ORIGINAL PAPER

Lemurs and macaques show similar numerical sensitivity

**Sarah M. Jones · John Pearson · Nicholas K. DeWind ·
David Paulsen · Ana-Maria Tenekedjieva ·
Elizabeth M. Brannon**

Abstract We investigated the precision of the approximate number system (ANS) in three lemur species (*Lemur catta*, *Eulemur mongoz*, and *Eulemur macaco flavifrons*), one Old World monkey species (*Macaca mulatta*) and humans (*Homo sapiens*). In Experiment 1, four individuals of each nonhuman primate species were trained to select the numerically larger of two visual arrays on a touchscreen. We estimated numerical acuity by modeling Weber fractions (w) and found quantitatively equivalent performance among all four nonhuman primate species. In Experiment 2, we tested adult humans in a similar procedure, and they outperformed the four nonhuman species but showed qualitatively similar performance. These results indicate that the ANS is conserved over the primate order.

Keywords Nonhuman primates · Numerical cognition · Comparative psychology · Number discrimination · Weber fraction · Prosimians

S. M. Jones · D. Paulsen · E. M. Brannon (✉)
Department of Psychology and Neuroscience, Duke University,
Durham, NC, USA
e-mail: Brannon@duke.edu

S. M. Jones (✉)
Psychology Department, St. Norbert College, De Pere, WI, USA
e-mail: Sarah.jones@snc.edu

J. Pearson · N. K. DeWind · A.-M. Tenekedjieva ·
E. M. Brannon
Center for Cognitive Neuroscience, Duke University, Durham,
NC, USA

J. Pearson · N. K. DeWind
Department of Neurobiology, Duke University Medical Center,
Durham, NC, USA

J. Pearson
Division of Neurosurgery, Duke University Medical Center,
Durham, NC, USA

Introduction

A large literature indicates that the ability to represent number is widespread throughout the animal kingdom. By and large, the focus in the field has been on providing evidence of numerical competence in wide-ranging taxa or using individual primate species as models for comparisons with humans using a wide variety of tasks (for reviews, see Beran 2008; Brannon 2006). Here, we take a different approach by systematically comparing multiple species with a single task and set of stimulus parameters in an attempt to provide a more quantitative assessment of species differences and similarities in numerical cognition.

Ratio-dependent performance in numerical discrimination is the hallmark signature of the approximate number system and has been shown to be common to humans and nonhuman primates, as well as a wide variety of other species. Experiments comparing humans with a single nonhuman animal species using identical stimuli have provided strong evidence for qualitative similarities in ratio-dependent numerical discrimination (e.g., Agrillo et al. 2012a, b; Beran et al. 2008, 2011a, b; Cantlon and Brannon 2006, 2007; Jordan and Brannon 2006a, b). Agrillo et al. (2012a, b) recently used identical procedures, stimuli, and values to compare quantity abilities of five fish species. Their results showed evidence of similar, ratio-dependent numerical abilities across species. Hanus and

Call (2007) also investigated the ability to select the larger of two quantities in multiple ape species: bonobos, chimpanzees, gorillas, and orangutans. All apes showed ratio-dependent performance such that they were increasingly successful at selecting the larger number of food items as ratio (small number/larger number) decreased. While these studies provide strong evidence for ratio dependence across species, none of them have parametrically tested whether species differ quantitatively in the acuity of the approximate number system.

Here, we attempt to quantitatively compare the numerical acuity of 5 primate species: humans, rhesus macaques, and three different species of lemur. Including prosimian primates in parametric comparisons of primate cognition is important in attempting to identify cognitive profiles of the primate ancestral state. The prosimian suborder of primates (lemurs, lorises, and galagos) diverged from other primates approximately 50 million years before monkeys diverged from the human ancestral line (75 vs. 25 million years ago; Horvath et al. 2008). Prosimians are considered to be among the most "evolutionarily conserved" primates, in so far as they have been hypothesized to be morphologically and behaviorally similar to the last common primate ancestor (Tattersall 1982; Yoder 2007). If prosimians share cognitive traits that are common among other primates, it is likely that these traits were present in the last common ancestor.

