Review

Author for correspondence:
Andreas Nieder
e-mail: andreas.nieder@uni-tuebingen.de

Evolution of cognitive and neural solutions enabling numerosity judgements: lessons from primates and corvids

Andreas Nieder

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

AN, 0000-0001-6381-0375

Brains that are capable of representing numerosity, the number of items in a set, have arisen repeatedly and independently in different animal taxa. This review compares the cognitive and physiological mechanisms found in a non-human primate, the rhesus macaque, and a corvid songbird, the carrion crow, in order to elucidate the evolutionary adaptations underlying numerical competence. Monkeys and corvids are known for their advanced cognitive competence, despite them both having independently and distinctly evolved endbrains that resulted from a long history of parallel evolution. In both species, numerosity is represented as an analogue magnitude by an approximate number system that obeys the Weber–Fechner Law. In addition, the activity of numerosity-selective neurons in the fronto-parietal association cortex of monkeys and the telencephalic associative area nidopallium caudolaterale of crows mirrors the animals' performance. In both species' brains, neuronal activity is tuned to a preferred numerosity, encodes the numerical value in an approximate fashion, and is best represented on a logarithmic scale. Collectively, the data show an impressive correspondence of the cognitive and neuronal mechanisms for numerosity representations across monkeys and crows. This suggests that remotely related vertebrates with distinctly developed endbrains adopted similar physiological solutions to common computational problems in numerosity processing.

This article is part of a discussion meeting issue 'The origins of numerical abilities'.

1. Introduction

(a) Evolution of numerical competence

The ability to judge the number of items in a set, its numerosity, is an essential aspect of numerical cognition. Different levels of numerical competence are widespread throughout the animal kingdom and have been demonstrated in such diverse taxa as insects [1,2], fishes [3–6], amphibians [7,8], birds [9–14] and mammals [15–17]. This may be surprising, given that some of these animal groups—such as insects and vertebrates—share a last common ancestor several hundreds of million years ago. The respective lineages have evolved in parallel over time, thereby acquiring very differently organized nervous systems. Despite different neural substrates giving rise to cognitive capabilities, rudimentary numerical capacities seem to be ubiquitous in advanced animals.

This points to a selective pressure for the evolution of numerical competence. This selective pressure could be explained if dealing with numerical information were to provide a survival advantage. Indeed, several studies examining animals in their ecological environments suggest that representing number enhances an animal's ability to reproduce [18,19], navigate [1], exploit food sources [7,20], hunt prey [21], avoid predation [22] and persist in social interactions [23–25].
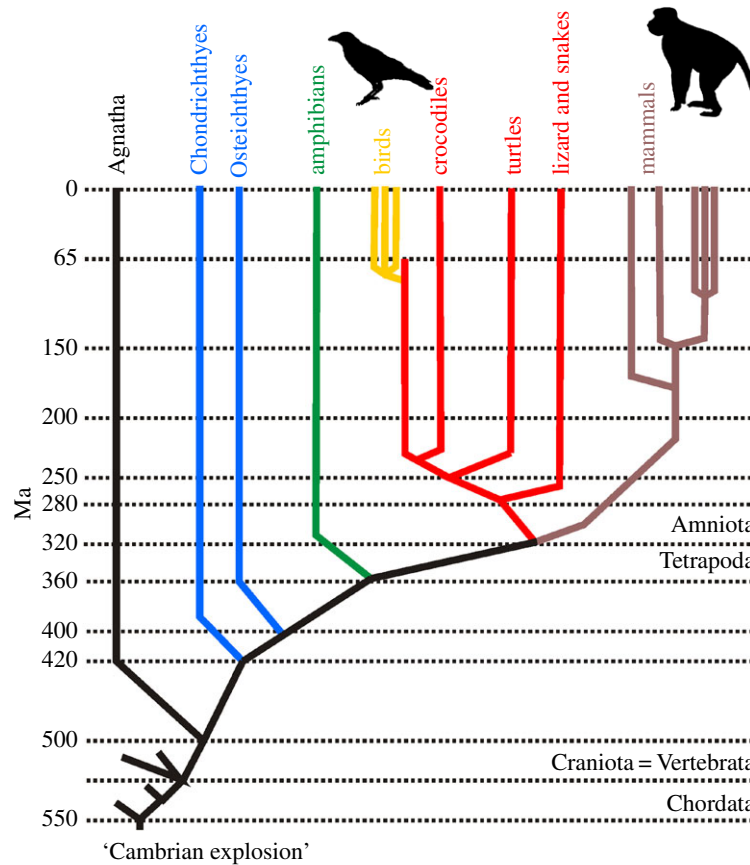
Figure 1. Simplified phylogenetic tree of the vertebrates. The traditional (but taxonomically incorrect) vertebrate groups are displayed: the fishes (*blue*) represented by the Chondrichthyes (cartilaginous fishes) and Osteichthyes (bony fish), the amphibians (*green*), the birds (*yellow*), the reptiles (*orange*) and the mammals (*brown*). These are all jawed vertebrates that are distinct from the jawless Agnatha (such as lampreys). From a taxonomic point of view, birds belong to the monophyletic class Reptilia. All displayed animal groups are vertebrates (also called 'Craniata'), which are characterized by a backbone. Vertebrates belong to the phylum Chordata, one of the major animal groups that evolved around the time of the so-called 'Cambrian explosion' about 550 million years ago. At that time also the largest animal phylum, the Arthropoda, containing also insects, diverged (not drawn). (Online version in colour.)

These examples illustrate that the ability to represent numerical values seems to provide a measurable benefit for different animal species. Numerical competence is of adaptive value, as it helps animals to survive and pass on their genes to the next generation. The process by which numerosity is extracted from sensory input and represented in cognition and brain is therefore of major interest in biology.

(b) Convergent evolution of intelligence in primates and corvids

The groups of primates and songbirds contain some of the most cognitively advanced species. Old World monkeys such as macaques (genus *Macaca*), and corvids such as crows (genus *Corvus*), are renowned for their superior cognitive flexibility. As mammals and primates, rhesus macaques are relatively closely related to humans with which they share a last common ancestor some 25 million years ago [26,27] (figure 1). Macaque monkeys grasp abstract categories and concepts [28], exhibit elaborate executive functions [29], interpret others' perceptions [30] and display complex social behaviours [31,32]. Wild macaques spontaneously discriminate the quantity of food items [33] and represent the numerosities 1–9 on an ordinal scale in controlled laboratory experiments [16].

Despite their distant relationship, corvids (ravens, crows and jays) are of similar intelligence as primates. The last common ancestor between mammals and birds, a cognitively humble reptilian-like stem amniote, lived about 320 million years ago [26,27,34]. Since then, primates and corvids have evolved in parallel (figure 1). Still, corvids exhibit superior object permanence [35,36], are able to rapidly extract general principles to guide behaviour [37,38], show a flexible capacity to remember the past and plan for the future [39], and take the states of conspecifics during social interactions into account [40,41]. Corvids have been shown to spontaneously discriminate the relative quantity of food items [42–44], and they can be trained to judge absolute numerical values [45–47]. Collectively, cognitive capabilities of corvids are in many respects on a par with those of many nonhuman primates.

(c) Evolution of mammalian and avian endbrains

In order to process abstract categories such as numerosity, high-level brain areas are required that can integrate sensory information across time and space and from many senses before planning motor commands. In primates, sophisticated circuitries in the six-layered neocortex of the parietal and frontal lobe fulfil these requirements. These classical association areas receive highly processed information from all sensory modalities and project to premotor structures. In the parietal and frontal association cortices, complementary brain research in humans and nonhuman primates identified a dedicated brain network for processing numerical information [48]
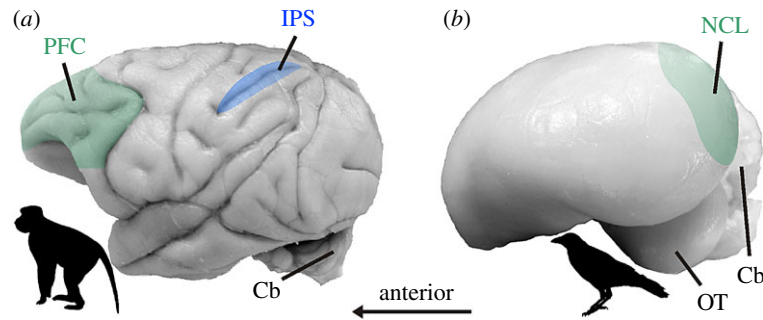
Figure 2. The brains of macaques and crows. (*a*) Lateral view of a macaque brain highlighting the prefrontal cortex (PFC, *green*) and the intraparietal sulcus (IPS, *blue*) on the surface of the cerebral cortex (neocortex). The cerebral cortex covers almost the entire brain. (*b*) Lateral view of a crow brain with the nidopallium caudolaterale (NCL) located inside the telencephalon colour coded. Cb, cerebellum; OT, optic tectum. (Online version in colour.)

