



Cite this article: Bryer MAH *et al.* 2021

The evolution of quantitative sensitivity. *Phil.*

Trans. R. Soc. B **377**: 20200529.

<https://doi.org/10.1098/rstb.2020.0529>

Received: 23 August 2021

Accepted: 28 October 2021

One contribution of 16 to a theme issue
'Systems neuroscience through the lens of
evolutionary theory'.

Subject Areas:

behaviour, cognition, neuroscience

Keywords:

brain evolution, Weber fraction,
quantity discrimination

Author for correspondence:

Jessica F. Cantlon

e-mail: jcantlon@andrew.cmu.edu

[†]These authors contributed equally to this study.

Electronic supplementary material is available online at <https://doi.org/10.6084/m9.figshare.c.5713120>.

The evolution of quantitative sensitivity

Margaret A. H. Bryer^{1,2,†}, Sarah E. Koopman^{3,†}, Jessica F. Cantlon¹, Steven T. Piantadosi², Evan L. MacLean^{4,5}, Joseph M. Baker⁶, Michael J. Beran⁷, Sarah M. Jones⁸, Kerry E. Jordan⁹, Salif Mahamane¹⁰, Andreas Nieder¹¹, Bonnie M. Perdue¹², Friederike Range¹³, Jeffrey R. Stevens¹⁴, Masaki Tomonaga¹⁵, Dorottya J. Ujfalussy^{16,17} and Jennifer Vonk¹⁸

¹Department of Psychology, Carnegie Mellon University, Pittsburgh, PA 15213, USA

²Department of Psychology, University of California-Berkeley, Berkeley, CA 94720, USA

³School of Psychology and Neuroscience, University of St. Andrews, St Andrews KY16 9AJ, UK

⁴School of Anthropology, and ⁵College of Veterinary Medicine, University of Arizona, Tucson, AZ 85719, USA

⁶Center for Interdisciplinary Brain Sciences Research, Division of Brain Sciences, Department of Psychiatry and Behavioral Sciences, School of Medicine, Stanford University, Stanford, CA 94305, USA

⁷Department of Psychology and Language Research Center, Georgia State University, Atlanta, GA 30302, USA

⁸Psychology Program, Berea College, Berea, KY 40403, USA

⁹Department of Psychology, Utah State University, Logan, UT 84322, USA

¹⁰Behavioral and Social Sciences Department, Western Colorado University, Gunnison, CO 81231, USA

¹¹Animal Physiology Unit, Institute of Neurobiology, University of Tübingen, Tübingen 72076, Germany

¹²Department of Psychology, Agnes Scott College, Decatur, GA 30030, USA

¹³Domestication Lab, Konrad Lorenz Institute of Ethology, Department of Interdisciplinary Life Sciences, University of Veterinary Medicine Vienna, Savoyenstrasse 1a, Vienna 1160, Austria

¹⁴Department of Psychology and Center for Brain, Biology and Behavior, University of Nebraska-Lincoln, Lincoln, NE 68588, USA

¹⁵Inuyama, Aichi 484-0000, Japan

¹⁶MTA-ELTE Comparative Ethology Research Group, and ¹⁷Department of Ethology, Eötvös Loránd University of Sciences (ELTE), Budapest 1117, Hungary

¹⁸Department of Psychology, Oakland University, Rochester, MI 48309, USA

id MAHB, 0000-0002-3088-6936; SEK, 0000-0001-8116-9913; JFC, 0000-0003-0207-5615; STP, 0000-0001-5499-4168; ELM, 0000-0001-7595-662X; MJB, 0000-0002-2504-0338; AN, 0000-0001-6381-0375; FR, 0000-0003-3127-5536; JRS, 0000-0003-2375-1360; MT, 0000-0002-9319-6991; DJU, 0000-0001-5363-753X; JV, 0000-0002-4362-5989

The ability to represent approximate quantities appears to be phylogenetically widespread, but the selective pressures and proximate mechanisms favouring this ability remain unknown. We analysed quantity discrimination data from 672 subjects across 33 bird and mammal species, using a novel Bayesian model that combined phylogenetic regression with a model of number psychophysics and random effect components. This allowed us to combine data from 49 studies and calculate the Weber fraction (a measure of quantity representation precision) for each species. We then examined which cognitive, socioecological and biological factors were related to variance in Weber fraction. We found contributions of phylogeny to quantity discrimination performance across taxa. Of the neural, socioecological and general cognitive factors we tested, cortical neuron density and domain-general cognition were the strongest predictors of Weber fraction, controlling for phylogeny. Our study is a new demonstration of evolutionary constraints on cognition, as well as of a relation between species-specific neuron density and a particular cognitive ability.

This article is part of the theme issue 'Systems neuroscience through the lens of evolutionary theory'.

1. Introduction

Quantitative sensitivity is an aspect of cognition that is ubiquitous among many species, and many researchers debate the nature of its evolutionary basis across

taxa, including in humans and other primates [1–6]. Baboons use numerical estimation to guide troop movement [7,8], desert ants and fiddler crabs navigate by keeping track of the number of steps they have taken [9,10], and social species like hyenas and lions vocalize or approach other conspecific groups only when their group has a numerical advantage [11–17]. A diverse range of animals—from primates to reptiles, fish and insects—can discriminate numerical quantities in laboratory tasks, for example, comparing computerized arrays or sequences of pure tones to peck, touch or approach the numerically larger set [2,4,18–22]. Moreover, animals represent numerical values cross-modally [23–27] and under conditions where dimensions such as area, density and duration are equated, uncorrelated with numerical value or otherwise controlled [2,4,20,28–30].

Behaviour in quantitative tasks is well predicted by the general psychophysical principle Weber's law: discrimination accuracy between the magnitudes of two stimuli is determined by their proportional difference [31,32]. Weber's law reliably predicts performance on nonsymbolic numerical tasks across species [7,20,29,33], through human development [34–36], and across cultures [37]. Weber fraction (w) is a measure of quantitative precision that represents the proportion difference between quantities that is needed to reliably discriminate them [36,38,39]. In an investigation of the basis of quantitative decision-making by wild baboons, number better predicted decisions than mass, with wild baboons exhibiting Weber fractions similar to those found for baboons in laboratory settings [7]. Insect number sense has been explored [40], especially in honeybees [41–44], with mixed results on relative versus absolute quantity discrimination that suggest different mechanisms may be at play than in mammals and birds. Bees do not show particularly high quantitative sensitivity; for example, with appetitive conditioning only, bees discriminated between 4 and 8 almost at chance level (54%), while appetitive and aversive conditioning improved discrimination only slightly at this 2:1 discrimination ratio (64%) [45].

The ubiquity of quantitative representation, and specifically numerical representation, suggests that number is useful in many animals' natural lives, but its phylogenetic basis is unclear. There are no prior studies that examine quantitative cognition in an explicitly phylogenetic framework across a broad range of species. Investigating species differences has long been one goal of comparative psychology, with integration of principles of evolutionary biology [46–48]. However, modern methods of phylogenetic regression across taxa are novel in the context of cognitive studies [49,50] and can provide insight into the phylogenetic signal or ancestral state of a cognitive trait, as well as if social, ecological or anatomical variables affect cognition when phylogeny is taken into account [48].

Here we develop a computational model of the phylogenetic distribution of quantitative sensitivity, and we examine potential causes of changes in quantitative sensitivity including brain morphology, ecology and domain-general cognition. This model is innovative in that it combines a phylogenetic regression—to control for relatedness between species—with a likelihood model based on the psychophysics of number. Specifically, we assume that performance on number tasks is derived from a linear internal mental scale with linearly increasing Gaussian noise. We model variation in species, individual animal and task, and combine these with a predictor. We

model phylogenetic relatedness as a correlation matrix among species effects, where a single parameter determines the degree to which species effects are determined by phylogeny. Together in a Bayesian data analysis, this allows us to test the likely influence of a single factor (e.g. an aspect of brain morphology or ecology), while controlling these other components. Thus, likely underlying causal factors can be inferred by exploring the phylogenetic variation in numerical ability across species and investigating which physiological and ecological factors best predict numerical competence—an approach used previously to investigate the evolution of self-control [49], domain-general cognitive ability [51], visual perspective taking [52], prosocial preferences [53], cultural intelligence [54] and temporal decision-making [55]. Identification of these factors may reveal the selection pressures that have shaped numerical understanding throughout evolution. The potential factors we investigate include brain size and neuronal number and density, socioecology and domain-general cognitive ability, detailed below.

(a) Features of the brain

One potential factor influencing the precision of quantitative representations is brain size, as it has been proposed that humans' exceptional cognitive abilities arise from our large brains [56–58]. Larger brains, despite their greater metabolic cost [59], provide the neural tissue necessary to solve more complex problems that an animal may encounter in its environment. Indeed, prior research has demonstrated a relation between some cognitive abilities and measures of brain size. Self-control and general cognitive ability (an aggregate score of invisible displacement, reversal learning, delayed response and other cognitive abilities) have been found to be strongly related to absolute brain size, supporting a hypothesis of cortical reorganization driving the evolution of cognitive abilities [49,51]. When brains increase in size, cortical reorganization becomes necessary as neurons in different parts of the brain are pushed apart. This reorganization can result in larger brains showing more modular organization than smaller brains [60,61]. Measures of relative brain size, which take body size into account, cannot capture the degree of cortical reorganization that may have occurred due to increases in absolute brain size. On the other hand, bigger bodies require more sensory and motor brain tissue; relative brain size accounts for body size and is thus a measure of brain size above and beyond that needed for sensory and motor control. Social learning, behavioural flexibility, tool use, innovation and problem-solving ability have been found to be correlated with relative brain size, supporting a hypothesis that encephalization (an increase in cortex size relative to body size) drives cognitive evolution [62–66]. Though the avian brain would appear to be a special case because it does not have a cortex, the nidopallium caudolaterale (NCL) has been identified as functionally equivalent to the mammalian prefrontal cortex [67,68].

