by Plexon). Spike sorting into single-unit waveforms was performed manually offline by using the Plexon system. The analysis includes all neurons which were recorded for at least four repetitions of each sample numerosity per protocol type (average repetition number was 33) and had a firing rate of at least 0.5 Hz during all periods. Each recording session lasted between 361 and 720 correct trials in ~2 h.

Data Analysis. Neuronal activity during the task was analyzed in the sample period. Neuronal response rates were measured in a 800-ms window, shifted by 100 ms from sample onset to account for the visual latency of most neurons. To determine numerosity selectivity of the neurons, a two-factorial analysis of variance (ANOVA) was performed with numerosity (1–5) and stimulation condition (standard or control) as factors. For each recording session, we used the highly variable standard protocol and one of the control protocols, which was alternated on a daily basis. Applying only two stimulus protocols was necessary to gain a viable number of repetitions of one and the same trial condition for proper statistical analysis before the crow was satisfied and stopped working. Irrespective of the control protocols, only few neurons responded to nonnumerical visual features of the stimulus displays, confirming that sensory parameters cannot account for the prominent numerosity effects. 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 "preferred numerosity" of a given cell.

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. Gauss functions were fit to the daily performance functions of each crow and to the neural filter functions of each numerosity-selective neuron. 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 the goodness-of-fit (r^2), and, therefore, the better that scale describes the data. These nonlinearly compressed scaling schemes were chosen because Stevens' power law ($S = k \times In$) postulates that sensation S is a power function of the stimulus magnitude I, whereas Fechner's Law ($S = k \times \log(I)$) proposes a logarithmic relationship.

To evaluate the behavioral significance of numerosity-selective neurons, discharges in correct and error trials were compared. Of all purely numerosity-selective neurons ($n = 98$), neurons with at least three error trials for their preferred numerosity ($n = 81$) were included in analyses of error trials. Discharge rates of single neurons to the preferred numerosity were compared in correct versus error trials (Wilcoxon test).

1. Scarf D, Hayne H, Colombo M (2011) Pigeons on par with primates in numerical competence. *Science* 334(6063):1664.
2. Bogale BA, Kamata N, Mioko K, Sugita S (2011) Quantity discrimination in jungle crows, Corvus macrorhynchos. *Anim Behav* 82(4):635–641.
3. Moll FW, Nieder A (2014) The long and the short of it: Rule-based relative length discrimination in carrion crows, Corvus corone. *Behav Processes* 107:142–149.
4. Lyon BE (2003) Egg recognition and counting reduce costs of avian conspecific brood parasitism. *Nature* 422(6931):495–499.
5. Hunt S, Low J, Burns KC (2008) Adaptive numerical competency in a food-hoarding songbird. *Proc Biol Sci* 275(1649):2373–2379.
6. Templeton CN, Greene E, Davis K (2005) Allometry of alarm calls: Black-capped chickadees encode information about predator size. *Science* 308(5730):1934–1937.
7. Dehaene S, Spelke E, Pinel P, Stanescu R, Tsivkin S (1999) Sources of mathematical thinking: Behavioral and brain-imaging evidence. *Science* 284(5416):970–974.
8. Piazza M, Izard V, Pinel P, Le Bihan D, Dehaene S (2004) Tuning curves for approximate numerosity in the human intraparietal sulcus. *Neuron* 44(3):547–555.
9. Jacob SN, Nieder A (2009) Tuning to non-symbolic proportions in the human frontoparietal cortex. *Eur J Neurosci* 30(7):1432–1442.
10. Sawamura H, Shima K, Tanji J (2002) Numerical representation for action in the parietal cortex of the monkey. *Nature* 415(6874):918–922.
11. Nieder A, Freedman DJ, Miller EK (2002) Representation of the quantity of visual items in the primate prefrontal cortex. *Science* 297(5587):1708–1711.
12. Nieder A, Diester I, Tudusciuc O (2006) Temporal and spatial enumeration processes in the primate parietal cortex. *Science* 313(5792):1431–1435.