(figure 2*a*). The intraparietal sulcus (IPS) and the dorsolateral prefrontal cortex (PFC) have been identified as key nodes of this network. Physiological parameters suggest that number information is processed hierarchically between the IPS and the PFC, with the IPS as the first cortical hub to extract quantitative information, and the PFC as a putative recipient of numerical information [49–52]. These reciprocally connected association areas hosting the number network not only are suited to extract abstract numerosities from sensory input, but also are particularly well suited to maintain information across time and to exert cognitive control [53–55]. Persistent activity is typically witnessed during delayed response tasks that include a short gap in time between a sensory stimulus and an instructed response [56,57]. This persistent activity is thought to be a neuronal correlate of working memory, which is the ability to briefly retain and manipulate information in mind. Through sustaining activity, neurons actively buffer and process information to bridge the gap until an adaptive output is selected. High-level numerical functions would be impossible without persistent activity.

Corvids, just as any other bird, have an independently evolved endbrain design affording twice as many neurons as a primate brain of equal mass [58]. During the more than 600 million years of independent evolution since the separation from the last common ancestor 320 million years ago [26,27,34], birds elaborated different parts of the endbrain (telencephalon) as highest integration centres. In both mammals and birds, the bulk of the endbrain stems from the ontogenetic mantle, the pallium, and thus shares common ancestry (*homology*) [59,60]. However, the overall architecture of the endbrain is very different and has developed independently to give rise to similar functions, probably owing to similar selection pressures (*homoplasy*, convergent evolution) [61,62].

As a consequence, a six-layered neocortex that endows primates with the highest levels of cognition is absent in the endbrain of birds and all other non-mammalian vertebrates. Instead of layers, the avian endbrain consists of nuclear organized circuits. They originate from a different part of the telencephalic pallium which does not give rise to the neocortex in mammals [61,63–65]. The so-called dorsal ventricular ridge (DVR) is one of the main components of the sauropsid (i.e. reptilian and avian) pallium that gives rise to associative telencephalic structures. A particularly integrative region originating from the DVR is the nidopallium, which also contains the nidopallium caudolaterale (NCL) (figure 2*b*). The NCL is a high-level cognitive structure in birds and considered to be a functional equivalent of the PFC [66,67]. In the same way

as the PFC, the NCL integrates highly processed sensory information for all modalities and projects to premotor structures, is modulated by dopamine, and interacts with limbic, visceral and memory-related structures. Recent single-unit recordings in behaving crows confirm the resemblance of the NCL with the PFC by showing that NCL neurons encode sensory and cognitive variables during working memory, but also participate in the translation of cognitive signals to motor behaviours [68,69]. As we shall see later, the NCL also plays an important role in numerical cognition.

Despite these functional similarities, the evolution of pallial structures and the putative homologies and homoplasies between birds and mammals are still highly debated and not fully understood [62,70–73]. For example, the nidopallium including the NCL is sometimes considered to be more related to the amygdaloid–claustrum complex [74], and the missing connections between NCL and the hippocampus emphasize structural differences compared with the mammalian PFC [75]. In addition, other avian telencephalic brain regions, such as the hyperpallium and mesopallium, could participate in numerical cognition.

Irrespective of these unresolved anatomical issues, primates and corvids provide a most interesting case of convergent evolution of numerical competence because of the independent evolution of the associative endbrain areas in both animal groups. How is numerosity represented in the brains of monkeys and crows? Could there be a common, superior code for numerosity evolved independently in the two animal groups, or do the solutions found during evolution differ? When species evolve in parallel, the neurophysiological solutions to a common behavioural problem may take different paths. Sound localization in birds and mammals provides an illustrative example for physiological differences resulting from parallel evolution [76]. The mechanisms of how location is computed from interaural time differences of a sound in homologous auditory brainstem nuclei are fundamentally different in birds and mammals. It is obvious that during evolution different solutions to the same problem—sound localization—have been implemented.

Could a similar scenario have happened for numerical representations? Our laboratory has conducted behavioural and neurophysiological experiments to establish how numerical quantity is represented in differently evolved endbrains. We trained monkeys and crows on the same numerosity tasks and under very similar conditions. The current review compares the respective findings and discusses the implications of this work for the evolution of abstract numerical representations.
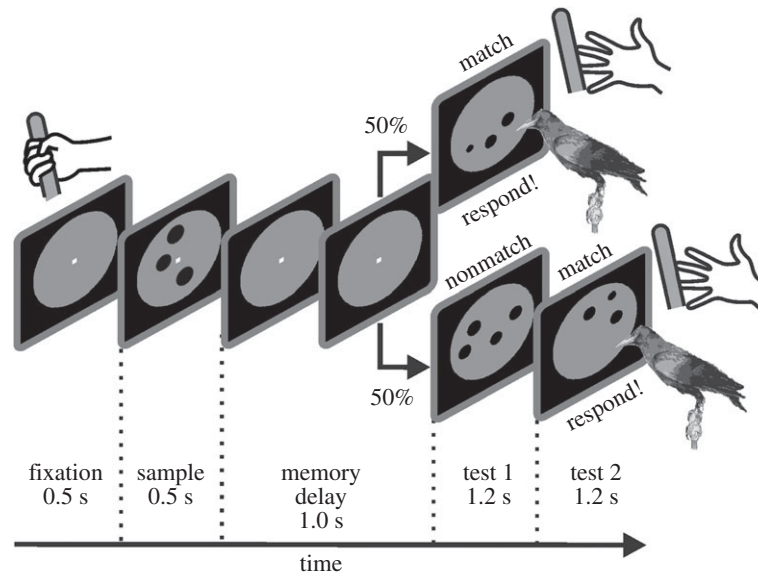
Figure 3. Delayed match-to-sample task used for both monkeys and crows with numerosity as discriminandum. A trial started when monkeys grabbed a manipulandum and fixated the fixation spot. Crows had to bring their head in position in front of the monitor to start a trial. A sample stimulus (here three dots) was followed by a delay period. The test stimulus contained either the same number of items ('match') or a different quantity of dots (here four dots) ('nonmatch'). Each nonmatch stimulus was followed by a match stimulus. Whenever a match stimulus appeared in the test phase, monkeys were required to release the lever to receive a reward, whereas crows needed to move their heads and peck at the display on the touch screen. Trials were pseudo-randomized and each numerosity was indicated with many different images.

2. Cognition of numerosity representation in monkeys and crows

(a) Conceptual grasp of numerical quantity in primates and crows

To allow a most direct comparison of behavioural (and later also neuronal) data, both monkeys [77] and crows [46] have been trained in equivalent operant conditioning protocols to perform computerized 'delayed match-to-sample tasks' (DMSTs) (figure 3). For every trial, a target numerosity (the *sample*) is initially presented on a computer monitor that the subject has to memorize over a brief delay period (without a stimulus). To earn a reward, the animals have to respond if the same target numerosity (the *match*) is shown again after a brief memory delay in the subsequent test phase. With equal probability, however, a deviant numerosity (a *nonmatch*) is presented in the test phase, requiring the subject to withhold responding. As a response to the match, monkeys were trained to immediately release a manipulandum that they grasped to start a trial. The crows, in contrast, were trained to keep their head still in front of a touch screen monitor throughout a trial, and peck with the beak at the match display as soon as it occurred. Based on the responses to the match and the nonmatch numerosity, the accuracy of the numerosity discrimination performance can be calculated.

This approach offers several advantages. First, it allows testing of the discrimination of specific and variable sample numerosities that change from trial to trial. Second, the animals are motivated to discriminate because of the reinforcement of correct trials (i.e. the reward), allowing testing of the scopes and limitations of their competence. Third, the computerized task permits the usage of arbitrary and controlled stimuli [46,77]. This is particularly important, given that non-numerical features of numerosity displays inevitably vary with an increase in the number of items.

For instance, in a spatial layout of dots (dot arrays) that need to be enumerated, the total area covered by the items and the overall density of the dots increases on average with increasing numbers of dots. On the other hand, when single items need to be enumerated one-by-one in a sequential layout (a protocol only used for monkeys so far), the time necessary to present the dots increases with more and more items. These and other potentially confounding sensory features have been equalized or controlled across numerosities to ensure that the animals are discriminating the stimuli based on the number of items rather than some low-level sensory cue. Finally, and as we shall see later, recording brain signals in animals under behavioural control is the most direct way to make sense of the brain signals that accompany (or rather cause) cognitive performance.

Using DMSTs, numerosity judgements in monkeys and crows have been explored in the visual domain by using multi-item arrays [46,47,77–81]. After training, rhesus monkeys (*Macaca mulatta*) and carrion crows (*Corvus corone*) proficiently discriminate the absolute values of visual numerosities from 1 to 5 and even from 1 to 30 items. Even a grasp of numerosity 0 was demonstrated in monkeys [82,83]. Hooded crows (*Corvus cornix*) were trained on an MST to discriminate the number of items in dot arrays [45]. Recently, carrion crows (*Corvus corone*) were trained to respond by pecking at a touch screen and they were able to successfully discriminate numerosities 1–5 and later up to 30 in a DMST irrespective of appearance and low-level visual features [46,47].