The limited data available on lemur cognition indicates that lemurs differ from apes and monkeys both quantitatively and qualitatively on measures of social intelligence. For example, lemurs do not exhibit within-group coalitions, deception, or elaborate visual communication (for a review, see Fichtel and Kappeler 2010). Monkeys outperform lemurs in visual coorientation with humans (Anderson and Mitchell 1999; but see Shepherd and Platt 2008; Ruiz et al. 2009), and vervet monkeys have been found to be superior to ring-tailed lemurs (and 1 brown lemur) on a reversal discrimination task (Rumbaugh and Arnold 1971).

On the other hand, a number of recent studies have found qualitative similarities between lemurs and monkeys in nonsocial cognitive tasks. For example, lemurs perform more like macaque monkeys than pigeons in a sequence learning task (Merritt et al. 2007). Lemurs are capable of transitive reasoning (MacLean et al. 2008), Piagetian object displacement (Deppe et al. 2009), and learning a reverse contingency in a food reward task (Genty et al. 2004). Lemurs also perform similarly to monkeys on assays of spatial memory (Erhart and Overdorff 2008; Luhrs et al. 2009; Picq 1993, 2007) and inferring causality (Santos et al. 2005a, b).

Only a few studies have investigated numerical cognition in prosimian primates (Lewis et al. 2005; Merritt et al. 2011; Santos et al. 2005a, b). Santos et al. (2005a, b)

investigated the spontaneous representation of number in lemurs (brown lemurs, mongoose lemurs, ring-tailed lemurs, and ruffed lemurs) with a modified version of the Wynn (1992) violation of expectancy paradigm. Lemurs watched as two lemons were hidden behind a screen; the screen was then lowered to reveal either a possible outcome (two lemons) or an impossible outcome (one lemon or three lemons). On average, lemurs looked longer at the impossible outcome than at the possible outcome, suggesting an ability to represent the number of lemons behind the screen. Lewis et al. (2005) modified another task designed for human infants (Feigenson and Carey 2003) to explore numerical abilities in lemurs. In this study, mongoose lemurs watched as grapes were hidden one at a time in a bucket that was filled with shredded paper. A subset of the food items were hidden in the container's false bottom. Lemurs were allowed to retrieve and consume the available food items, and additional search time was then measured. If lemurs expected additional grapes to be in the bucket, they should have continued to manually search after retrieving the available food items. Results indicated that search time was dependent on the number of grapes that should have been in the bucket and reflected ratio-dependent discrimination (e.g., lemurs searched longer after retrieving 2 grapes if 4 grapes had been hidden, a 1:2 ratio of retrieved grapes to presented grapes, but they did not search longer after retrieving 2 grapes if 3 grapes had been hidden). Recently, Jones and Brannon (2012) used a third task to directly test whether spontaneous food choices in lemurs reveal ratio dependence or alternatively set-size limitations as have been reported in human infants and rhesus monkeys (Feigenson et al. 2002; Hauser et al. 2000). When human infants are given a choice between two sequentially baited buckets, they choose the bucket with the greater amount of food but only when the quantities are small. A similar phenomenon has been reported in rhesus macaques (Hauser et al. 2000). In contrast, Jones and Brannon (2012) found evidence for ratio dependence in the spontaneous quantity discriminations of lemurs and no evidence for set-size limitations.

In contrast to the studies reviewed above which used food items as the discriminative stimuli and examined spontaneous numerical discrimination in lemurs, Merritt et al. (2011) used a touchscreen task and trained two ring-tailed lemurs to order visual arrays based on numerosity. Lemurs were trained to select the smaller of two arrays of dots (values 1–4), using trial-unique stimuli and surface area controls. After successfully learning to order pairs of the values 1–4, the lemurs were able to extrapolate this ordinal rule to novel numerical values (arrays of 5–9 dots) replicating the basic pattern of results found by Brannon and Terrace in macaque monkeys (Brannon and Terrace 1998, 2000). While lemur performance in this ordinal

numerical comparison task was qualitatively similar to that of monkeys, accuracy was generally lower for the ring-tailed lemurs compared to the macaques. It is unclear, however, whether these differences reflect true species differences in capacity or instead reflect task parameters and training history (e.g., Savage and Snowdon 1989).