Criticisms of using brain size measurements in comparative biology include that these measures do not actually capture variation in brain size across taxa well [69,70]. Some researchers contend that the number of cortical neurons, rather than brain size, is a better proxy for cognitive ability [71]. Recent advancements have made the quantification of neurons more accurate and efficient [72,73], leading to the discovery, spearheaded by Herculano-

Houzel, that number of neurons does not scale with brain size the same way in different taxa [74–81]. For example, as brain size increases, the number of neurons in the primate brain increases more than the number of neurons in the rodent brain, meaning that similarly sized primate and rodent brains contain different numbers of neurons [82–84]. The number or density of cortical/pallium neurons may, therefore, be a more accurate proxy for quantitative ability than brain volume, particularly when comparing species across a range of taxonomic groups.

Ratio-dependent neural responses to numerical stimuli have been demonstrated in the NCL of corvids [33,85], comparable to those found in the intraparietal sulcus area of the cortex in humans [86,87] and nonhuman primates [27]. Given the vast differences in avian and mammalian brain structure, it may also be fruitful to explore other avian brain regions with a comparatively high neuronal density as potential proxies for cognitive ability. Cerebellar neuronal scaling rules also differ across taxonomic groups [84], with songbird and parrot cerebellum neuronal densities (as well as pallium neuronal densities) being higher than those of mammals [79]. Mammals from four orders showed the same coordinated scaling of neuron number in both the cerebral and cerebellar cortices, supporting arguments for integrated functions and selective pressures [88]. Indeed, some researchers argue that the cerebellum has been neglected in studies of vertebrate brain evolution and cognition [89–91].

(b) Socioecology

Socioecological selective pressures could also influence quantitative ability directly since both dietary and social complexity have been previously linked to differences in cognition. Diet is related to spatial memory in seed-caching birds [92–94], Central and South American primates [95] and lemurs [96]. Bonobos, whose food source is more uniformly distributed in their habitat and who arguably rely on a relatively consistent food source, are more risk averse than chimpanzees, whose food source is more patchily distributed [97]. This and other differences in ecology between bonobos and chimpanzees [98,99] have implications for the cognitive profile of the last common ancestor between humans and *Pan* [100]. More broadly, the foraging required to maintain a particular diet has been hypothesized as one of the driving forces in the evolution of primate cognition [101]. Successful foragers of diverse taxa accurately decide when to keep foraging in the same area and when to move to a new area [102]. This likely requires a finely tuned representation of quantity and the ability to accurately compare multiple quantities. Therefore, dietary factors such as the percentage of fruit in the diet, which reflects the diversity of foraging done by a species, may be related to the precision of species' quantitative representations. A larger home range size and/or a longer day journey length involves more complex navigation of heterogeneous environments on a larger scale and may require more quantitative sensitivity in making ecological decisions with high cost or high gain outcomes in terms of energy acquisition or expenditure.

Navigating complex social systems influences multiple components of cognition in social vertebrate taxa. The social intelligence (or social brain) hypothesis suggests that social complexity drove brain evolution, especially in primates [103–105]. Survival in large social groups requires the ability to keep track of multiple individuals and their relationships,

as well as judge whether to engage with another group or not. When baboons disagree about the direction of troop movement, undecided baboons more often choose to go in the direction with more supporters [8], behaviour which is driven by approximate number representations [7]. Additionally, in species that engage in reciprocal altruism, having precise quantity representations could help them keep track of the quantities being exchanged and thus avoid cheaters [106]. In intergroup encounters, assessment of the number of opponents is advantageous across multiple taxa [12,15,16]. Within-group social decision-making, in the context of navigation, predator avoidance, cooperative hunting and mating behaviours also involves quantitative processing [5]. Being able to make accurate quantity judgements, due to a well-tuned quantity representation system, would likely be adaptive for a highly social species. Social group factors like group size could, therefore, be related to numerical precision.

(c) Domain-general cognitive ability

Finally, domain-general cognitive abilities could influence performance on the quantity discrimination tasks used to measure numerical precision. These tasks could place demands on executive functions like attention, memory and self-control that, while distinct from the number sense itself, influence an animal's ability to discriminate between quantities. For example, species with low self-control could perform poorly on quantity discrimination tasks because they are less able to inhibit incorrect responses.

Our goal was to evaluate these factors quantitatively in mammals and birds ($N = 33$ species) using a novel modelling approach that integrates Bayesian, phylogenetic and number psychophysics components. Specifically, the model jointly fits effects of predictors across species on Weber fraction (w), as well as the influence of phylogenetic relatedness, while controlling for species, individual, study and task effects. With the wealth of quantitative discrimination studies across taxa, we implemented this model in the context of a meta-analysis.

2. Methods

(a) Animal quantity discrimination data compilation

To find studies from which Weber fractions could be calculated, Google Scholar literature searches were conducted of 'animal' and 'quantity discrimination' in March 2015 and May 2017. Studies were included in these analyses if they used one of these three paradigms: match-to-sample, go/no-go or ordinal tasks with visual array stimuli (controlled array paradigm); food stimuli tasks with sequential presentation of food (sequential paradigm); and food stimuli tasks with simultaneous presentation of food (simultaneous paradigm). Data were obtained either from the primary literature or, when possible, directly from the author(s). In total, data were obtained from 49 studies for 33 species and 672 subjects. See electronic supplementary material, table S1 for a complete list of the studies included in these analyses.

Number representation precision was quantified by the Weber fraction (w), a value that quantifies the representational precision, assuming the underlying approximate number system is formalized as Gaussian distributions along a mental number line [107]. Smaller Weber fractions correspond to narrower Gaussian curves and more precise quantity representations, while larger Weber fractions reflect increased noise and overlap between Gaussian curves of numbers being compared and poorer quantity discrimination. In our model, Weber fraction was jointly estimated with

other parameter effects (species, study, subject, task) per species in a Bayesian model implemented in the Stan programming language (see details under Bayesian analyses section).

(b) Phylogenetic tree construction

Phylogenetic trees were generated in R using the ape package [108], based on information from The Timetree of Life [109,110], incorporating a 3 Mya divergence of the North Island robin and South Island robin [111], and a 15 000-year-ago divergence of dogs and wolves [112].

(c) Brain and socioecological data

Most endocranial volume (ECV) and socioecological variable data were acquired from a previous large-scale comparative study [49] (see electronic supplementary material, table S1 for primary sources). Additional ECV and socioecological data were obtained from the literature. When ECV could not be found, it was calculated from brain mass based on brain tissue density of 1.036 g cm^{-3} [113]. Residual brain volume (RBV) is the residuals from a phylogenetic regression of ECV on body mass [49]. Actual cortical/pallium neuron data and cerebellum neuron data for 12 species were obtained from the literature (see electronic supplementary material, table S1 for primary sources). For cerebellum neuron number, an outlier, the African elephant, is included, as models run with the African elephant excluded for cerebellum neuron number and density did not show a major shift in effect (see electronic supplementary material, figure S2 for scatterplot of w by cerebellum neuron number with and without the elephant outlier). For the primate species without cortical neuron data, the number of cortical neurons was calculated from brain mass using the following scaling rule: no. of cortex neurons = $37\,813\,551.018 \times \text{brain mass}^{0.891}$ [114]. For the primate species without cerebellum neuron number data, the number of cerebellar neurons was calculated from brain mass using the following scaling rule: no. of cerebellum neurons = $69\,640\,042.656 \times \text{brain mass}^{0.936}$ [114]. Calculating cortical or cerebellar neuronal density values for species with missing values was less straightforward because neuronal density and neuronal number are decoupled in primates, i.e. unlike in nonprimate mammals, more neurons does *not* mean larger neurons (and lower density) in primates [75,115]. Therefore, our sample size for neuronal density is smaller than that for neuronal number.

Socioecological variable data, specifically group size, percentage fruit in diet, home range and day journey length, were compiled from the literature (see electronic supplementary material, table S1). Group size was compiled across all taxa, while percentage fruit in the diet, home range and day journey length were compiled for primate species only. Finding comparable social and ecological variable data across primate, nonprimate mammal and avian taxa is not straightforward: for example, finding values for percentage frugivory for corvids is complicated by lack of reliance on fruit and corvid diverse diets with variable and sometimes qualitative mention of fruit-eating [116,117]. For bird taxa generally, other social and ecological variables not applicable to mammals may be more relevant selective pressures; for example, degree of food caching.

Some of the variables were transformed in order to meet the assumptions of the statistical analyses performed: ECV, group size, home range, day journey, cortical neuron number, cerebellar neuron number, cortical neuron density and cerebellar neuron density were log-transformed.

(d) Domain-general cognitive ability data

Self-control data were obtained for 17 species from a large-scale comparative study (species score was composite score calculated

from performance on A-not-B task and cylinder task [49] and for two species from studies of corvids [118] and parrots [119] (species score was performance on cylinder task). Domain-general cognitive ability data were obtained for 10 primate genera from a meta-analysis [120]. In the meta-analysis study [120], a high general cognitive ability score indicated poor performance, so we reversed the scores in our analyses for clarity (so that high score indicated good performance).