If animals have a concept of numerosity, they are expected to discriminate any possible numerosity, even those they have not been previously trained on. This was tested in a study in which two monkeys were initially only trained to discriminate numerosities 1–5 and then confronted with novel numerosities [81]. After proficient performance with numerosities 1–5, one monkey was confronted with novel numerosities 6, 7 and 8 in transfer trials in which it could not have learned the correct

response. The monkey continued to discriminate these novel numerosities with comparable accuracy to the previously learned small set sizes. To further demonstrate an abstract knowledge of the quantity concept, both monkeys in this study [81] only trained to small numerosities were abruptly confronted (i.e. from one day to the other) with numerosities ranging up to 30. Both monkeys showed spontaneous generalization to novel large numerosities, showing the same discrimination characteristics as for the well-trained small numerosities. Both the transfer and the generalization test argue in favour of a conceptual quantitative understanding. For crows, Smirnova *et al.* [45] reported successful transfer. After managing numerosities 1−4, hooded crows successfully transferred their discrimination to stimulus sets with novel values of 5−8. These studies show that monkeys and crows can apply the meaning of absolute numerosity to novel set sizes.

(b) Signatures of numerosity discrimination

What are the limits of animals' discrimination ability, and which numerical value can they discriminate from the immediately adjacent one? With nonmatch test stimuli consisting of one number higher and lower in equal probability (e.g. 3 and 5 for sample numerosity 4), the average discrimination accuracy in both monkeys and crows is close to perfect for the smallest numerosity, but declines rapidly towards large values, indicating that it becomes more difficult for them to discriminate progressively larger numerosities. Using 60% correct performance as criterion (which was significantly better than chance for the number of trials per numerosity based on a binomial test), the discrimination threshold was found to be between 4 and 5 items in monkeys [79], and slightly lower in crows [46]. Thus, with minimum numerical distance of 1 between match and nonmatch, the animals reliably discriminated numerical values 1 to approximately 4, but failed for numerosities of 5 and higher.

However, when the numerical distance between match and nonmatch is increased, performance recovers systematically [78]. For instance, if the discrimination between 3 and 4 is very defective, the discrimination between 3 and 5 is more accurate and the discrimination between 3 and 6 still better, indicating that discrimination of numerical values is easier for numerosity pairs that are more distant. This effect is termed the 'numerical distance effect'. As a consequence, discriminability is a function of the ratio between match and nonmatch numerosity (i.e. ratio-dependent) [80]. How accurately a specific target numerosity can be discriminated from any other numerosity can be evaluated by the monkeys' and crows' detailed numerosity discrimination behaviour. If the response probability to nonmatch numerosities relative to specific sample numerosities is plotted, bell-shaped performance curves result (figure 4). The numerical distance effect is reflected by the finding that the performance distributions have a certain width based on the approximate estimation. Both macaques and crows exhibit a numerical distance effect of similar magnitude, as witnessed by similar width of the performance curves in both monkeys and crows [46,78].

In addition to the distance effect, another numerical effect is evident in both species: the higher the numerical values, the larger the numerical distance between match and nonmatch numerosities must be in order for the numerosity pair to be just discriminable [47,78,81]. For instance, the discrimination between 2 and 3 is easier than discrimination

between 3 and 4, despite the numerical distance being 1 in both cases. This increase in the numerical distance needed for discrimination with increasing numerical values is called the 'numerical size effect'. The size effect causes the monkeys' and crows' bell-shaped performance curves to become broader with increasing values of the target numerosities (figure 4). In fact, a closer analysis shows that the distance between match and nonmatch numerosities grows in proportion to the numerical values for them to maintain just discriminable. For instance, if 6 can just be discriminated from 4, then 12 can just be discriminated from 8. Exactly the same proportional size relationship has been found for the performance of monkeys and crows [47,78,81].

(c) The approximate number system obeys Weber's Law

Both the numerical distance and size effect are classical signatures of the approximate number system (ANS) which represents the number of items in a set as a noisy mental magnitude. The ANS allows an estimation of set sizes of any numerical value, but becomes systematically less precise with increasing numerical values [48,84,85]. This effect is what Ernst Heinrich Weber captured for sensory magnitude discriminations in his law that is named after him. Weber's Law is a hallmark of the ANS [84,86], and numerosity discriminations in monkeys and crows obey it.

According to this law, the Weber fraction as a measure of discriminability (the quotient of the just noticeable numerical difference divided by the sample numerosity) is a constant across all magnitude values. Performance in both monkeys and crows exhibits constant Weber fractions across a broad range of set sizes (except for the very small numerosities 1 and 2). Jordan & Brannon [80] reported Weber fractions of 0.47 and 0.48 for two rhesus monkeys performing a DMST. Similarly, Merten & Nieder [81] found Weber fractions of 0.60 and 0.51, respectively, for two rhesus monkeys. Humans that were prevented from counting and tested in the same study with the same protocol on average showed very similar Weber fractions of 0.55 for numerosities larger than 6 [81].

Similar Weber fractions have been found for two carrion crows. For numerosities 1−5, the crows showed almost identical Weber fractions of 0.51 and 0.47, respectively [46]. Interestingly, the same two crows' Weber fractions markedly increased to 1.12 and 1.51 when they were tested with larger and numerically more distant numerosities 4, 7, 12, 20 and 30 [47]. Because the crows were not forced to discriminate as precisely as in the previous study [46], discrimination performance decreased (and therefore Weber fractions increased). This indicates that the choice of the numerosities the animals have to compare has an impact on precision.

The ANS has been demonstrated repeatedly in several animals [46,47,77−80,87]. Even in humans, this ancient system surfaces again in tasks that prevent counting [81,85]. Moreover, for humans who have never learned to count symbolically in their culture, the ANS is the only system available to estimate absolute numerosity [88−90]. Collectively, these behavioural data suggest that the ANS is an evolutionarily ancient nonverbal system which is able to assess set size across the animal kingdom, including humans.

The ANS found in monkeys and crows contrasts with the object file system. The object file system is thought to implicitly keep track of a small number of up to four items by
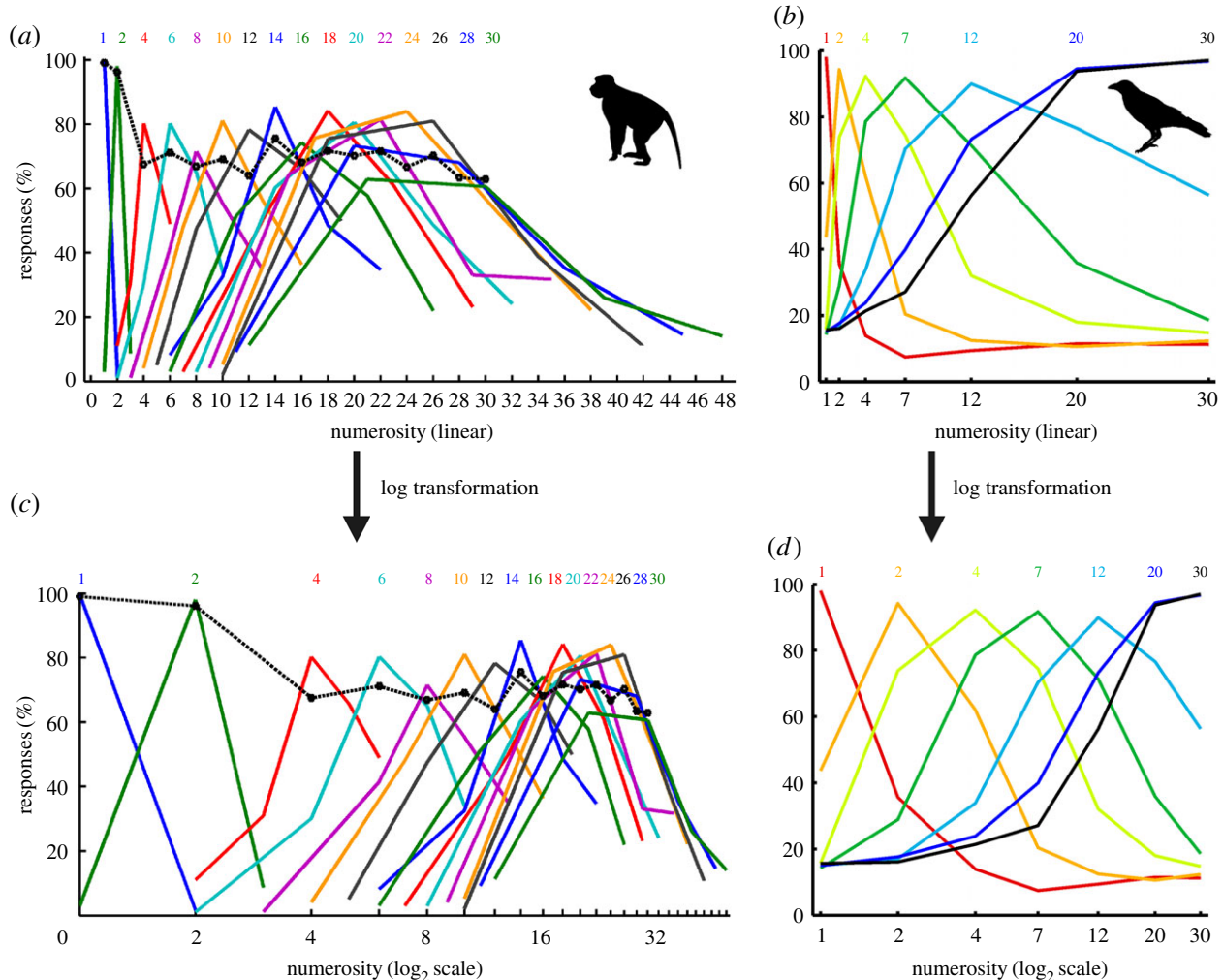
Figure 4. Behavioural performance functions of monkeys and crows for absolute numerosity discriminations during a DMST. (a) Performance functions of a representative monkey to sample numerosities 1–30. The bell-shaped functions reflect the percentage of trials in which a monkey judged displays in the test period as containing the same number of items as the sample numerosity. Colour indicates the numerosity of sample stimulus; the x-axis shows the test 1 numerosity. The data point at the centre of each coloured function indicates the correct performance in the match trials for the sample numerosities (shown in the same colour above each curve). The data points to the left and the right of the centre of each bell-shaped curve reflect erroneous responses to nonmatch numerosities. The same data are plotted on a linear number scale (top) and on a logarithmic scale (bottom). Note that the logarithmic scaling results in more symmetric curves when statistically tested across all functions. Data from [81]. (b) Performance functions of two crows to sample numerosities 1–30. Same layout as in (a). Data from [47]. (Online version in colour.)