NEUROSCIENCE

13. Sawamura H, Shima K, Tanji J (2010) Deficits in action selection based on numerical information after inactivation of the posterior parietal cortex in monkeys. *J Neurophysiol* 104(2):902–910.
14. Nieder A (2012) Supramodal numerosity selectivity of neurons in primate prefrontal and posterior parietal cortices. *Proc Natl Acad Sci USA* 109(29):11860–11865.
15. Vallentin D, Nieder A (2008) Behavioral and prefrontal representation of spatial proportions in the monkey. *Curr Biol* 18(18):1420–1425.
16. Nieder A (2013) Coding of abstract quantity by 'number neurons' of the primate brain. *J Comp Physiol A Neuroethol Sens Neural Behav Physiol* 199(1):1–16.
17. Jacob SN, Nieder A (2014) Complementary roles for primate frontal and parietal cortex in guarding working memory from distractor stimuli. *Neuron* 83(1):226–237.
18. Viswanathan P, Nieder A (2015) Differential impact of behavioral relevance on quantity coding in primate frontal and parietal neurons. *Curr Biol* 25(10):1259–1269.
19. Evans SE (2000) *Evolutionary Developmental Biology of the Cerebral Cortex*, eds Bock G, Cardew G (John Wiley & Sons, Chichester, UK), pp 109–113.
20. Jarvis ED, et al.; Avian Brain Nomenclature Consortium (2005) Avian brains and a new understanding of vertebrate brain evolution. *Nat Rev Neurosci* 6(2):151–159.
21. Butler AB, Reiner A, Karten HJ (2011) Evolution of the amniote pallium and the origins of mammalian neocortex. *Ann N Y Acad Sci* 1225:14–27.
22. Divac I, Mogensen J, Björklund A (1985) The prefrontal 'cortex' in the pigeon. Biochemical evidence. *Brain Res* 332(2):365–368.
23. Güntürkün O (2005) The avian 'prefrontal cortex' and cognition. *Curr Opin Neurobiol* 15(6):686–693.
24. Rose J, Colombo M (2005) Neural correlates of executive control in the avian brain. *PLoS Biol* 3(6):e190.
25. Veit L, Nieder A (2013) Abstract rule neurons in the endbrain support intelligent behaviour in corvid songbirds. *Nat Commun* 4:2878.
26. Veit L, Hartmann K, Nieder A (2014) Neuronal correlates of visual working memory in the corvid endbrain. *J Neurosci* 34(23):7778–7786.
27. Dehaene S, Changeux JP (1993) Development of elementary numerical abilities: A neuronal model. *J Cogn Neurosci* 5(4):390–407.
28. Nieder A, Miller EK (2003) Coding of cognitive magnitude: Compressed scaling of numerical information in the primate prefrontal cortex. *Neuron* 37(1):149–157.
29. Merten K, Nieder A (2009) Compressed scaling of abstract numerosity representations in adult humans and monkeys. *J Cogn Neurosci* 21(2):333–346.
30. Viswanathan P, Nieder A (2013) Neuronal correlates of a visual "sense of number" in primate parietal and prefrontal cortices. *Proc Natl Acad Sci USA* 110(27):11187–11192.
31. Emery NJ, Clayton NS (2004) The mentality of crows: Convergent evolution of intelligence in corvids and apes. *Science* 306(5703):1903–1907.
32. Calabrese A, Woolley SM (2015) Coding principles of the canonical cortical microcircuit in the avian brain. *Proc Natl Acad Sci USA* 112(11):3517–3522.
33. Bullock TH (1984) Comparative neuroscience holds promise for quiet revolutions. *Science* 225(4661):473–478.
34. Hoffmann A, Rüttler V, Nieder A (2011) Ontogeny of object permanence and object tracking in the carrion crow, Corvus corone. *Anim Behav* 82:359–367.
35. Committee on Care and Use of Laboratory Animals (1996) *Guide for the Care and Use of Laboratory Animals* (Natl Inst Health, Bethesda), DHHS Publ No (NIH) 85-23.
36. Karten HJ, Hodos W (1967) *A Stereotaxic Atlas of the Brain of the Pigeon: (Columba Livia)* (Johns Hopkins Press, Baltimore).