(e) Bayesian analyses

We used a Bayesian model that combines a phylogenetic regression with a likelihood that is grounded in number psychophysics. Specifically, the regression assumes that there was a linear effect of each single predictor (outlined above) on the mean log Weber fraction for a species. The phylogenetic component includes a correlation that derives from the phylogenetic tree, but the strength of this correlation is determined by a free parameter, lambda (λ) [48,121–123]. Finally, the inferred Weber fractions w are converted into a binary prediction by following the common linear-scale-variable psychophysical model which assumes a number n is represented with standard deviation wn . This model predicts binary accuracy judgements on specific numerical comparisons, which is the observed data for the model. In addition, the model includes parameters to capture variation in species, tasks and individuals via a simple mixed-effect component. Thus, the model jointly fits effects of predictors across species on w , as well as the influence of phylogenetic relatedness employing a mixed-effect component similar to previous phylogenetic approaches [124], to control for individuals, study, species and task. Task refers to experimental task (i.e. paradigm), while study indicates a specific research study. A graphical model for our analysis is shown in figure 1. This shows the relationships between the key variables, with plates (rectangles) illustrating arrays of variables. For example, lambda, a parameter that controls the strength of the phylogenetic component, determines the species effects and is constant across all species, while x , the predictor of interest, takes on a different value for each species and, along with the regression coefficients β_0 and β_1 , determines the mean value μ of each Weber fraction. This model importantly allows us to quantify the effect of each predictor under tight controls for these other factors, as well as a principled likelihood that is based on a formalization of approximate number perception.

Our model differs from standard phylogenetic regressions in that the quantity we measure, Weber fraction, is not directly observed. It is a latent variable that predicts accuracy only via a particular psychophysical linking function. We have a fixed covariance matrix V based on one phylogenetic tree which indicates how correlated species should be based on their phylogeny. That matrix V is mixed with the identity matrix via lambda (e.g. $M = \lambda \times V + (1 - \lambda) \times \text{identity}$) to create the inferred covariance matrix M among species. If $\lambda = 0$, each species effect is independent (sampled from standard normal), while if $\lambda = 1$, each species' adjustment to the overall Weber fraction is determined by the phylogeny's predicted covariance. In the model, lambda is fitted simultaneously with all other parameters (only the matrix V is fixed). The species-level adjustments to log Weber fraction are sampled from a normal with covariance M . In the model, these species effects are added to the grand overall (log) Weber fraction, effects for each task, individual and study. And then the (log) Weber fraction thus derived is used to predict the correct/incorrect outcomes on each number comparison using a standard psychophysical model. The only data we observe are V and the accuracies, while the effective Weber fraction for each trial is latent, as is the species effects on this Weber fraction.

legend		
component of model		explanation or examples
β_0	intercept	intercept of regression
β_1	coefficient	coefficient of regression
x	predictor	examples: cortical neuron number, frugivory, self-control
μ	species mean	
λ	phylogenetic signal	phylogenetic component of regression representing evolution by Brownian motion, i.e. genetic drift
species effect	species of animal subject	example: random effect caused by 'being a jackdaw'
study effect	research study	example: Cox and Montrose's (2016) study of quantity discrimination in rats
subject effect	individual research subject	example: an individual long-tailed macaque housed at the German Primate Center tested in Schmitt and Fischer's (2011) study
task effect	experimental paradigm	simultaneous, sequential, controlled array
w	Weber fraction	a latent variable that predicts accuracy via a psychophysical linking function

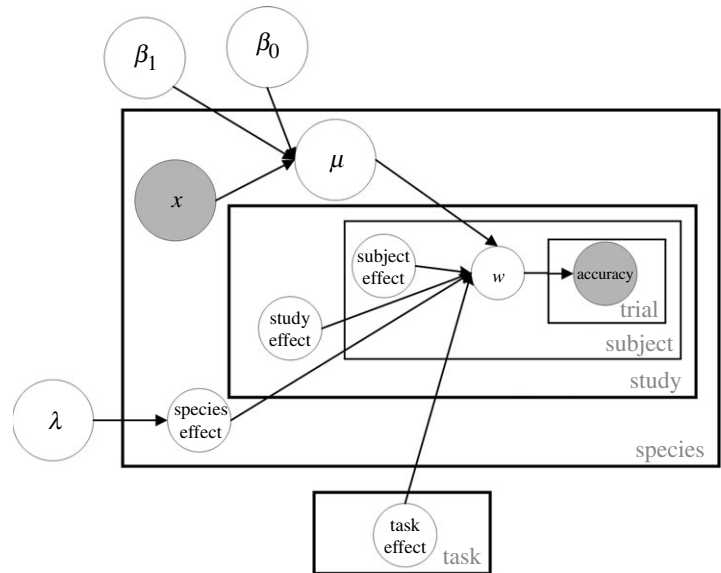


Figure 1. Graphical diagram of the Bayesian model implemented in this study. The model jointly fits effects of predictors (x) across species on w , as well as the influence of phylogenetic relatedness (λ), while controlling for species, subject, study and task effects. The inferred Weber fractions are converted into a binary prediction by following the common linear-scale-variable psychophysical model that predicts binary accuracy judgements on specific numerical comparisons, which are the observed data for the model.

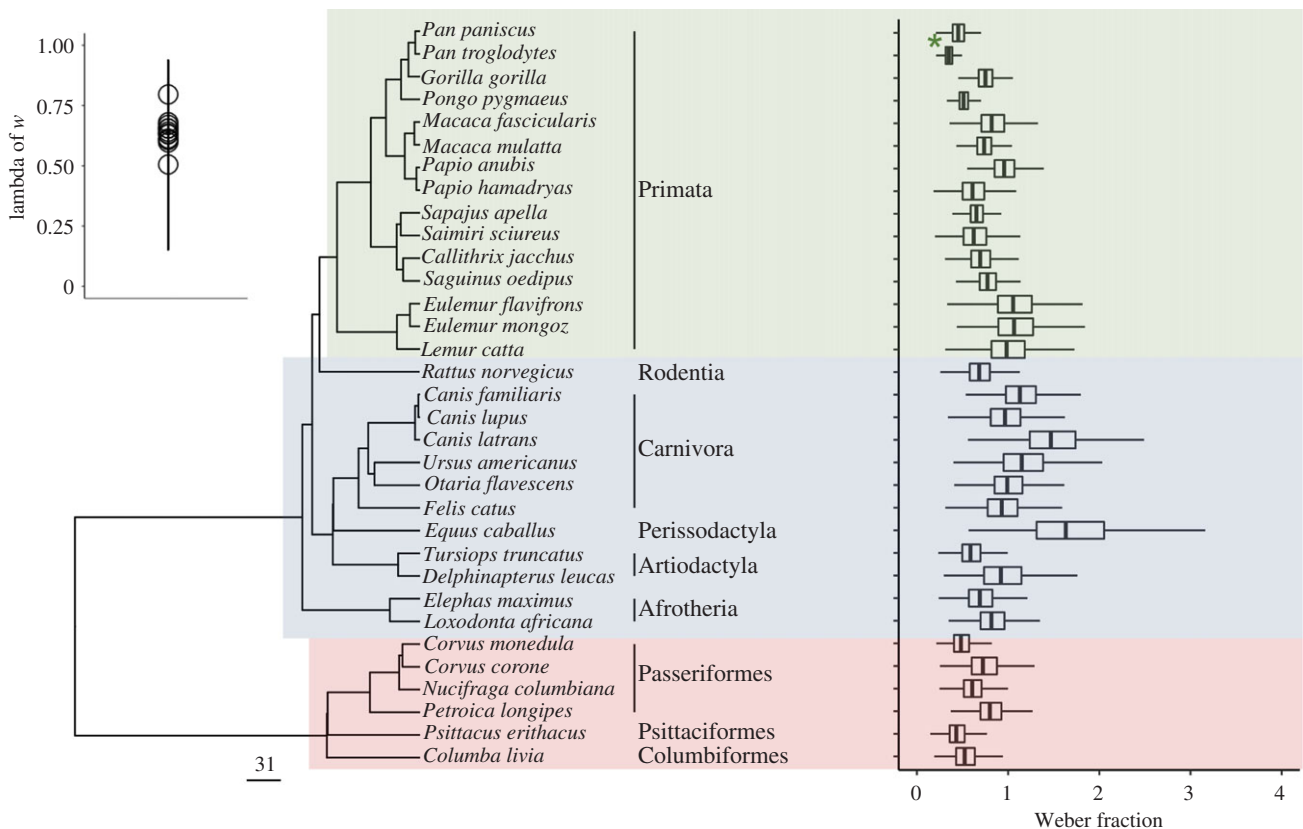


Figure 2. Phylogenetic tree of the species included in these analyses and the overall Weber fraction for each species (posterior quantiles shown, generated via a model with no predictors and only species, study, subject and task effects). These Weber fractions are based on quantity discrimination performance data from 672 individual animals across 33 species. Scale bar represents 31 Myr. Primates are highlighted in green, nonprimate mammals in blue and birds in red. Human Weber fraction (from [127]) indicated with green asterisk for comparison. The inset shows the lambda of w across all predictors with mean minimum and maximum of posterior distributions for lambda shown. (Online version in colour.)

We implemented the model in the Stan programming language (Stan v. 2.19.1, using the R interface rstan [125]), which employs Hamiltonian Monte Carlo, specifically the no-U-turn sampling (NUTS) algorithm [126]. Rhat values, meaning the agreement of multiple chains, were used to assess convergence; we found all Rhat values < 1.05 (indicating that between-chain variance was similar to within-chain variance).