assigning single markers (object files) to each to-be-enumerated element of a set. The object file system has been suggested for relative numerosity discriminations, particularly in human infants and chicks [91]. However, neither monkeys nor crows show signs of an object file system during absolute numerosity discriminations.

(d) Logarithmic representation

The bell-shaped performance functions resulting from absolute numerosity judgements also help to clarify how numerical quantity is represented along a number continuum. Quantitative examination of the shapes of the behavioural performance functions in monkeys and crows reveals that these performance functions are not symmetric when plotted on a linear number scale [46,47,78,81]. Instead, they are skewed, with steep slopes towards smaller numerosities, and shallow slopes towards larger numerosities (figure 4). This indicates that absolute numerosity representations are not appropriately described on a linear scale. However, when the same performance functions are plotted on a logarithmically compressed scale, the results are symmetric, bell-shaped Gauss functions (figure 4)

[46,47,78,81]. Because the logarithm of the responses is normally distributed, numerosity representations are also called lognormal distributions.

Interestingly, Siegler & Booth [92] also observed a logarithmic number representation in pre-school children. These researchers also found evidence that the logarithmic scale shifts to a linear number representation during school education. It seems that the logarithmic scale is the evolutionary default for nonverbal numerosity representations before number symbols require a linear transformation.

(e) Cross-format and cross-modal numerosity judgements in monkeys

As an abstract quantity category, numerosity is expected to be represented irrespective of the spatio-temporal presentation format. For example, three apples lying in a bowl or three light flashes in a sequence are both instances of 'three'. An animal is therefore expected to not only discriminate the number of items that are being represented across space (as in item arrays, i.e. ∴), but also over time (as in event sequences, i.e. ● - ● - ●).

In a combined behavioural/electrophysiological experiment, two monkeys were first trained to match the numerosity of dot arrays. Once they mastered simultaneously displayed items in spatial arrays, trials with sequential sample numerosities were added. One to four single dots appearing one-by-one in the sample phase were shown to the monkey while carefully controlling for temporal factors [93]. The monkeys had to memorize the sequential number of dots and respond if the same numerosity was displayed in the test period. Once the monkeys learned to discriminate sequential numerosity 2 from 4 (and vice versa), they were immediately able to transfer discrimination to the novel sequential numerosity 3 versus 2 and 3 versus 4 without further training. This indicates that monkeys can understand the concept of sequential numerosity and apply it together with simultaneous numerosity protocols. However, the performance data showed that discrimination of the number of sequentially presented items was more difficult for both monkeys.

Numerosity is also independent of the sensory modality of the items that make up a set. Three light flashes or three calls are both instances of 'three' despite the elements being visual items in the first, and acoustic events in the second case. To find out if monkeys can perform such cross-modal numerosity judgements, Jordan et al. [94] presented rhesus monkeys with a sample sequence of 1–9 visual items or acoustic sounds. After sample presentation, the monkeys correctly chose the matching visual test arrays that contained the same number for both visual and auditory sample numerosities. Moreover, they were just as accurate in matching across sensory modalities (auditory–visual) as they were within a single modality (visual–visual). As expected for the ANS, performance was dependent on the ratio between the displayed match and the nonmatch numerosities. Similar cross-modal numerosity discriminations have been reported in an electrophysiological study with monkeys trained to discriminate 1–4 visual or acoustic items in a DMST [95].

Jordan and colleagues [94] went an important step further and also presented samples that consisted of interspersed visual and acoustic items. For example, monkeys had to match two visual items and two sounds to the test numerosity 4. In the first 150-trial test session, the monkeys already performed with above chance accuracy for these bimodal visual–auditory samples. This demonstrates that nonhuman primates can cross-modally enumerate the number of visual objects they see and the number of sounds they hear. They show this capacity over a relatively large range of numerosities and use the ANS to solve the task.

3. Neurophysiology of numerosity representations in monkeys and crows

(a) Neurons specifically responsive to numerical quantity

The performance data reviewed showed that the behavioural signatures of numerosity discriminations found in monkeys and crows are more or less identical and point to an ANS that obeys the Weber Law. However, what looks like a fundamental similarity may turn out to be a superficial result of very different underlying neuronal mechanisms. After all, the endbrain circuitries giving rise to high-level cognitive capabilities in primates and corvids are a product of convergent evolution and therefore differ fundamentally. Studying behaviour alone cannot tell us whether numerical judgements truly depend on the same neural representations and if they really follow the Weber Law [96,97]. The behavioural outcome of a numerical discrimination task may simply be the result of multiple, diverse scaling schemes at different processing stages. What looks like numerosity estimation emerging from the same code for ANS in both primate and corvid species may have quite different neuronal realizations.

Animals trained on a well-controlled DMST with numerosity as discriminandum provide an ideal opportunity to study brain processes of numerical competence. The DMST has been used for many decades in primate neurophysiology studies [56] and is also easily mastered by corvids [98]. To investigate how single neurons in the endbrain of monkeys and crows function, we recorded single neuron activity while the animals assessed and memorized numerical values. Recording neuronal activity simultaneously with behavioural performance presents a rich and direct opportunity for experimental analysis of the neuronal foundations of numerical cognition that would not be possible in untrained animals.

Neurons recorded in the associative endbrains of both monkeys and crows that discriminated visual numerosity showed single neurons that selectively responded to numerosity displays during the sample period and during memorization of numerosity in the delay period [46,77,78,99,100]. Such numerosity-selective neurons, also termed 'number neurons', increased the number of action potentials (i.e. their activity) during these time intervals in response to the presentation or memorization of numerical quantities (figure 5a,b). Importantly, changes in the physical appearance of the displays had no effect on the activity of numerosity-selective neurons [46,77,78,99,100]; all the neurons cared about was numerical information. In monkeys, a relatively large proportion of about 20–30% of randomly recorded neurons in the IPS (particularly in ventral intraparietal area, VIP) and the lateral PFC signalled numerosity. A very similar proportion of neurons responded to numerosity in the NCL of behaving crows.

(b) Properties of the activity of numerosity-selective neurons

The neuronal data show an impressive correspondence of the neuronal code found in the primate and avian brain. Numerosity-selective neurons are tuned to the number of items both in the fronto-parietal association cortex of monkeys and the NCL of crows (figure 5a,b). This means that neurons respond with maximum activity to one of the presented quantities—a neuron's 'preferred numerosity'—and progressively decrease activity as the displayed quantity departs from the preferred number [46,77,78,99,100].

The result of this way of encoding a quantity category is a bell-shaped distribution of activity (i.e. action potential rates) around the respective preferred numerosity of each tuned neuron. A neuron responds strongest to its preferred numerosity, but it also responds to slightly smaller and larger numerosities relative to the preferred numerosity. Numerosity-selective neurons therefore encode the numerical value in an approximate fashion, just as animals do in their behaviour.