Lemur performance in the Merritt et al. (2011) study showed the hallmark signature of the approximate number system (ANS), wherein numerical discrimination, was modulated by ratio in accordance with Weber's law. According to Weber's law, the increase in intensity required to detect a change to a stimulus is a function of the magnitude of the initial intensity. The internal Weber fraction, w, specifies the precision of the internal quantity representation. A Weber fraction, w, can be roughly translated to the ratio that would result in 75 % correct performance using the formula $w = n_L/n_S - 1$, in which n_L is the larger numerosity and n_S is the smaller numerosity being compared. In other words, w can be thought of as the proportion of n_S that needs to be added to n_S to allow successful discrimination of the two values 75 % of the time. Thus, a Weber fraction of $w = 0.2$ indicates that the subject would require that $n_L = 0.2\ n_S + n_S$ to allow successful discrimination on 75 % of trials. In the models presented here, w is an estimated parameter and is thus conceptually similar but computationally distinct from the formula $w = n_L/n_S - 1$. However, it is also important to note that Weber fractions are not hard and fast invariants. The absolute value will vary based on the model used to estimate w and may also vary based on a multitude of different factors such as the duration of stimulus presentation or the range of numerical values used. These factors mean that it is essential to use the same task parameters when attempting to make species comparisons.

To compare the precision of numerical representations in the following experiments, we used Weber fractions computed by fitting two models of numerosity encoding: the linear and log models. While it is empirically observed that the precision of numerical discriminations decreases with numerical magnitude (but is fixed as a proportion of the magnitude), the two models reproduce this finding in distinct ways. The linear model presented by Pica et al. (2004), hereafter referred to as w_{linear}, assumes that the numerosity of a set of n dots is represented internally by a Gaussian random variable with standard deviation proportional to its mean. That is, the internal estimates of number $\tilde{n}$ are normally distributed about the stimulus value, with constant coefficient of variation (=std/mean) w_{lin}.

$$p(\tilde{n}|n) = N\left(n, (wn)^2\right) = \frac{1}{\sqrt{2\pi(wn)^2}} e^{-\frac{(\tilde{n}-n)^2}{2(wn)^2}}$$

In the log model, presented in Piazza et al. (2004), the subjective representation of number is normally distributed on a log scale with fixed precision, resulting in a standard deviation that grows with number on the original (uncompressed) scale. $p(\eta|n) = N(\log(n), w^2) = \frac{1}{\sqrt{2\pi w^2}} e^{-\frac{(\eta-\log(n))^2}{2w^2}}$.

It is further assumed that animals' decisions result from a process of randomly drawing one sample from the distribution of each encoded stimulus and choosing the option corresponding to the larger sample. Thus, the number of errors grows with the overlap of the distributions encoding the internal representations of the stimuli.

In Experiment 1, we tested three species of lemurs and rhesus macaques in an identical ordinal touchscreen task to more quantitatively assess species differences in Weber fraction for numerical discrimination. Prosimian primates are a suborder of primates with tremendous diversity in social structure and ecology. Despite their close phylogenetic relatedness, prosimians primates differ widely in their home range, social group size, and diet. By directly contrasting the performance of lemur species that differ in these traits, we are able to test the hypothesis that differences in ecology or social structure cause large differences in nonsocial cognition. To this end, we tested three lemur species: blue-eyed black lemurs, ring-tailed lemurs, and mongoose lemurs. Ring-tailed lemurs (Lemur catta) live in large social groups with a well-defined hierarchy, have a broad home range, and consume fruit, leaves, flowers, bark, sap, insects, and animals (Sauther et al. 1999). Blue-eyed black lemurs (Eulemur macaco flavifrons) live in smaller social groups of 6–10 individuals, have a variable home range, and eat primarily fruit, leaves, and flowers (Volampeno et al. 2011). Mongoose lemurs (Eulemur mongoz) live in small groups consisting of a mating pair and up to three dependent offspring, have a small home range, and eat primarily fruit, flowers, and nectar (Curtis and Zaramody 1998). If selective pressure to discriminate between quantities of discrete food items results in more precise number representations, mongoose lemurs and blue-eyed black lemurs would be expected to outperform ring-tailed lemurs by virtue of a dietary reliance on fruit. Alternatively, if selective pressures to discriminate between a large number of conspecifics in one's group and/or a rival group results in greater numerical acuity, ring-tailed lemurs (and to a lesser extent, blue-eyed black lemurs) would be expected to outperform mongoose lemurs. Finally, if spatial abilities associated with a large home range result in a more evolved magnitude representation system, the broad ranging ring-tailed lemurs should outperform other species.