3. Results

Figure 2 shows a phylogenetic tree of the species included in these analyses and overall Weber fraction, with uncertainty represented, for each species. Larger Weber fractions reflect worse performance—a Weber fraction of 0.2 would mean an animal can reliably discriminate quantities as fine as 10 versus 12

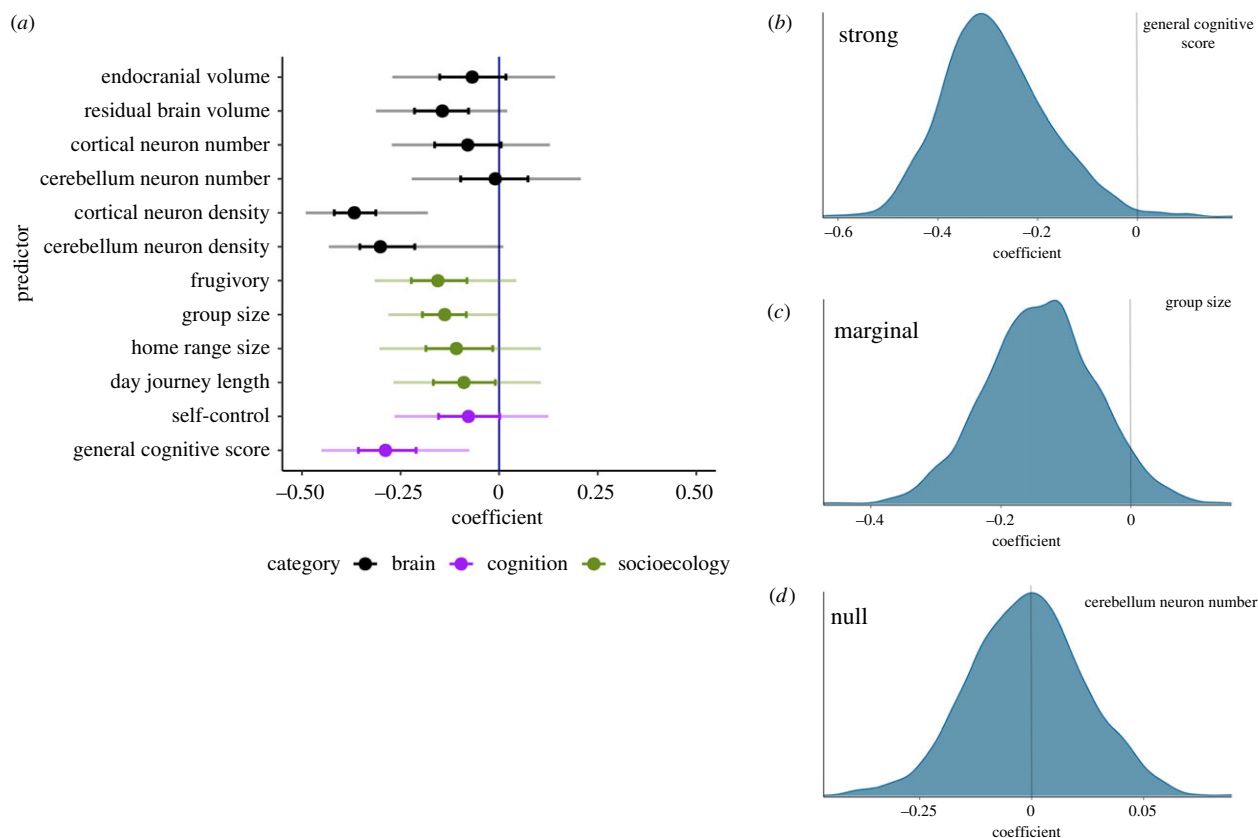


Figure 3. (a) Posterior distribution for the coefficient for each predictor from a model run with 33 species. 95% credible interval represents the central 95% of the posterior distribution (there is a 95% probability that the parameter lies within the 95% credible interval), while the 50% credible interval represents the central 50% of the posterior distribution. A parameter with no effect would be centred at 0 (as we found for cerebellum neuron number) with wide symmetrical error bars. To further characterize the distribution of posterior samples, posterior kernel density plots of a (b) strong effect (general cognitive score), (c) marginal effect (group size) and (d) null effect (cerebellum neuron number) are presented. Kernel density plots show posterior samples with chains merged.

whereas a Weber fraction of 0.5 affords reliable discrimination only to 10 versus 15. The ranges in Weber fractions in each of the three groups broadly agree with previous studies in primates [7,128–130], other mammals and birds [131]. Birds (in red) have Weber fractions that are lower and narrower posterior distributions than those of nonprimate mammals (blue) and some primates (green). Among nonprimate mammals, the horse and the coyote show highest estimated Weber fractions with the widest posterior distributions. Among primates, smaller Weber fractions with narrow posterior distributions are found for some Central and South American primates, and some Afro-Eurasian primates, including most apes. The human Weber fraction [127] is shown with a green asterisk for comparison with the other species. The inset in figure 2 shows lambda of w across all predictors (controlled), indicating the existence of a phylogenetic correlation for w across brain, socioecological and cognitive predictors, with uncertainty. Lambda indicates whether the species-level adjustments to Weber fraction are correlated in the way the phylogeny would predict or not; therefore, the inset in figure 2 provides evidence that $\lambda > 0$, which means that species are not independent. Subsequent results reported below take into account this uncertainty in phylogenetic correlation for w .

In figure 3, the intervals illustrate the posterior distribution for the coefficient for each predictor. Posterior distributions in Bayesian analysis quantify what we should believe about the true value of the parameters, i.e. what we normatively should believe about the size of each effect. Specifically, figure 3 shows posterior uncertainty intervals

that indicate the means that give the best estimate with a spread that shows how confident to be in that estimate. In our model, parameters are chosen so that their value quantifies a theoretically driven hypothesis, such as the likely influence of a predictor on numerical acuity. We plot posterior quantiles, which provide the range of likely true values for the parameters (e.g. the 50% quantiles show the most central half of probability, or belief, in the parameter's value). Bayesian analysis does not typically ask whether variables are exactly equal to zero—as in a null hypothesis test, or standard frequentist regression—because that has zero probability of being true in the world. Instead, these ranges illustrate the model's best guess (value) and confidence (range) in each effect. In figure 3, from a model with 33 species, the posterior distributions demonstrate that w decreases (quantity discrimination improves) with increased cortical neuron density, cerebellar neuron density and general cognitive score; weaker effects are also shown for RBV, frugivory and group size. A parameter with no effect would be centred at 0 (as cerebellum neuron number is) with wide symmetrical error bars. These findings indicate that a subset of neural and general cognitive variables are related to the evolution of quantitative sensitivity among species. A model with only primate species ($n = 15$) showed some effects, including cerebellar neuron density and self-control, were weakened or removed; cerebellar neuron density shifted to no effect (centred at zero) (electronic supplementary material, figure S1). Two effects were strengthened when only primates were included in the model: primate species

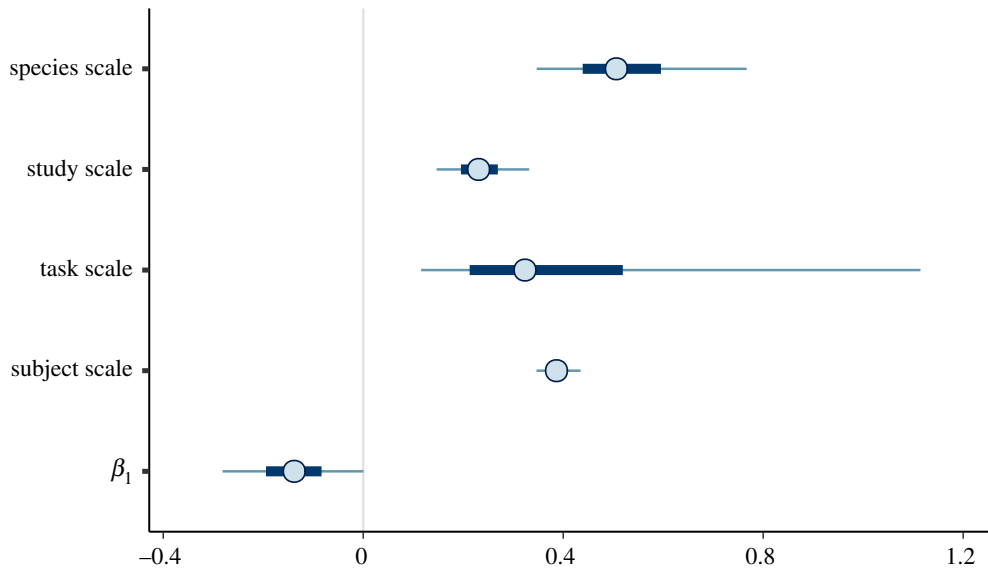


Figure 4. Posterior credible intervals (circles, posterior means; thick segments, 50% intervals; thin lines, 90% intervals) for the scale parameters and one predictor parameter chosen as an example (group size) run with 33 species. Each of these scales is a factor that multiplies the predictor parameter (e.g. for species, study, task, subject), and β_1 multiplies the predictor (e.g. group size), making them comparable in indicating how much change in $\log w$ there is for each change in the predictor.