Since a single neuron can only represent a very small range of numerical values, a population of neurons with neurons each tuned to different numerosities is required to cover the continuum of values from 1 upwards. Interestingly, the
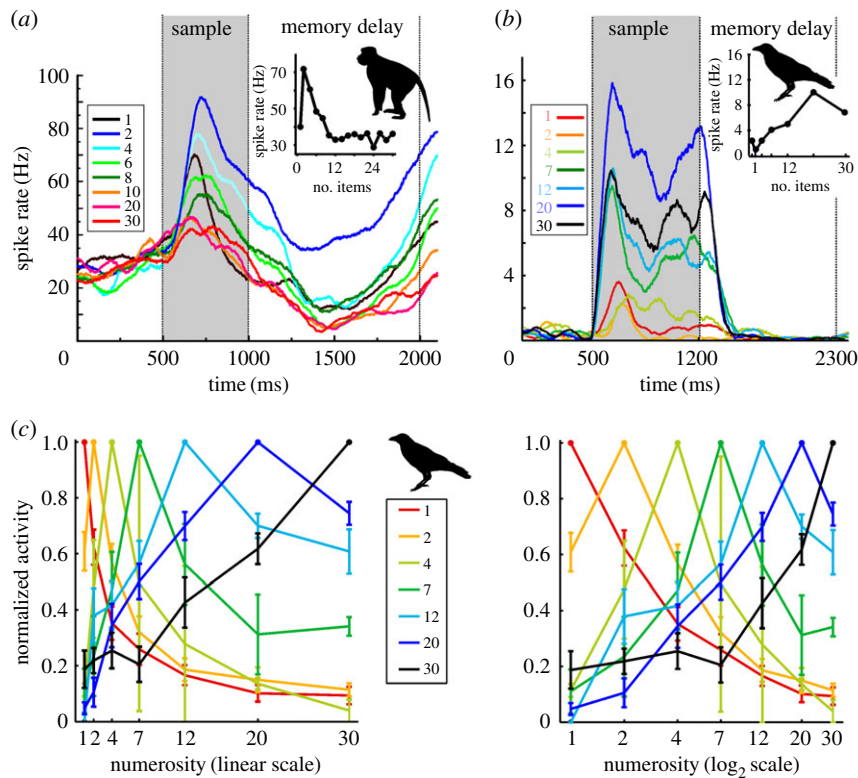
8

Figure 5. Response characteristics of numerosity-selective neurons in monkeys and crows. (a) Neurons of the monkey association cortex (here the PFC) selectively respond to specific numerosities. The time course of neuronal activity during trials shows that this example neuron responded highest whenever two dots were shown: it preferred number 2. Firing rates dropped gradually for numbers more distant from number 2, indicating neuronal tuning to numbers. Note that this neuron signals numerosity 2 both during sample presentation and in working memory during the following delay phase. Discharges to all even numbers (and 1) up to 30 for many stimulus repetitions are plotted as averaged spike density functions (neural activity is averaged and smoothed and plotted over time per trials, and only a selection of the numbers are shown for clarity). Colours correspond to specific tested numbers. Data from [100]. (b) Neurons of the corvid NCL also respond to preferred numerosities. This NCL neuron is tuned to numerosity 20. Same layout as in (a). Data from [99]. (c) Normalized (bell-shaped) tuning functions averaged for all neurons preferring the same numerosity (indicated by colour) when plotted on a linear (left) or logarithmic number scale (right) (error bar ± s.e.m.). Data from [99]. (Online version in colour.)

neurons' sequentially-arranged overlapping tuning curves preserved an inherent order of numerical values (figure 5c) [46,77,78,99,100]. This is helpful because numerosities are not isolated categories, but exist in relation to one another (for example, 3 is greater than 2 and less than 4); numerical values need to be sequentially ordered to allow meaningful quantity assignments.

This fuzzy selectivity of a population of numerosity-selective neurons explains the numerical distance effect, i.e. why the animals which have to base their judgement on the responses of neurons cannot precisely discriminate similar numerical values. Because the tuning curves overlap quite a bit, a pair of similar numerosities may activate neurons tuned to different preferred numerosity in about equal amounts, thus creating an ambiguous signal that causes an animal to mix up numerosities and make many errors. However, if very dissimilar numerosities have to be discriminated, the tuning functions of the respectively activated neurons will hardly overlap. In this situation, neurons provide an animal with unambiguous signals which leads to high discrimination accuracy. Because the neuronal tuning functions of numerosity-selective neurons are also broader with increasing preferred numerosity, the same logic can also explain the additional numerical size effect [48].

The tuning curves of numerosity-selective neurons show an even more fundamental correspondence between monkeys and crows, as do behavioural findings in the respective species (figure 5c). Much like the behavioural data, the neural filter

functions are lognormal distributions: only when plotted on a logarithmic scale do they become symmetric [46,78,99,100]. Moreover, the variance (i.e. width) of neural distributions is constant with increasing preferred numerosity when the data are plotted on a logarithmic scale. In terms of the scaling scheme, the neural data in monkeys and crows mirror each other, and the behavioural findings in both species.

Interestingly, numerosity tuning curves were also indirectly reconstructed from blood-flow measurements using functional magnetic resonance imaging (fMRI) in humans. These functions show the same logarithmic relationship [52,101,102]. This suggests that nonverbal numerosity representations in the brains of humans, and nonhuman primates and crows share fundamental characteristics.

(c) Neurons relevant for behaviour

If numerosity-selective neurons are truly relevant for the animals' numerosity discrimination performance, their activity should reflect successful and erroneous trails. More precisely, if the neurons do not signal their preferred numerosity clearly with high action potential rates, the animals evaluating imprecisely tuned neurons would be prone to make errors. Indeed, when either monkeys or crows made judgement errors, neuronal activity to the preferred numerosity is significantly reduced, for both neurons of the fronto-parietal network in primates and those of the corvid NCL [46,77,78,99,100]. This provides evidence that numerosity-selective neurons are at

least partly responsible for the numerical capabilities found in primates and corvids.

(d) Abstract numerosity neurons in monkeys

Sequential or cross-modal numerosity discriminations so far have only been combined with electrophysiological recordings in monkeys. As expected for an abstract representation of numerosity, both simultaneous and sequential numerosity formats were encoded in the parietal lobe. Neurons represented the number of simultaneously and sequentially presented dots [86] or the number of executed hand movements [50]. In a cross-modal discrimination study mentioned earlier, groups of neurons encoded the number of either auditory pulses, visual items, or both [88]. Interestingly, a significant proportion of neurons in the PFC were tuned to the same number of visual and auditory items. For instance, a neuron responded maximally whenever three dots in a row or three sounds in a row were presented. This argues that cells in the PFC can represent absolute numerosity irrespective of sensory modalities in a supra-modal fashion. While some neurons in the monkey association cortex constitute very abstract numerosity detectors, such data await demonstration in crows.

4. Evolution has a taste for logarithmic representation

At both the behavioural and the neural level, primates and corvids with their distinctly evolved endbrains represent the number of items in a set as a noisy numerical value in accordance with Weber's Law. What is more, these animals' numerosity judgements (figure 4) and the tuning curves of numerosity-selective neurons in the primate neocortex and the avian nidopallium (figure 5) vary logarithmically with the number of elements in a set [46,78,99,100]. This logarithmic variation can be reconciled when considering Gustav Theodor Fechner's [103] extension of Weber's Law. Fechner postulated that linear increments in sensation S are proportional to the logarithm of stimulus magnitude I, a relationship known as the Weber–Fechner Law ($S = k \log I$). Surprisingly, this fundamental law, which is largely valid for general sensory phenomena, also holds true for derived cognitive magnitudes such as numerosity [78]. Even distance perception and time perception vary logarithmically with the length of distance and the time interval, respectively, and so do many other physiological parameters [104].

Already, prior to neuronal recordings studies, computational network models have predicted tuned and logarithmically scaled 'numerosity detectors' for analogue number representations [105]. This numerosity detector model encoded numerosity from parallel input which best mirrors the processing of numerosities from dot arrays discussed here for monkeys and crows. Later, Verguts & Fias [106] showed that an initially uncommitted neural backpropagation network also developed logarithmically scaled numerosity units under unsupervised learning when collections of dots were provided as input.

The logarithmic representation resulting in symmetric bell-shaped distributions has its advantages. First, it causes the variability (i.e. width) of the representations to be constant across, and independent of, the target number. This ensures that discrimination accuracy is proportional to the magnitude

of the numerosities at stake. This leads to high precision for small numerosities at the expense of fuzzy large-numerosity representations. The adaptive value of such a representation for an animal is evident: for example, whether an animal faces one or two opponents is crucial information to decide whether to attack or retreat and therefore deserves precise judgement. In contrast, whether it faces 20 or 40 adversaries is irrelevant for such a decision because it should take to its heels regardless [107]. A second advantage refers to the range of encoded numerosities based on a limited number of neurons. Even if noisy, a logarithmic compression still allows the processing of almost arbitrarily large numbers, and with relatively few neurons. These advantages may explain why the physiology of numerosity representation obeys the Weber–Fechner Law in animal groups with a very different phylogenetic history resulting in independently evolved endbrain structures.

These behavioural and neurophysiological data for absolute numerosity judgements in different animal species show an impressive correspondence and help to resolve a classical debate in psychophysics: the mental number line for nonverbal numerical information is logarithmic rather than linear, and not just in human and nonhuman primates, but probably across vertebrates and maybe even beyond. It suggests that this way of coding numerical information has evolved based on convergent evolution because it exhibits a superior solution to a common computational problem.

5. Outlook

Until very recently, the macaque monkey was the only species for which we had single neuron data in behaving animals. Several studies showed how neurons in the behaving monkey brain encode and process numerical information, thus complementing data in humans via indirect functional imaging measures. In contrast to advances in nonhuman primates, investigations into the neurophysiology of numerical competence in corvids have only just begun, and much still needs to be explored.