Experiment 1

Methods

Subjects

Subjects were 4 male individuals from each of four species: mongoose lemurs (*E. mongoz*, mean age 18.68 years, standard deviation 5.92), ring-tailed lemurs (*L. catta*, mean age 8.41 years, standard deviation 6.05), blue-eyed black lemurs (*E. macaco flavifrons*, mean age 12.30 years, standard deviation 3.95), and rhesus macaques (*Macaca mulatta*, mean age 9.21 years, standard deviation 4.10). All animals were experimentally naïve. Animals were neither food- nor water-restricted.

Apparatus

Prosimian subjects were housed in indoor enclosures at the Duke University Lemur Center (DLC). Animals were housed either singly or in pairs, and those housed in pairs were either separated from their cage-mate during testing or had a cage-mate who was trained not to interfere. A custom-built, stainless steel testing station (86 cm × 43 cm × 35 cm) containing a 15-inch touch-sensitive computer monitor, a desktop computer, and a reward delivery system was used to test prosimian primates. The apparatus was wheeled into each subject's home enclosure for all training and testing sessions. A clear Plexiglas panel with circular openings (diameter 5 cm) over each stimulus location covered the screen to prevent unnecessary contact with areas of the touchscreen that did not contain stimuli. Lemurs were trained to sit on a small plastic crate in front of the cart to engage in the task.

Macaques were housed either singly or in pairs in a vivarium. Those housed in pairs were separated for testing. All animals had been trained to use the touchscreen for a previous task, but had no numerical discrimination training. A 15-inch touch-sensitive computer monitor and a reward delivery system were attached to the front of each macaque's cage for each test session. Macaques were kept on water-restricted diets for participation in other tasks, as approved by an Institutional Animal Care and Use Committee.

Task

The experimental program was identical for the four species. A custom-built REALbasic program displayed the stimuli and recorded responses. Subjects initiated each trial by touching a red square presented in the bottom center of the screen. Following the start response, two yellow squares, each containing an array of red dots, were presented in two central screen locations. The location of the correct stimulus (left vs. right) was determined randomly. Subjects were rewarded for selecting the array containing the larger number of dots. Correct responses elicited positive visual and auditory feedback and a food reward. Incorrect responses elicited a warning tone and a black screen for 6 s. All trials were followed by a 3-s inter-trial interval (ITI). For lemurs, correct responses were rewarded with a 190-mg fruit punch-flavored TestDiet® pellet (Division of Land O'Lakes, Purina Feed, LLC, Richmond, IN, USA), and for macaques, the same fruit flavored pellet or a miniature M&M or Reese's Pieces candy.

Stimuli

Stimuli were 200 × 200 pixel yellow squares within which 1–9 elements were randomly arranged (Fig. 1). The computer program randomly selected the size and location of the dots in each array and generated trial-unique images with the restriction that dots did not overlap with other dots or with the boundary of the square in which they were contained. To ensure that the task could not be solved by attending to the cumulative surface area of the arrays or the size of the individual elements, the physical sizes of the elements were varied such that on 50 % of trials, element
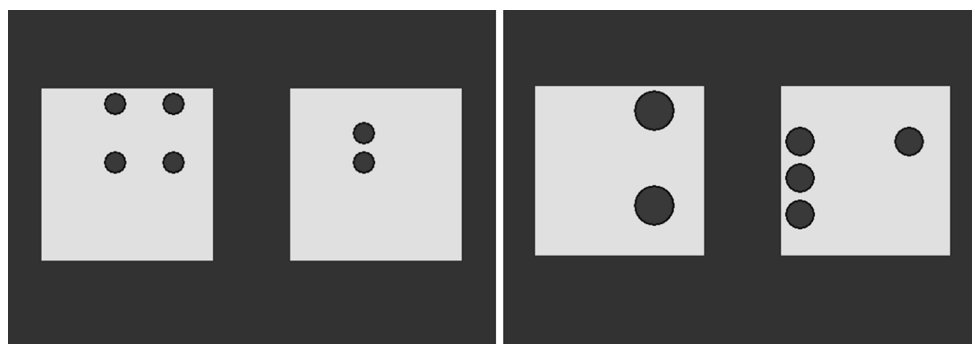

Fig. 1 Example stimuli

size was equated between the two choice stimuli and on the other 50 % of trials, cumulative surface area was equated. The cumulative surface area as well as the individual element size of the choice stimuli varied between 174 and 7,854 pixels.