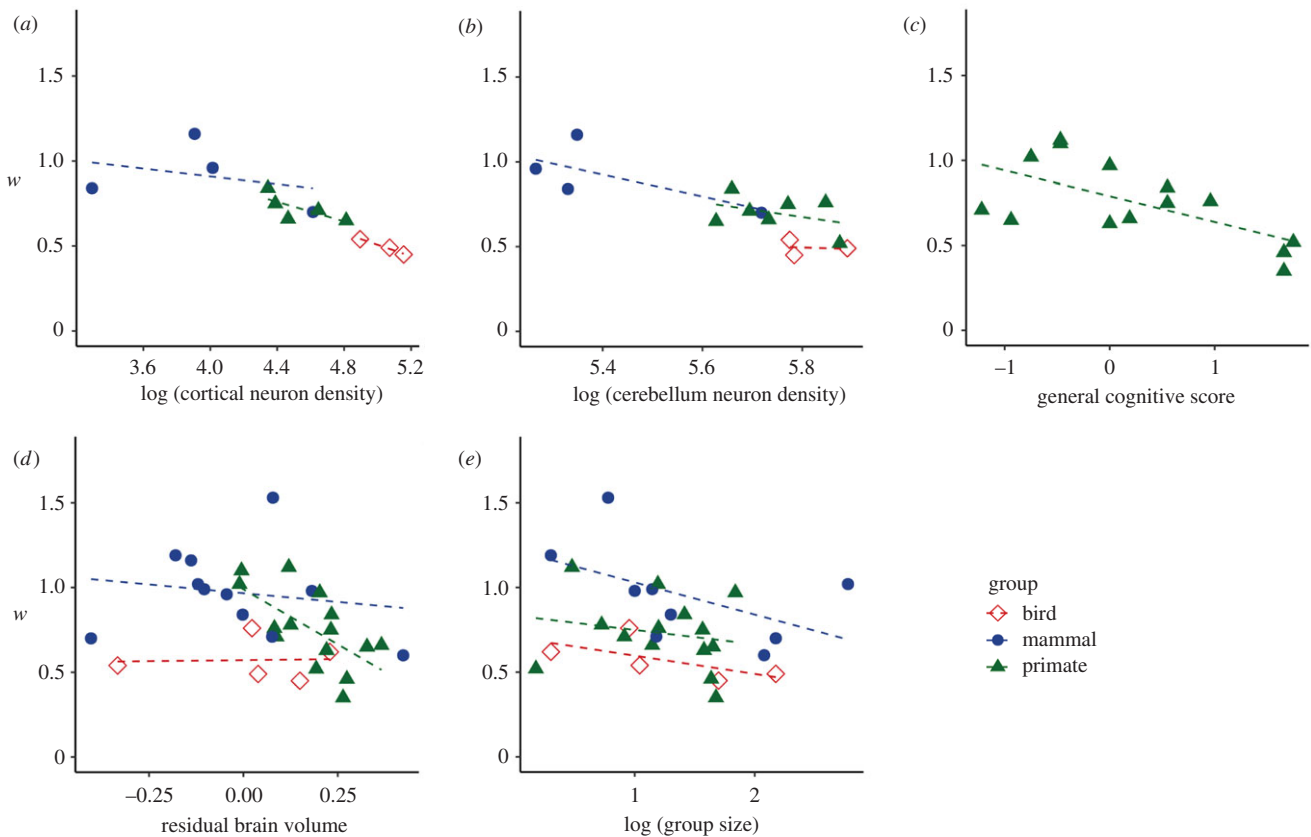


Figure 5. Scatterplots show species w (generated via a model with no predictors and only species, study, subject and task effects) by (a) cortical neuron density, (b) cerebellum neuron density, (c) general cognitive score, (d) residual brain volume and (e) group size. Mammals are blue circles, primates are green triangles and birds are red diamonds. (Online version in colour.)

w decreases (quantity discrimination improves) with increased RBV and increased cerebellum neuron number (electronic supplementary material, figure S1).

Figure 4 shows the posterior credible intervals on the scale parameters for species, subject, study and task, as well as the posterior credible interval of one of the predictors (group size). Figure 4 indicates that species and individual subject contribute to performance. Task, with wide uncertainty, also contributes to performance.

Scatterplots shown in figure 5 allow us to visualize each predictor by effectively estimating w independently and examining its relationship with key predictors, by taxonomic group. The species w values in figure 5 were generated via a model with no predictors and only species, study, subject and task effects (i.e. the same model as for figure 1). The cortical neuron density (figure 5a) effect is driven by primates and birds, while the cerebellum neuron density effect is driven by nonprimate mammals (figure 5b). Better (lower) w was

related to better general cognitive score in primates (figure 5c). For weaker effects, the RBV effect is driven by primates (figure 5d) and the group size effect is driven by all three taxonomic groups (figure 5e). The largest group size (with a relatively high w) belongs to the South American sea lion.

We acknowledge that w for different tasks may be related to predictors differently: for example, self-control may specifically be related to the sequential task w . In the sequential numerical cognition task, animals wait while a set of items is baited one-by-one over time, and this is a general task demand on attention and response control (e.g. [129]). Testing these predictions about task differences is outside the scope of our model because by assumption in the model, each task scales the value of w the same across all other predictors. Future work with different modelling approaches should address these questions as different quantitative tasks likely have unique relations to some predictors—particularly if the task requirements have consistent effects on performance among individuals of a species.

4. Discussion

The question of how strongly phylogeny predicts variation in cognition across species has a contentious intellectual history. For nearly a century, many psychologists considered evolution to be irrelevant to the study of learning because the behaviours of all species were thought to obey similar laws [132,133]. Ultimately, however, it was shown that the laws of learning often depend on the biological preparedness of the organism [134]. Comparative psychologists have since made inferences about species-specific biological influences and possible evolutionary relations among cognitive faculties but often relied on few species for comparative study [46]. By contrast, modern phylogenetic modelling methods applied across hundreds of individuals from dozens of species provide a new empirical test of phylogenetic constraints on cognition [48]. This is the first study to measure the origins of quantitative cognition with these methods. We found contributions of phylogeny to quantity discrimination performance across taxa, indicating evolutionary constraints on quantitative cognition. Additionally, a subset of neuronal and cognitive variables predicted species' quantitative sensitivity—the strongest predictors were neuron density and general cognitive ability. The results indicate that when selecting an animal from the world at random, we can roughly predict its Weber fraction by knowing its species.

An individual's Weber fraction was related to its species-typical cortical neuron density. Individuals from species with higher cortical neuron density had more precise Weber fractions. Thus, one constraint on an individual's quantitative cognition is the biological capacity for information processing in their brain, as determined genetically and developmentally for each species [74–81]. Additionally, quantitative precision was related to neuron density in the cerebellum, a brain structure that has been overlooked in studies of vertebrate brain evolution and cognition [89–91]. We found that density of neurons was a more accurate proxy for quantitative ability than brain volume when comparing species across taxa. Caveats to the interpretation of the neuron density findings include (i) the number of species with cortical and cerebellar neuron density values is lower than for neuronal number, and (ii) the relationship between neuron number and neuron density differs

across animal groups. An increase in number of neurons in a primate brain structure does *not* mean larger neurons (and lower density), whereas in most nonprimate mammals more neurons means larger neurons and lower density [75,115]. Our study is a rare demonstration of a relation between neuron number or neuron density and a particular cognitive ability. Previous within-species comparisons showed that neuron number in multiple brain regions did not predict performance on a battery of behavioural tasks in mice [135], and though raccoons who performed best on a puzzle box task had more cells in their hippocampus than lower performing individuals, this difference may have been driven by glial cells [136]. However, cross-species comparisons in primates and birds suggest that neuron number has more behavioural explanatory power than cranial capacity, based on the correlation between cortical or pallial neuron number and performance on a self-control task [71]. Our cross-species finding from birds and mammals implies that quantitative sensitivity is yoked to species-specific developmental programmes for neuronal density; therefore, some species are well-equipped to develop precise quantitative sensitivity whereas others may be unable to do so.

Our finding that primate species' quantitative sensitivity improved with their domain-general cognition score indicates that general cognitive functions, perhaps in tandem with specialized quantitative functions, impacted the evolution of quantitative precision across species. In the original domain-general cognition study [120], primates' performance on tasks like object discrimination, reversal learning, sorting, oddity learning and delayed response reliably approximated general cognitive ability at the genus level in a Bayesian latent variable analysis. This 'general cognitive ability' variable could represent a constellation of cognitive capacities that distinguishes species, or it could represent attention or memory functions that affect animals' overall task performance [105]. The current study expands the suite of tasks that relate to general cognitive ability to include quantitative sensitivity, which provides further support for the hypothesis that some species perform better than others across tasks independently of task content. These conclusions about general cognitive factors are limited to primates and it is an open question whether these patterns generalize to other mammals and birds. However, it is remarkable that these general cognitive scores may mediate differences in quantitative cognition between species as closely related as, for example, an olive baboon and a capuchin monkey.

It is important to highlight that these anatomical, ecological and cognitive factors are likely interrelated [51,137,138], meaning that some factors may play a role in mediating others. For instance, the influence of foraging pressures may be difficult to separate independently from group size. However, in general, separating out the independent *causal* contributions of correlated components is difficult, even more so in analyses like these where all variables were not available for all species. Our analysis, therefore, focused on evaluating each factor in a simplified framework where each is tested independently. Additional data and methods are needed to disentangle how social, ecological and brain selection pressures are intertwined in affecting cognition [139] and for exploration of mosaic evolution of cognition, in which social and/or ecological variables shape particular domains differently [100,140]. Fully formalized theories of how such causal factors jointly interact may permit computational modelling of further relationships in the future.

Our novel analysis shows that biological features of a species' evolutionary history likely modulate the development of individuals' numerical cognition, a crucial finding that emphasizes the importance of phylogenetic constraints on cognition. Natural selection has biologically prepared some species to develop high neuronal densities and general cognitive capacities that yield precise quantitative representations. These data begin to reveal the evolutionary pressures that shaped numerical cognition across species and bring us closer to understanding the evolutionary precursors that sparked human mathematical cognition.

Data accessibility. Data files, Stan models and R scripts used in these analyses are available via the Open Science Framework (OSF) at:

<https://osf.io/j9sv8/> (doi:10.17605/OSF.IO/J9SV8) and via Carnegie Mellon University's repository KiltHub, doi:10.1184/R1/17144432.

Authors' contributions. J.F.C., E.L.M., S.T.P. and S.E.K. conceived the study. S.E.K. and M.A.H.B. compiled data. J.F.C., J.M.B., M.J.B., S.M.J., K.E.J., S.M., A.N., B.M.P., F.R., J.R.S., M.T., D.J.U. and J.V. contributed data. S.T.P. and S.E.K. wrote code. S.E.K., M.A.H.B. and S.T.P. analysed data. S.E.K., J.F.C., M.A.H.B. and S.T.P. prepared the manuscript.