First, so far only visual multi-item arrays have been used as numerosity stimuli in crows [46,99]. To pin down the level of coding abstraction, numerosities need to be presented also in other modalities. Crows can learn delayed cross-modal tasks and many NCL neurons associate visual and auditory stimuli [108,109]. Thus, cross-modal numerical tasks similar to those used in monkeys [94,95] are feasible.

Second, numerosity can be presented simultaneously in spatial arrays, or sequentially in item successions. In monkeys, simultaneous and sequential enumeration mechanisms seem to differ in the IPS [93]. Recording from crows trained to perform both simultaneous and sequential protocols could clarify if this distinction is also present in the NCL.

Third, we now know that NCL neurons signal numerosity, but this brain area may not be the only one involved in extracting and maintaining numerical quantity information. In the primate, numerical information seems to be processed in a hierarchical fashion from upstream parietal to downstream prefrontal areas [48]. Since the NCL is commonly compared to the PFC, it begs the question if the avian brain also contains an input processing stage equivalent to the IPS.

Fourth, the classical idea of a 'sense of number' [110–112] suggests that we and animals are endowed with a hard-wired

faculty to perceive the number of items. This idea argues that animals assess numerosity spontaneously and without the need to learn. Indeed, numerosity-selective neurons were recently reported in numerically naive monkeys [113,114], providing compelling evidence for this idea. To generalize the number sense concept beyond primates, similar investigations are needed in numerically naive crows, but so far all experiments done in corvids have been performed in numerosity-trained animals.

Last but not least, monkeys can learn to discriminate not only countable numerosities, but also empty sets [82,83]. Monkeys represent empty sets as quantitative categories at the lower end of a numerical continuum. Moreover, neurons in the IPS encode empty sets predominantly as a category distinct from numerosities, whereas PFC neurons represent empty sets like countable numerosities, exhibiting numerical distance and size effects [83,115]. It would be interesting to

know if crows also can learn to grasp empty sets as numerical quantity, and whether and how numerosity 0 is also represented by NCL neurons.

With the addition of such data, a more concise picture will surface that will allow us to fully answer the question as to which physiological solutions for numerical competence implemented during the evolution of crows and primates are similar in all respects. Only more comparative approaches in neuroscience will help to decipher the presence of evolutionarily stable neuronal mechanisms [116].

References

1. Dacke M, Srinivasan MV. 2008 Evidence for counting in insects. *Anim. Cogn.* **11**, 683–689. (doi:10.1007/s10071-008-0159-y)

2. Gross HJ, Pahl M, Si A, Zhu H, Tautz J, Zhang S, Tanimoto H. 2009 Number-based visual generalisation in the honeybee. *PLoS ONE* **4**, e4263. (doi:10.1371/journal.pone.0004263)

3. Agrillo C, Piffer L, Bisazza A. 2011 Number versus continuous quantity in numerosity judgments by fish. *Cognition* **119**, 281–287. (doi:10.1016/j.cognition.2010.10.022)

4. Gómez-Laplaza LM, Gerlai R. 2012 Activity counts: the effect of swimming activity on quantity discrimination in fish. *Front Psychol.* **3**, 484. (doi:10.3389/fpsyg.2012.00484)

5. Bisazza A, Tagliapietra C, Bertolucci C, Foà A, Agrillo C. 2014 Non-visual numerical discrimination in a blind cavefish (*Phreatichthys andruzzii*). *J. Exp. Biol.* **217**, 1902–1909. (doi:10.1242/jeb.101683)

6. Potrich D, Sovrano VA, Stancher G, Vallortigara G. 2015 Quantity discrimination by zebrafish (*Danio rerio*). *J. Comp. Psychol.* **129**, 388–393. (doi:10.1037/com0000012)

7. Uller C, Jaeger R, Guidry G, Martin C. 2003 Salamanders (*Plethodon cinereus*) go for more: rudiments of number in an amphibian. *Anim. Cogn.* **6**, 105–112. (doi:10.1007/s10071-003-0167-x)

8. Krusche P, Uller C, Dicke U. 2010 Quantity discrimination in salamanders. *J. Exp. Biol.* **213**, 1822–1828. (doi:10.1242/jeb.039297)

9. Koehler O. 1951 The ability of birds to 'count.' *Bull. Anim. Behav.* **9**, 41–45.

10. Emmerton J, Renner JC. 2006 Scalar effects in the visual discrimination of numerosity by pigeons. *Learn. Behav.* **34**, 176–192. (doi:10.3758/BF03193193)

11. Rugani R, Fontanari L, Simoni E, Regolin L, Vallortigara G. 2009 Arithmetic in newborn chicks. *Proc. R. Soc. B* **276**, 2451–2460. (doi:10.1098/rspb.2009.0044)

12. Scarf D, Hayne H, Colombo M. 2011 Pigeons on par with primates in numerical competence. *Science* **334**, 1664. (doi:10.1126/science.1213357)

13. Rugani R, Cavazzana A, Vallortigara G, Regolin L. 2013 One, two, three, four, or is there something more? Numerical discrimination in day-old domestic chicks. *Anim. Cogn.* **16**, 557–564. (doi:10.1007/s10071-012-0593-8)

14. Rugani R, Vallortigara G, Regolin L. 2014 From small to large: numerical discrimination by young domestic chicks (*Gallus gallus*). *J. Comp. Psychol.* **128**, 163–171. (doi:10.1037/a0034513)

15. Davis H, Albert M. 1986 Numerical discrimination by rats using sequential auditory stimuli. *Anim. Learn. Behav.* **14**, 57–59. (doi:10.3758/BF03200037)

16. Brannon EM, Terrace HS. 1998 Ordering of the numerosities 1 to 9 by monkeys. *Science* **282**, 46–49. (doi:10.1126/science.282.5389.746)

17. Vonk J, Beran MJ. 2012 Bears 'count' too: quantity estimation and comparison in black bears (*Ursus americanus*). *Anim. Behav.* **84**, 231–238. (doi:10.1016/j.anbehav.2012.05.001)

18. Lyon BE. 2003 Egg recognition and counting reduce costs of avian conspecific brood parasitism. *Nature* **422**, 95–99. (doi:10.1038/nj6927-095a)

19. Carazo P, Fernández-Perea R, Font E. 2012 Quantity estimation based on numerical cues in the mealworm beetle (*Tenebrio molitor*). *Front. Psychol.* **3**, 502. (doi:10.3389/fpsyg.2012.00502)

20. Stancher G, Rugani R, Regolin L, Vallortigara G. 2015 Numerical discrimination by frogs (*Bombina orientalis*). *Anim. Cogn.* **18**, 219–229. (doi:10.1007/s10071-014-0791-7)

21. Nelson XJ, Jackson RR. 2012 The role of numerical competence in a specialized predatory strategy of an araneophagic spider. *Anim. Cogn.* **15**, 699–710. (doi:10.1007/s10071-012-0498-6)

22. Gómez-Laplaza LM, Gerlai R. 2011 Can angelfish (*Pterophyllum scalare*) count? Discrimination between different shoal sizes follows Weber's law. *Anim. Cogn.* **14**, 1–9. (doi:10.1007/s10071-010-0337-6)

23. McComb K, Packer C, Pusey AE. 1994 Roaring and numerical assessment in contests between groups of female lions, *Panthera leo*. *Anim. Behav.* **47**, 379–387. (doi:10.1006/anbe.1994.1052)

24. Wilson ML, Hauser MD, Wrangham, RW. 2001 Does participation in intergroup conflict depend on numerical assessment, range location, or rank for wild chimpanzees? *Anim. Behav.* **61**, 1203–1216. (doi:10.1006/anbe.2000.1706)

25. Benson-Amram S, Heinen K, Dryer SL, Holekamp KE. 2011 Numerical assessment and individual call discrimination by wild spotted hyaenas, *Crocuta crocuta*. *Anim. Behav.* **82**, 743–752. (doi:10.1016/j.anbehav.2011.07.004)

26. Kumar S, Hedges SB. 1998 A molecular timescale for vertebrate evolution. *Nature* **392**, 917–920. (doi:10.1038/31927)

27. Hedges SB. 2002 The origin and evolution of model organisms. *Nat. Rev. Genet.* **3**, 838–849. (doi:10.1038/nrg929)

28. Miller EK, Nieder A, Freedman DJ, Wallis JD. 2003 Neural correlates of categories and concepts. *Curr. Opin. Neurobiol.* **13**, 198–203. (doi:10.1016/S0959-4388(03)00037-0)

29. Stoet G, Snyder LH. 2009 Neural correlates of executive control functions in the monkey. *Trends Cogn. Sci.* **13**, 228–234. (doi:10.1016/j.tics.2009.02.002)

30. Flombaum JI, Santos LR. 2005 Rhesus monkeys attribute perceptions to others. *Curr. Biol.* **15**, 447–452. (doi:10.1016/j.cub.2004.12.076)

31. Subiaul F, Cantlon JF, Holloway RL, Terrace HS. 2004 Cognitive imitation in rhesus macaques. *Science* **305**, 407–410. (doi:10.1126/science.1099136)