Procedure

Subjects were tested Monday through Friday for one session per day. Two individuals of each species were trained to select the larger numerosity using the pairs {1 vs. 9, 1 vs. 3, 3 vs. 9}. The other two individuals were trained to select the larger numerosity using a different set: the pairs {2 vs. 8, 2 vs. 4, 4 vs. 8}. Each training session contained 20 trial-unique sets per numerosity pair each presented 3 times for a total of 60 trials. Each subject was trained until accuracy reached 75 % or higher for two sessions, or until 30 training sessions had been completed. After the criterion was met, each animal moved on to the test phase of the experiment, which was characterized by expansion of the stimuli to all 36 numerosity pairs (3 training pairs plus 33 novel pairs) made possible by combining the numerosities 1–9. These values were chosen to reflect small and large values, and a broad range of ratios and because these values were used in prior studies of numerical competence using similar methods (e.g., Brannon and Terrace 1998). Differential reinforcement was used throughout both the training and test phases. Lemurs were tested with 15 72-trial sessions, and monkeys were tested with 30 36-trial sessions. The monkeys received fewer trials per day due to their involvement in other research using consumable rewards.

Analyses

Bayesian analysis

We performed a hierarchical Bayesian analysis (Gelman and Hill 2007; Wagenmakers et al. 2010) using Markov Chain Monte Carlo (MCMC) simulation via the RJAGS package in R. That is, we sampled the posterior distributions of the Weber fractions, w, for each animal, conditioned on its observed choices. We modeled these Weber fractions as drawn from species-specific gamma distributions with distinct means and standard deviations. We assumed these two hyperparameters to be likewise gamma-distributed (with scale and rate parameters (4, 3) and (3, 4), respectively, though results were robust against alternative choices of hyperparameters, as detailed in Online Resource 1).

Specifically, we ran two Markov chains of $N = 20,000$ samples each, thinning each chain's samples by a factor of 3 following a burn-in period of 5,000 samples. This was done to reduce the autocorrelations of successive samples, thereby ensuring that the samples were indeed independent draws from the posterior distributions for w. We verified convergence of this process by examination of the chains' autocorrelation functions and use of the Geweke diagnostic, which compares means early and late in the chain (which should be the same at convergence) (Best et al. 1995; Cowles and Carlin 1996; Gelman et al. 2004) via the CODA package in R.

To compare models, we used the method of Bayes factors. For a given pair of models M_1 and M_2, the Bayes factor is defined as the ratio of the likelihoods for the two models

$$BF = {p(D|M_1)}/{p(D|M_2)}$$

where D is the observed data. This is related to the posterior odds ratio of the two models $p(M_1|D)/p(M_2|D)$ by the ratio of the priors of the two models and so represents the relative probability of each model apart from prior considerations (i.e., due to evidence). We calculated Bayes factors for four models: (1) the linear model with species-specific distributions of w, (2) the linear model with a single, cross-species distribution of w, (3) the logarithmic model with species-specific distributions for w, and (4) the logarithmic model with a single, cross-species distribution of w. While a variety of approaches to calculating Bayes factors exist (e.g., Kass and Raftery 1995; Wagenmakers et al. 2010), we used the Laplace–Metropolis method, a numerical approximation (based on the method of steepest descent) to the integral over model parameters that defines the likelihood (Kass and Raftery 1995) (see Online Resource 1), which is a good approximation for large datasets such as ours. In interpreting these results, we followed the recommendations in Kass and Raftery (1995), in which 2log(BF) (with log the natural logarithm) gives evidence in favor of the experimental hypothesis that is considered "positive" for values in [2, 5], "strong" in (5, 10], and "very strong" for 2log(BF) > 10. In Table 1, we report log likelihoods relative to the best model, $LL_{rel} = LL_M - LL_{max}$. Strength of evidence for model A relative to model B may be then calculated as 2log(BF) = $2(LL_A - LL_B)$, where LL_A and LL_B are the numbers in the table.