Competing interests. We declare we have no competing interests.

Funding. This work was supported by the National Science Foundation Graduate Research Fellowship Programme grant no. DGE-1419118 to S.E.K., and NSF 2000759 from the Division of Research on Learning in Formal and Informal Settings (DRL) to S.T.P., the Austrian Science Fund (FWF project no. P33928_B) to F.R., the Alfred P. Sloan Foundation Fellowship grant no. 220020300 to J.F.C., National Institutes of Health grant no. R01 HD085996 to J.F.C. and S.T.P. and the James S. McDonnell Foundation.

References

- Gallistel CR. 1990 *The organization of learning*. Cambridge, MA: Bradford Books/MIT Press.
- Agrillo C, Dadda M, Serena G, Bisazza A. 2009 Use of number by fish. *PLoS ONE* **4**, e4786. (doi:10.1371/journal.pone.0004786)
- Agrillo C, Bisazza A. 2017 Understanding the origin of number sense: a review of fish studies. *Phil. Trans. R. Soc. B* **B373**, 20160511. (doi:10.1098/rstb.2016.0511)
- Scarf D, Hayne H, Colombo M. 2011 Pigeons on par with primates in numerical competence. *Science* **334**, 1664. (doi:10.1126/science.1213357)
- Nieder A. 2020 The adaptive value of numerical competence. *Trends Ecol. Evol.* **35**, 605–617. (doi:10.1016/j.tree.2020.02.009)
- Cantlon JF. 2017 How evolution constrains human numerical concepts. *Child Dev. Perspect.* **12**, 65–71. (doi:10.1111/cdep.12264)
- Piantadosi ST, Cantlon JF. 2017 True numerical cognition in the wild. *Psychol. Sci.* **28**, 462–469. (doi:10.1177/0956797616686862)
- Strandburg-Peshkin A, Farine DR, Couzin ID, Crofoot MC. 2015 Shared decision-making drives collective movement in wild baboons. *Science* **348**, 1358–1361. (doi:10.1126/science.aaa5099)
- Walls ML, Layne JE. 2009 Direct evidence for distance measurement via flexible stride integration in the fiddler crab. *Curr. Biol.* **19**, 25–29. (doi:10.1016/j.cub.2008.10.069)
- Wittlinger M, Wehner R, Wolf H. 2006 The ant odometer: stepping on stilts and stumps. *Neuroforum* **12**, 240–241. (doi:10.1515/nf-2006-0307)
- Benson-Amram S, Heinen VK, 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)
- Bonanni R, Natoli E, Cafazzo S, Valsecchi P. 2011 Free-ranging dogs assess the quantity of opponents in intergroup conflicts. *Anim. Cogn.* **14**, 103–115. (doi:10.1007/s10071-010-0348-3)
- Cresswell W, Quinn JL. 2004 Faced with a choice, sparrowhawks more often attack the more vulnerable prey group. *Oikos* **104**, 71–76. (doi:10.1111/j.0030-1299.2004.12814.x)
- Kitchen DM. 2004 Alpha male black howler monkey responses to loud calls: effect of numeric odds, male companion behaviour and reproductive investment. *Anim. Behav.* **67**, 125–139. (doi:10.1016/j.anbehav.2003.03.007)
- McComb K, Packer C, Pusey A. 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)
- 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)
- Wilson ML, Kahlenberg SM, Wells M, Wrangham RW. 2012 Ecological and social factors affect the occurrence and outcomes of intergroup encounters in chimpanzees. *Anim. Behav.* **83**, 277–291. (doi:10.1016/j.anbehav.2011.11.004)
- Gatto E, Carlesso D. 2019 Spontaneous quantity discrimination in crickets. *Ethology* **125**, 613–619. (doi:10.1111/eth.12912)
- Miletto Petrazzini ME, Bertolucci C, Foà A. 2018 Quantity discrimination in trained lizards (*Podarcis sicula*). *Front. Psychol.* **9**, 274. (doi:10.3389/fpsyg.2018.00274)
- 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)
- Cantlon JF, Brannon EM. 2007 Basic math in monkeys and college students. *PLoS Biol.* **5**, 2912–2919. (doi:10.1371/journal.pbio.0050328)
- Lucon-Xiccato T, Gatto E, Bisazza A. 2018 Quantity discrimination by treefrogs. *Anim. Behav.* **139**, 61–69. (doi:10.1016/j.anbehav.2018.03.005)
- Beran MJ. 2012 Quantity judgments of auditory and visual stimuli by chimpanzees (*Pan troglodytes*). *J. Exp. Psychol. Anim. Behav. Process.* **38**, 23–29. (doi:10.1037/a0024965)
- Jordan KE, Brannon EM, Logothetis NK, Ghazanfar AA. 2005 Monkeys match the number of voices they hear to the number of faces they see. *Curr. Biol.* **15**, 1034–1038. (doi:10.1016/j.cub.2005.04.056)
- 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)
- Meck WH, Church RM. 1983 A mode control model of counting and timing processes. *J. Exp. Psychol. Anim. Behav. Process.* **9**, 320–334. (doi:10.1037/0097-7403.9.3.320)
- 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)
- Brannon EM, Terrace HS. 1998 Ordering of the numerosities 1 to 9 by monkeys. *Science* **282**, 746–749. (doi:10.1126/science.282.5389.746)
- 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)
- Nieder A, Diester I, Tudusciuc O. 2006 Temporal and spatial enumeration processes in the primate parietal cortex. *Science* **43**, 1431–1435. (doi:10.1126/science.1130308)
- Akre KL, Johnsen S. 2014 Psychophysics and the evolution of behavior. *Trends Ecol. Evol.* **29**, 291–300. (doi:10.1016/j.tree.2014.03.007)
- Piantadosi ST. 2016 A rational analysis of the approximate number system. *Psychon. Bull. Rev.* **23**, 877–886. (doi:10.3758/s13423-015-0963-8)
- Nieder A. 2017 Evolution of cognitive and neural solutions enabling numerosity judgements: lessons from primates and corvids. *Phil. Trans. R. Soc. B* **373**, 20160514. (doi:10.1098/rstb.2016.0514)
- Brannon EM, Van De Walle GA. 2001 The development of ordinal numerical competence in young children. *Cogn. Psychol.* **43**, 53–81. (doi:10.1006/cogp.2001.0756)
- Cantlon J, Fink R, Safford K, Brannon EM. 2007 Heterogeneity impairs numerical matching but