32. Chang SW, Gariépy JF, Platt ML. 2013 Neuronal reference frames for social decisions in primate frontal cortex. *Nat. Neurosci.* **16**, 243–250. (doi:10.1038/nn.3287)

33. Hauser MD, Carey S, Hauser LB. 2000 Spontaneous number representation in semi-free-ranging rhesus

monkeys. *Proc. R. Soc. Lond. B* **267**, 829–833. (doi:10.1098/rspb.2000.1078)

34. Wang Z *et al.* 2013 The draft genomes of soft-shell turtle and green sea turtle yield insights into the development and evolution of the turtle-specific body plan. *Nat. Genet.* **45**, 701–706. (doi:10.1038/ng.2615)

35. Zucca P, Milos N, Vallortigara G. 2007 Piagetian object permanence and its development in Eurasian jays (*Garrulus glandarius*). *Anim. Cogn.* **10**, 243–258. (doi:10.1007/s10071-006-0063-2)

36. 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. (doi:10.1016/j.anbehav.2011.05.012)

37. Veit L, Nieder A. 2013 Abstract rule neurons in the endbrain support intelligent behaviour in corvid songbirds. *Nat. Commun.* **4**, 2878. (doi:10.1038/ncomms3878)

38. Smirnova A, Zorina Z, Obozova T, Wasserman E. 2015 Crows spontaneously exhibit analogical reasoning. *Curr. Biol.* **25**, 256–260. (doi:10.1016/j.cub.2014.11.063)

39. Raby CR, Alexis DM, Dickinson A, Clayton NS. 2007 Planning for the future by western scrub-jays. *Nature* **445**, 919–921. (doi:10.1038/nature05575)

40. Emery NJ, Clayton NS. 2001 Effects of experience and social context on prospective caching strategies by scrub jays. *Nature* **414**, 443–446. (doi:10.1038/35106560)

41. Bugnyar T, Reber SA, Buckner C. 2016 Ravens attribute visual access to unseen competitors. *Nat. Commun.* **7**, 10506. (doi:10.1038/ncomms10506.)

42. Ujfalussy D, Miklósi A, Bugnyar T, Kotrschal K. 2013 Role of mental representations in quantity judgments by jackdaws (*Corvus monedula*). *J. Comp. Psychol.* **128**, 11–20. (doi:10.1037/a0034063)

43. Bogale BA, Aoyama M, Sugita S. 2014 Spontaneous discrimination of food quantities in the jungle crow, *Corvus macrorhynchos*. *Anim. Behav.* **94**, 73–78. (doi:10.1016/j.anbehav.2014.05.012)

44. Tornick JK, Callahan ES, Gibson BM. 2015 An investigation of quantity discrimination in Clark's nutcrackers (*Nucifraga columbiana*). *J. Comp. Psychol.* **129**, 17–25. (doi:10.1037/a0037863)

45. Smirnova AA, Lazareva OF, Zorina ZA. 2000 Use of number by crows: investigation by matching and oddity learning. *J. Exp. Anal. Behav.* **73**, 163–176. (doi:10.1901/jeab.2000.73-163)

46. Ditz HM, Nieder A. 2015 Neurons selective to the number of visual items in the corvid songbird endbrain. *Proc. Natl Acad. Sci. USA* **112**, 7827–7832. (doi:10.1073/pnas.1504245112)

47. Ditz HM, Nieder A. 2016 Numerosity representations in crows obey the Weber-Fechner law. *Proc. R. Soc. B* **283**, 20160083. (doi:10.1098/rspb.2016.0083)

48. Nieder A. 2016 The neuronal code for number. *Nat. Rev. Neurosci.* **17**, 366–382. (doi:10.1038/nrn.2016.40)

49. Dehaene S, Spelke E, Pinel P, Stanescu R, Tsivkin S. 1999 Sources of mathematical thinking: behavioral and brain-imaging evidence. *Science* **284**, 970–974. (doi:10.1126/science.284.5416.970)

50. Sawamura H, Shima K, Tanji J. 2002 Numerical representation for action in the parietal cortex of the monkey. *Nature* **415**, 918–922. (doi:10.1038/415918a)

51. Nieder A, Miller EK. 2004 A parieto-frontal network for visual numerical information in the monkey. *Proc. Natl Acad. Sci. USA* **101**, 7457–7462. (doi:10.1073/pnas.0402239101)

52. 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. (doi:10.1016/j.neuron.2004.10.014)

53. Vallentin D, Bongard S, Nieder A. 2012 Numerical rule coding in the prefrontal, premotor, and posterior parietal cortices of macaques. *J. Neurosci.* **32**, 6621–6630. (doi:10.1523/JNEUROSCI.5071-11.2012)

54. Jacob SN, Nieder A. 2014 Complementary roles for primate frontal and parietal cortex in guarding working memory from distractor stimuli. *Neuron* **83**, 226–237. (doi:10.1016/j.neuron.2014.05.009)

55. Ott T, Jacob SN, Nieder A. 2014 Dopamine receptors differentially enhance rule coding in primate prefrontal cortex neurons. *Neuron* **84**, 1317–1328. (doi:10.1016/j.neuron.2014.11.012)

56. Fuster JM, Alexander GE. 1971 Neuron activity related to short-term memory. *Science* **173**, 652–654. (doi:10.1126/science.173.3997.652)

57. Merten K, Nieder A. 2012 Active encoding of decisions about stimulus absence in primate prefrontal cortex neurons. *Proc. Natl Acad. Sci. USA* **109**, 6289–6294. (doi:10.1073/pnas.1121084109)

58. Olkowicz S, Kocourek M, Lučan RK, Porteš M, Fitch WT, Herculano-Houzel S, Němec P. 2016 Birds have primate-like numbers of neurons in the forebrain. *Proc. Natl Acad. Sci. USA* **113**, 7255–7260 (doi:10.1073/pnas.1517131113)

59. Karten HJ. 2015 Vertebrate brains and evolutionary connectomics: on the origins of the mammalian 'neocortex'. *Phil. Trans. R. Soc. B.* **370**, 20150060. (doi:10.1098/rstb.2015.0060)

60. Dugas-Ford J, Ragsdale CW. 2015 Levels of homology and the problem of neocortex. *Annu. Rev. Neurosci.* **38**, 351–368. (doi:10.1146/annurev-neuro-071714-033911)

61. Jarvis ED *et al.* 2005 Avian brains and a new understanding of vertebrate brain evolution. *Nat. Rev. Neurosci.* **6**, 151–159. (doi:10.1038/nrn1606)

62. Butler AB, Reiner A, Karten HJ. 2011 Evolution of the amniote pallium and the origins of mammalian neocortex. *Ann. NY Acad. Sci.* **1225**, 14–27. (doi:10.1111/j.1749-6632.2011.06006.x)

63. Puelles L, Kuwana E, Puelles E, Bulfone A, Shimamura K, Keleher J, Smiga S, Rubenstein JL. 2000 Pallial and subpallial derivatives in the embryonic chick and mouse telencephalon, traced by the expression of the genes Dlx-2, Emx-1, Nkx-2.1, Pax-6, and Tbr-1. *J. Comp. Neurol.* **424**, 409–438. (doi:10.1002/1096-9861(20000828)424:3<409::AID-CNE3>3.0.CO;2-7)

64. Reiner A *et al.* 2004 Revised nomenclature for avian telencephalon and some related brainstem nuclei.

J. Comp. Neurol. **473**, 377–414. (doi:10.1002/cne.20118)

65. Güntürkün O, Bugnyar T. 2016 Cognition without cortex. *Trends Cogn. Sci.* **20**, 291–303. (doi:10.1016/j.tics.2016.02.001)

66. Divac I, Mogensen J. 1985 The prefrontal 'cortex' in the pigeon catecholamine histofluorescence. *Neuroscience* **15**, 677–682. (doi:10.1016/0306-4522(85)90069-7)

67. Güntürkün O. 2005 The avian 'prefrontal cortex' and cognition. *Curr. Opin. Neurobiol.* **15**, 686–693. (doi:10.1016/j.conb.2005.10.003)

68. Veit L, Hartmann K, Nieder A. 2017 Spatially tuned neurons in corvid nidopallium caudolaterale signal target position during visual search. *Cereb. Cortex* **27**, 1103–1112. (doi:10.1093/cercor/bhv299)

69. Nieder A. 2017 Inside the corvid brain—probing the physiology of cognition in crows. *Curr. Opin. Behav. Sci.* 2017, **16**, 8–14 (doi:10.1016/j.cobeha.2017.02.005)

70. Medina L, Reiner A. 2000 Do birds possess homologues of mammalian primary visual, somatosensory and motor cortices? *Trends Neurosci.* **23**, 1–12. (doi:10.1016/S0166-2236(99)01486-1)

71. Reiner A. 2013 You are who you talk with—a commentary on Dugas-Ford *et al.* PNAS, 2012. *Brain Behav. Evol.* **81**, 146–149. (doi:10.1159/000348281)

72. Medina L, Abellán A, Desfilis E. 2013 A never-ending search for the evolutionary origin of the neocortex: rethinking the homology concept. *Brain Behav. Evol.* **81**, 150–153. (doi:10.1159/000348282)

73. Faunes M, Botelho JF, Ahumada Galleguillos P, Mpodozis J. 2015 On the hodological criterion for homology. *Front. Neurosci.* **9**, 223. (doi:10.3389/fnins.2015.00223)

74. Martinez-Garcia F, Novejarque A, Lanuza, E. 2007 Evolution of the amygdala in vertebrates. In *Evolution of nervous systems*, vol. 2 (ed. JH Kaas), pp. 255–334. Oxford, UK: Oxford University Press.

75. Rattenborg NC, Martinez-Gonzalez D, Roth 2nd TC, Pravosudov VV. 2011 Hippocampal memory consolidation during sleep: a comparison of mammals and birds. *Biol. Rev. Camb. Philos. Soc.* **86**, 658–691. (doi:10.1111/j.1469-185X.2010.00165.x)

76. Schnupp JW, Carr CE. 2009 On hearing with more than one ear: lessons from evolution. *Nat. Neurosci.* **12**, 692–697. (doi:10.1038/nn.2325)

77. Nieder A, Freedman DJ, Miller EK. 2002 Representation of the quantity of visual items in the primate prefrontal cortex. *Science* **297**, 1708–1711. (doi:10.1126/science.1072493)

78. Nieder A, Miller EK. 2003 Coding of cognitive magnitude: compressed scaling of numerical information in the primate prefrontal cortex. *Neuron* **37**, 149–157. (doi:10.1016/S0896-6273(02)01144-3)

79. Nieder A, Miller EK. 2004 Analog numerical representations in rhesus monkeys: evidence for parallel processing. *J. Cogn. Neurosci.* **16**, 889–901. (doi:10.1162/089892904970807)

80. Jordan KE, Brannon EM. 2006 Weber's Law influences numerical representations in rhesus macaques (*Macaca mulatta*). *Anim. Cogn.* **9**, 159–172. (doi:10.1007/s10071-006-0017-8)

81. Merten K, Nieder A. 2009 Compressed scaling of abstract numerosity representations in adult humans and monkeys. *J. Cogn. Neurosci.* **21**, 333–346. (doi:10.1162/jocn.2008.21032)

82. Merritt DJ, Rugani R, Brannon EM. 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. (doi:10.1037/a0015231)

83. Ramirez-Cardenas A, Moskaleva M, Nieder A. 2016 Neuronal representation of numerosity zero in the primate parieto-frontal number network. *Curr. Biol.* **26**, 1285–1294. (doi:10.1016/j.cub.2016.03.052)

84. Dehaene S. 1992 Varieties of numerical abilities. *Cognition* **44**, 1–42. (doi:10.1016/0010-0277(92)90049-N)

85. Whalen J, Gallistel CR, Gelman R. 1999 Nonverbal counting in humans: the psychophysics of number representation. *Psychol. Sci.* **10**, 130–137. (doi:10.1111/1467-9280.00120)

86. Moyer RS, Landauer TK. 1967 Time required for judgements of numerical inequality. *Nature* **215**, 1519–1520. (doi:10.1038/2151519a0)

87. Cantlon JF, Brannon EM. 2006 Shared system for ordering small and large numbers in monkeys and humans. *Psychol. Sci.* **17**, 401–406. (doi:10.1111/j.1467-9280.2006.01719.x)

88. Gordon P. 2004 Numerical cognition without words: evidence from Amazonia. *Science* **306**, 496–499. (doi:10.1126/science.1094492)

89. Pica P, Lemer C, Izard V, Dehaene S. 2004 Exact and approximate arithmetic in an Amazonian indigene group. *Science* **306**, 499–503. (doi:10.1126/science.1102085)

90. Frank MC, Everett DL, Fedorenko E, Gibson E. 2008 Number as a cognitive technology: evidence from Pirahã language and cognition. *Cognition* **108**, 819–824. (doi:10.1016/j.cognition.2008.04.007)

91. Rugani R, Regolin L, Vallortigara G. 2008 Discrimination of small numerosities in young chicks. *J. Exp. Psychol. Anim. Behav. Process.* **34**, 388–399. (doi:10.1037/0097-7403.34.3.388)

92. Siegler RS, Booth JL. 2004 Development of numerical estimation in young children. *Child Dev.* **75**, 428–444. (doi:10.1111/j.1467-8624.2004.00684.x)

93. Nieder A, Diester I, Tudusciuc O. 2006 Temporal and spatial enumeration processes in the primate parietal cortex. *Science* **313**, 1431–1435. (doi:10.1126/science.1130308)

94. Jordan KE, Maclean EL, Brannon EM. 2008 Monkeys match and tally quantities across senses. *Cognition* **108**, 617–625. (doi:10.1016/j.cognition.2008.05.006)

95. Nieder A. 2012 Supramodal numerosity selectivity of neurons in primate prefrontal and posterior parietal cortices. *Proc. Natl Acad. Sci. USA* **109**, 11 860–11 865. (doi:10.1073/pnas.1204580109)

96. MacKay DM. 1963 Psychophysics of perceived intensity: a theoretical basis for Fechner's and Stevens' laws. *Science* **139**, 1213–1216. (doi:10.1126/science.139.3560.1213-a)

97. Johnson KO, Hsiao SS, Yoshioka T. 2002 Neural coding and the basic law of psychophysics. *Neuroscientist* **8**, 111–121. (doi:10.1177/107385840200800207)

98. Veit L, Hartmann K, Nieder A. 2014 Neuronal correlates of visual working memory in the corvid endbrain. *J. Neurosci.* **34**, 7778–7786. (doi:10.1523/JNEUROSCI.0612-14.2014)

99. Ditz HM, Nieder A. 2016 Sensory and working memory representations of small and large numerosities in the crow endbrain. *J. Neurosci.* **36**, 12 044–12 052. (doi:10.1523/JNEUROSCI.1521-16.2016)

100. Nieder A, Merten K. 2007 A labeled-line code for small and large numerosities in the monkey prefrontal cortex. *J. Neurosci.* **27**, 5986–5993. (doi:10.1523/JNEUROSCI.1056-07.2007)

101. Jacob SN, Nieder A. 2009 Tuning to non-symbolic proportions in the human frontoparietal cortex. *Eur. J. Neurosci.* **30**, 1432–1442. (doi:10.1111/j.1460-9568.2009.06932.x)

102. Kersey AJ, Cantlon JF. 2017 Neural tuning to numerosity relates to perceptual tuning in 3-6-year-old children. *J. Neurosci.* **37**, 512–522. (doi:10.1523/JNEUROSCI.0065-16.2017)

103. Fechner GT. 1860 *Elemente der psychophysik* [*Elements of psychophysics*], vol. III (in German). Leipzig, Germany: Breitkopf and Härtel.

104. Buzsáki G, Mizuseki K. 2014 The log-dynamic brain: how skewed distributions affect network operations. *Nat. Rev. Neurosci.* **15**, 264–278. (doi:10.1038/nrn3687)

105. Dehaene S, Changeux JP. 1993 Development of elementary numerical abilities: a neuronal model. *J. Cogn. Neurosci.* **5**, 390–407. (doi:10.1162/jocn.1993.5.4.390)

106. Verguts T, Fias W. 2004 Representation of number in animals and humans: a neural model. *J. Cogn. Neurosci.* **16**, 1493–1504. (doi:10.1162/0898929042568497)

107. Wilson ML, Britton NF, Franks NR. 2002 Chimpanzees and the mathematics of battle. *Proc. R. Soc. Lond. B* **269**, 1107–1112. (doi:10.1098/rspb.2001.1926)

108. Moll FW, Nieder A. 2015 Cross-modal associative mnemonic signals in crow endbrain neurons. *Curr. Biol.* **25**, 2196–2201 (doi:10.1016/j.cub.2015.07.013)

109. Veit L, Pidpruzhnykova G, Nieder A. 2015 Associative learning rapidly establishes neuronal representations of upcoming behavioral choices in crows. *Proc. Natl Acad. Sci. USA* **112**, 15 208–15 213. (doi:10.1073/pnas.1509760112)

110. Danzig T. 1930 *Number—the language of science*. New York, NY: The Free Press.

111. Burr D, Ross J. 2008 A visual sense of number. *Curr. Biol.* **18**, 425–428. (doi:10.1016/j.cub.2008.02.052.)

112. Cicchini GM, Anobile G, Burr DC. 2016 Spontaneous perception of numerosity in humans. *Nat. Commun.* **7**, 12536. (doi:10.1038/ncomms12536)

113. 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**, 11 187–11 192 (doi:10.1073/pnas.1308141110)

114. Viswanathan P, Nieder A. 2015 Differential impact of behavioral relevance on quantity coding in primate frontal and parietal neurons. *Curr. Biol.* **25**, 1259–1269. (doi:10.1016/j.cub.2015.03.025)

115. Nieder A. 2016 Representing something out of nothing: the dawning of zero. *Trends Cogn. Sci.* **20**, 830–842. (doi:10.1016/j.tics.2016.08.008)

116. Bullock TH. 1984 Comparative neuroscience holds promise for quiet revolutions. *Science* **225**, 473–478. (doi:10.1126/science.6740319)