- not numerical ordering in preschool children. *Dev. Sci.* **10**, 431–440. (doi:10.1111/j.1467-7687.2007.00597.x)
36. Gallistel CR, Gelman R. 1992 Preverbal and verbal counting and computation. *Cognition* **44**, 43–74. (doi:10.1016/0010-0277(92)90050-R)
37. 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)
38. Halberda J, Odic D. 2015 The precision and internal confidence of our approximate number thoughts. *Math. Cognit. Learn.* **1**, 305–333. (doi:10.1016/B978-0-12-420133-0.00012-0)
39. Piantadosi ST. 2016 Efficient estimation of Weber's *W. Behav. Res. Methods* **48**, 42–52. (doi:10.3758/s13428-014-0558-8)
40. Bortot M, Regolin L, Vallortigara G. 2021 A sense of number in invertebrates. *Biochem. Biophys. Res. Commun.* **564**, 37–42. (doi:10.1016/j.bbrc.2020.11.039)
41. Gross HJ, Pahl M, Si A, Zhu H, Jurgen T, Zhang S. 2009 Number-based visual generalisation in the honeybee. *PLoS Biol.* **4**, e4263. (doi:10.1371/journal.pone.0004263)
42. Chittka L, Geiger K. 1995 Can honey bees count landmarks? *Anim. Behav.* **49**, 159–164. (doi:10.1016/0003-3472(95)80163-4)
43. Howard SR, Avarguès-Weber A, Garcia JE, Greentree AD, Dyer AG. 2018 Numerical ordering of zero in honey bees. *Science* **360**, 1124–1126. (doi:10.1126/science.aar4975)
44. Bortot M, Agrillo C, Avarguès-Weber A, Bisazza A, Petrazzini MEM, Giurfa M. 2019 Honeybees use absolute rather than relative numerosity in number discrimination. *Biol. Lett.* **15**, 20190138. (doi:10.1098/rsbl.2019.0138)
45. Howard SR, Avarguès-Weber A, Garcia JE, Greentree AD, Dyer AG. 2019 Surpassing the subitizing threshold: appetitive–aversive conditioning improves discrimination of numerosities in honeybees. *J. Exp. Biol.* **222**, 205658. (doi:10.1242/jeb.205658)
46. Shettleworth SJ. 2009 The evolution of comparative cognition: is the snark still a boojum? *Behav. Processes* **80**, 210–217. (doi:10.1016/j.beproc.2008.09.001)
47. Shettleworth SJ. 2010 *Cognition, evolution, and behavior*. Oxford, UK: Oxford University Press.
48. Maclean EL *et al.* 2012 How does cognition evolve? Phylogenetic comparative psychology. *Anim. Cogn.* **15**, 223–238. (doi:10.1007/s10071-011-0448-8)
49. MacLean EL *et al.* 2014 The evolution of self-control. *Proc. Natl Acad. Sci. USA* **111**, E2140–E2148. (doi:10.1073/pnas.1323533111)
50. Primates M *et al.* 2019 Establishing an infrastructure for collaboration in primate cognition research. *PLoS ONE* **14**, e0223675. (doi:10.1371/journal.pone.0223675)
51. Deaner RO, Isler K, Burkart J, Van Schaik C. 2007 Overall brain size, and not encephalization quotient, best predicts cognitive ability across non-human primates. *Brain. Behav. Evol.* **70**, 115–124. (doi:10.1159/000102973)
52. Maclean EL, Sandel AA, Bray J, Oldenkamp RE, Reddy RB. 2013 Group size predicts social but not nonsocial cognition in lemurs. *PLoS ONE* **8**, 66359. (doi:10.1371/journal.pone.0066359)
53. Burkart JM *et al.* 2014 The evolutionary origin of human hyper-cooperation. *Nat. Commun.* **5**, 4747. (doi:10.1038/ncomms5747)
54. Street SE, Navarrete AF, Reader SM, Laland KN. 2017 Coevolution of cultural intelligence, extended life history, sociality, and brain size in primates. *Proc. Natl Acad. Sci. USA* **114**, 7908–7914. (doi:10.1073/pnas.1620734114)
55. Stevens JR. 2014 Evolutionary pressures on primate intertemporal choice. *Proc. R. Soc. B* **281**, 20140499. (doi:10.1098/rspb.2014.0499)
56. Herculano-Houzel S. 2012 The remarkable, yet not extraordinary, human brain as a scaled-up primate brain and its associated cost. *Proc. Natl Acad. Sci. USA* **109**, 10 661–10 668. (doi:10.1073/pnas.1201895109)
57. Kaplan H, Hill K, Lancaster J, Hurtado AM. 2000 A theory of human life history evolution: diet, intelligence, and longevity. *Evol. Anthropol.* **9**, 156–185. (doi:10.1002/1520-6505(2000)9:4<156::AID-EVAN5>3.0.CO;2-7)
58. Seyfarth RM, Cheney DL. 2002 What are big brains for? *Proc. Natl Acad. Sci. USA* **99**, 4141–4142. (doi:10.1073/pnas.082105099)
59. Isler K, Van SC. 2006 Metabolic costs of brain size evolution. *Biol. Lett.* **2**, 557–560. (doi:10.1098/rsbl.2006.0538)
60. Herculano-Houzel S. 2011 Not all brains are made the same: new views on brain scaling in evolution. *Brain Behav. Evol.* **78**, 22–36. (doi:10.1159/000327318)
61. Kaas JH. 2006 Evolution of the neocortex. *Curr. Biol.* **16**, 910–914. (doi:10.1016/j.cub.2006.09.057)
62. Benson-Amram S, Dantzer B, Stricker G, Swanson EM, Holekamp KE. 2016 Brain size predicts problem-solving ability in mammalian carnivores. *Proc. Natl Acad. Sci. USA* **113**, 2532–2537. (doi:10.1073/pnas.1505913113)
63. Lefebvre L, Reader SM, Sol D. 2004 Brains, innovations and evolution in birds and primates. *Brain Behav. Evol.* **63**, 233–246. (doi:10.1159/000076784)
64. Reader SM, Laland KN. 2002 Social intelligence, innovation, and enhanced brain size in primates. *Proc. Natl Acad. Sci. USA* **99**, 4436–4441. (doi:10.1073/pnas.062041299)
65. Sol D, Bacher S, Reader SM, Lefebvre L. 2008 Brain size predicts the success of mammal species introduced into novel environments. *Am. Nat.* **172**, S63–S71. (doi:10.1086/588304)
66. Sol D, Duncan RP, Blackburn TM, Cassey P, Lefebvre L. 2005 Big brains, enhanced cognition, and response of birds to novel environments. *Proc. Natl Acad. Sci. USA* **102**, 5460–5465. (doi:10.1073/pnas.0408145102)
67. Kalt T, Diekamp B, Güntürkün O. 1999 Single unit activity during a Go/NoGo task in the 'prefrontal cortex' of pigeons. *Brain Res.* **839**, 263–278. (doi:10.1016/S0006-8993(99)01727-8)
68. Güntürkün O. 2005 The avian 'prefrontal cortex' and cognition. *Curr. Opin. Neurobiol.* **15**, 686–693. (doi:10.1016/j.conb.2005.10.003)
69. Smaers JB *et al.* 2021 The evolution of mammalian brain size. *Sci. Adv.* **7**, eabe2101. (doi:10.1126/sciadv.abe2101)
70. van Schaik CP, Van Triki Z, Bshary R, Heldstab S. 2021 A farewell to EQ: a new brain size measure for comparative primate cognition. *Brain Behav. Evol.* **96**, 1–19. (doi:10.1159/000517013)
71. Herculano-Houzel S. 2017 Numbers of neurons as biological correlates of cognitive capability. *Curr. Opin. Behav. Sci.* **16**, 1–7. (doi:10.1016/j.cobeha.2017.02.004)
72. Herculano-Houzel S, Lent R. 2005 Isotropic fractionator: a simple, rapid method for the quantification of total cell and neuron numbers in the brain. *J. Neurosci.* **25**, 2518–2521. (doi:10.1523/JNEUROSCI.4526-04.2005)
73. Herculano-Houzel S. 2012 The isotropic fractionator: a fast, reliable method to determine numbers of cells in the brain or other tissues. *NeuroMethods* **67**, 391–403. (doi:10.1007/7657_2011_13)
74. Neves K *et al.* 2014 Cellular scaling rules for the brain of afrotherians. *Front. Neuroanat.* **8**, 5. (doi:10.3389/fnana.2014.00005)
75. Herculano-Houzel S, Collins CE, Wong P, Kaas JH. 2007 Cellular scaling rules for primate brains. *Proc. Natl Acad. Sci. USA* **104**, 3562–3567. (doi:10.1073/pnas.0611396104)
76. Herculano-Houzel S, Mota B, Lent R. 2006 How to build a bigger brain: cellular scaling rules for rodent brains. *Proc. Natl Acad. Sci. USA* **103**, 12 138–12 143. (doi:10.1073/pnas.0604911103)
77. Sarko DK, Catania KC, Leitch DB, Kaas JH, Herculano-Houzel S. 2009 Cellular scaling rules of insectivore brains. *Front. Neuroanat.* **3**, 8. (doi:10.3389/neuro.05.008.2009)
78. Gabi M, Collins CE, Wong P, Torres LB, Kaas JH, Herculano-Houzel S. 2010 Cellular scaling rules for the brains of an extended number of primate species. *Brain Behav. Evol.* **76**, 32–44. (doi:10.1159/000319872)
79. Olkowicz S, Kocourek M, Lučan RK, Porteš M, Fitch WT, Herculano-Houzel S, Némec 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)
80. Dos Santos SE *et al.* 2017 Cellular scaling rules for the brains of marsupials: not as 'primitive' as expected. *Brain. Behav. Evol.* **89**, 48–63. (doi:10.1159/000452856)
81. Jardim-Messeder D *et al.* 2017 Dogs have the most neurons, though not the largest brain: trade-off between body mass and number of neurons in the cerebral cortex of large carnivorous species. *Front. Neuroanat.* **11**, 118. (doi:10.3389/fnana.2017.00118)
82. Herculano-Houzel S. 2009 The human brain in numbers: a linearly scaled-up primate brain. *Front.*

- Hum. Neurosci.* **3**, 1–11. (doi:10.3389/neuro.09.031.2009)
83. Herculano-Houzel S. 2011 Brains matter, bodies maybe not: the case for examining neuron numbers irrespective of body size. *Ann. NY Acad. Sci.* **1225**, 191–199. (doi:10.1111/j.1749-6632.2011.05976.x)
84. Herculano-Houzel S, Catania K, Manger PR, Kaas JH. 2015 Mammalian brains are made of these: a dataset of the numbers and densities of neuronal and nonneuronal cells in the brain of Glires, Primates, Scandentia, Eulipotyphlans, Afrotherians and Artiodactyls, and their relationship with body mass. *Brain Behav. Evol.* **86**, 145–163. (doi:10.1159/000437413)
85. Ditz HM, Nieder A. 2016 Sensory and working memory representations of small and large numerosities in the crow endbrain. *J. Neurosci.* **36**, 12044. (doi:10.1523/JNEUROSCI.1521-16.2016)
86. Cantlon JF, Brannon EM, Carter EJ, Pelphrey KA. 2006 Functional imaging of numerical processing in adults and 4-y-old children. *PLoS Biol.* **4**, 844–854. (doi:10.1371/journal.pbio.0040125)
87. Cantlon JF, Libertus ME, Pinel P, Dehaene S, Brannon EM, Pelphrey KA. 2009 The neural development of an abstract concept of number. *J. Cogn. Neurosci.* **21**, 2217–2229. (doi:10.1162/jocn.2008.21159)
88. Herculano-Houzel S. 2010 Coordinated scaling of cortical and cerebellar numbers of neurons. *Front. Neuroanat.* **4**, 12. (doi:10.3389/fnana.2010.00012)
89. Barton RA. 2012 Embodied cognitive evolution and the cerebellum. *Phil. Trans. R. Soc. B* **1599**, 2097–2107. (doi:10.1098/rstb.2012.0112)
90. Barton RA, Venditti C. 2014 Rapid evolution of the cerebellum in humans and other great apes. *Curr. Biol.* **24**, 2440–2444. (doi:10.1016/j.cub.2014.08.056)
91. Smaers JB, Turner AH, Gómez-Robles A, Sherwood CC. 2018 A cerebellar substrate for cognition evolved multiple times independently in mammals. *Elife* **7**, 1–19. (doi:10.7554/eLife.35696)
92. Balda RP, Kamil AC. 1989 A comparative study of cache recovery by three corvid species. *Anim. Behav.* **38**, 486–495. (doi:10.1016/S0003-3472(89)80041-7)
93. Balda RP, Kamil AC, Bednekoff PA. 1996 Predicting cognitive capacity from natural history. In *Current ornithology*, vol. 13 (eds V Nolan Jr, ED Ketterson), pp. 33–66. Boston, MA: Springer.
94. Pravosudov VV, Roth II TC. 2013 Cognitive ecology of food hoarding: the evolution of spatial memory and the hippocampus. *Annu. Rev. Ecol. Evol. Syst.* **44**, 173–193. (doi:10.1146/annurev-ecolsys-110512-135904)
95. Platt ML, Brannon EM, Briese TL, French JA. 1996 Differences in feeding ecology predict differences in performance between golden lion tamarins (*Leontopithecus rosalia*) and Wied's marmosets (*Callithrix kuhli*) on spatial and visual memory tasks. *Anim. Learn. Behav.* **24**, 384–393. (doi:10.3758/BF03199010)
96. Rosati AG, Rodriguez K, Hare B. 2014 The ecology of spatial memory in four lemur species. *Anim. Cogn.* **17**, 947–961. (doi:10.1007/s10071-014-0727-2)
97. Heilbronner SR, Rosati AG, Stevens JR, Hare B, Hauser MD. 2008 A fruit in the hand or two in the bush? Divergent risk preferences in chimpanzees and bonobos. *Biol. Lett.* **4**, 246–249. (doi:10.1098/rsbl.2008.0081)
98. Rosati AG, Hare B. 2012 Chimpanzees and bonobos exhibit divergent spatial memory development. *Dev. Sci.* **15**, 840–853. (doi:10.1111/j.1467-7687.2012.01182.x)
99. Rosati AG, Hare B. 2013 Chimpanzees and bonobos exhibit emotional responses to decision outcomes. *PLoS ONE* **8**, e63058. (doi:10.1371/journal.pone.0063058)
100. Rosati AG. 2017 Foraging cognition: reviving the ecological intelligence hypothesis. *Trends Cogn. Sci.* **21**, 691–702. (doi:10.1016/j.tics.2017.05.011)
101. Milton K. 1981 Distribution patterns of tropical plant foods as an evolutionary stimulus to primate mental development. *Am. Anthropol.* **83**, 534–548. (doi:10.1525/aa.1981.83.3.02a00020)
102. Stephens DW, Krebs JR. 1986 *Foraging theory*. Princeton, NJ: Princeton University Press.
103. Byrne R, Whiten A. 1988 *Machiavellian intelligence: social expertise and the evolution of intellect in monkeys, apes, and humans*. Oxford, UK: Clarendon Press.
104. Dunbar RIM. 1988 The social brain hypothesis. *Evol. Anthropol.* **6**, 178–190. (doi:10.1002/(SICI)1520-6505(1998)6:5<178::AID-EVAN5>3.0.CO;2-8)
105. Jolly A. 1966 Lemur social behavior and primate intelligence. *Science* **153**, 501–506. (doi:10.1126/science.153.3735.501)
106. Stevens JR, Hauser MD. 2004 Why be nice? Psychological constraints on the evolution of cooperation. *Trends Cogn. Sci.* **8**, 60–65. (doi:10.1016/j.tics.2003.12.003)
107. Halberda J. 2011 *What is Weber Fraction?* See <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.672.2472&rep=rep1&type=pdf>.
108. Popescu AA, Huber KT, Paradis E. 2012 ape 3.0: New tools for distance-based phylogenetics and evolutionary analysis in R. *Bioinformatics* **28**, 1536–1537. (doi:10.1093/bioinformatics/bts184)
109. Hedges SB, Dudley J, Kumar S. 2006 TimeTree: a public knowledge-base of divergence times among organisms. *Bioinformatics* **22**, 2971–2972. (doi:10.1093/bioinformatics/btl505)
110. Kumar S, Hedges SB. 2011 Timetree2: species divergence times on the iPhone. *Bioinformatics* **27**, 2023–2024. (doi:10.1093/bioinformatics/btr315)
111. Miller HC, Lambert DM. 2006 A molecular phylogeny of New Zealand's *Petroica* (Aves: Petroicidae) species based on mitochondrial DNA sequences. *Mol. Phylogenet. Evol.* **40**, 844–855. (doi:10.1016/j.ympev.2006.04.012)
112. Savolainen P. 2002 Genetic evidence for an East Asian origin of domestic dogs. *Science* **298**, 1610–1613. (doi:10.1126/science.1073906)
113. Stephan H, Frahm H, Baron G. 1981 New and revised data on volumes of brain structures in insectivores and primates. *Folia Primatol.* **35**, 1–29. (doi:10.1159/000155963)
114. Herculano-Houzel S, Kaas JH. 2011 Gorilla and orangutan brains conform to the primate cellular scaling rules: implications for human evolution. *Brain Behav. Evol.* **77**, 33–44. (doi:10.1159/000322729)
115. Herculano-Houzel S. 2020 What modern mammals teach about the cellular composition of early brains and mechanisms of brain evolution. In *Evolutionary neuroscience*, 2nd edn (eds JH Kaas), pp. 349–375. London, UK: Academic Press.
116. Seed A, Emery N, Clayton N. 2009 Intelligence in corvids and apes: a case of convergent evolution? *Ethology* **115**, 401–420. (doi:10.1111/j.1439-0310.2009.01644.x)
117. Goodwin D. 1986 *Crows of the world*. London, UK: British Museum (Natural History).
118. Kabadayi C, Taylor LA, Von Bayern AMP, Ravens OM. 2014 New Caledonian crows and jackdaws parallel great apes in motor self-regulation despite smaller brains. *Proc. Natl Acad. Sci. USA* **111**, 2140–2148. (doi:10.1098/rstb.2014.0104)
119. Kabadayi C, Krasheninnikova A, O'Neill L, van de Weijer J, Osavath M, von Bayern AMP. 2017 Are parrots poor at motor self-regulation or is the cylinder task poor at measuring it? *Anim. Cogn.* **20**, 1137–1146. (doi:10.1007/s10071-017-1131-5)
120. Deane RO, Van SC, Johnson V. 2006 Do some taxa have better domain-general cognition than others? A meta-analysis of nonhuman primate studies. *Evol. Psychol.* **4**, 149–196. (doi:10.1177/147470490600400114)
121. Lynch M. 1991 Methods for the analysis of comparative data in evolutionary biology. *Evolution* **45**, 1065–1080. (doi:10.1111/j.1558-5646.1991.tb04375.x)
122. Freckleton RP, Harvey PH, Pagel M. 2002 Phylogenetic analysis and comparative data: a test and review of evidence. *Am. Nat.* **160**, 712–726. (doi:10.1086/343873)
123. Harvey PH, Pagel MD. 1991 *The comparative method in evolutionary biology*. Oxford, UK: Oxford University Press.
124. Housworth EA, Martins EP, Lynch M. 2004 The phylogenetic mixed model. *Am. Nat.* **163**, 84–96. (doi:10.1086/380570)
125. Stan Development Team. 2020 *RStan: the R interface to Stan*. See <http://mc-stan.org/>.
126. Hoffman MD, Gelman A. 2014 The No-U-Turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo. *J. Mach. Learn. Res.* **15**, 1593–1623.
127. Halberda J, Ly R, Wilmer JB, Naiman DQ, Germine L. 2012 Number sense across the lifespan as revealed by a massive Internet-based sample. *Proc. Natl Acad. Sci. USA* **109**, 11 116–11 120. (doi:10.1073/pnas.1200196109)
128. 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)
129. Barnard AM, Hughes KD, Gerhardt RR, DiVincenti L, Bovee JM, Cantlon JF. 2013 Inherently analog quantity representations in olive baboons (*Papio anubis*). *Front. Psychol.* **4**, 1–11. (doi:10.3389/fpsyg.2013.00253)
130. Tomonaga M. 2008 Relative numerosity discrimination by chimpanzees (*Pan troglodytes*): evidence for approximate numerical representations. *Anim. Cogn.* **11**, 43–57. (doi:10.1007/s10071-007-0089-0)
131. 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)
132. Garcia J. 1993 Misrepresentation of my criticisms of Skinner. *Am. Psychol.* **48**, 1158. (doi:10.1037/0003-066X.48.11.1158.a)
133. Roediger III HL. 2004 What happened to behaviorism. *APS*, 1 March. See <https://www.psychologicalscience.org/observer/what-happened-to-behaviorism/comment-page-1>.
134. Seligman ME. 1970 On the generality of the laws of learning. *Psychol. Rev.* **77**, 406–418. (doi:10.1037/h0029790)
135. Neves K, Guercio GD, Anjos-Travassos Y, Costa S, Perozzo A, Montezuma K, Herculano-Houzel S, Panizzutti R. 2020 The relationship between the number of neurons and behavioral performance in Swiss mice. *Neurosci. Lett.* **735**, 135202. (doi:10.1016/j.neulet.2020.135202)
136. Jacob J *et al.* 2021 Cytoarchitectural characteristics associated with cognitive flexibility in raccoons. *J. Comp. Neurol.* **529**, 3375–3388. (doi:10.1002/cne.25197)
137. Gittleman JL. 1986 Carnivore brain size, behavioral ecology, and phylogeny. *J. Mammal.* **67**, 23–36. (doi:10.2307/1380998)
138. Harvey PH, Clutton-Brock TH, Mace GM. 1980 Brain size and ecology in small mammals and primates. *Proc. Natl Acad. Sci. USA* **77**, 4387–4389. (doi:10.1073/pnas.77.7.4387)
139. Cunningham E, Janson C. 2007 A socioecological perspective on primate cognition, past and present. *Anim. Cogn.* **10**, 273–281. (doi:10.1007/s10071-007-0078-3)
140. DeCasien AR, Higham JP. 2019 Primate mosaic brain evolution reflects selection on sensory and cognitive specialization. *Nat. Ecol. Evol.* **3**, 1483–1493. (doi:10.1038/s41559-019-0969-0)