Results

Training

Twelve of the fifteen animals met the criterion on the training values in fewer than 30 sessions. There was a large amount of inter-subject variability in the number of trials needed to reach criterion (range = 411 and 2,204 trials).

Table 1 Model fits

Experiment	Linear model, no difference between distribution	Linear model, different distributions	Log model, no difference between distributions	Log model, different distributions
1	0	−5.39	−21.35	−25.46
2	−6.96	0	−31.63	−20.09

Differences from maximum log likelihood (best fit) for each of four models. Log Bayes factors for comparing any two models may be calculated as twice the difference between the relevant columns, with positive values favoring the first model and negative values the second. Log Bayes factors of >10 are considered strong evidence, while those <2 are considered negligible. Thus, in Experiment 1, models in which w's are drawn from species-specific distributions are less preferred than those positing a single distribution of w's across all species, with the two linear models outperforming the two log models. In contrast, in Experiment 2, results dramatically favor models in which w's are drawn from species-specific distributions, indicating that the underlying distribution of Weber fractions in humans differed significantly from the underlying distribution Weber fractions for nonhumans

Two blue-eyed black lemurs never reached criterion and were moved on to testing after 30 training sessions. Due to experimenter error, a third blue-eyed black lemur was moved on to testing after 8 training sessions (480 trials), despite having only reached 75 % once rather than twice for the pair 1 v. 3.

A 2 (training values: {2,4,8} vs. {1,3,9}) × 4 (species) ANOVA on number of trials to reach criterion revealed no significant main effect of species, $F(3, 15) = 2.16$, NS, or training values, $F(1, 15) = 0.66$, NS. There was, however, a significant interaction between species and training values, $F(3, 15) = 4.77$, $P < 0.05$. This interaction reflects the fact that for two species (mongoose lemurs and rhesus macaques), animals trained on the values 1, 3, and 9 required more trials to reach criterion than the animals trained with the values 2, 4, and 8. In contrast, the other two species (blue-eyed black lemurs and ring-tailed lemurs) showed the opposite pattern. A larger sample size would be necessary to determine whether this interaction reflects meaningful species differences or instead individual differences.

We conducted separate analyses on trials when (1) cumulative surface area was equated across both stimuli and (2) dot size was equated across both stimuli. Overall, we found that accuracy on the training pairs was above chance regardless of whether the cumulative surface area of the stimuli was equated [$M = 64.41$ %, $t(15) = 8.89$, $P < 0.0001$] or the size of the elements in each stimulus was equated [$M = 66.11$ %, $t(15) = 9.59$, $P < 0.0001$]. A paired samples t test comparing trials in which the cumulative surface area of the stimuli was equated to trials on which the size of the elements in each stimulus was equated revealed no significant difference in accuracy, $t(15) = -2.01$, NS.

Testing

Each of the 16 subjects selected the larger value more often than predicted by chance in test (all P values <0.001).

Average accuracy across test sessions ranged from 59 % to 91 % correct ($M = 72$ %, SD = 0.08). To compare average accuracy across the four species, we conducted a one-way ANOVA. Results revealed no significant main effect of species, $F(3, 12) = 0.105$, NS.

For the 8 subjects trained using values 2, 4, and 8, we compared accuracy on previously trained pairs (2 vs. 4 and 4 vs. 8) to novel pairs with the same ratio (1 vs. 2, 3 vs. 6). A paired samples t test revealed no significant difference in accuracy, $t(7) = 0.87$, NS. For the 8 subjects trained using values 1, 3, and 9, we compared accuracy on previously trained pairs (1 vs. 3 and 3 vs. 9) to a novel pair with the same ratio (2 vs. 6). A paired samples t test revealed no significant difference in accuracy, $t(7) = 2.14$, NS.

Ratio-dependent performance Like the macaques and ring-tailed lemurs in Merritt et al. (2011), all four species showed a clear effect of ratio on accuracy, consistent with Weber's law (Fig. 2). The 36 number pairs result in 27 unique numerical ratios derived by dividing the larger
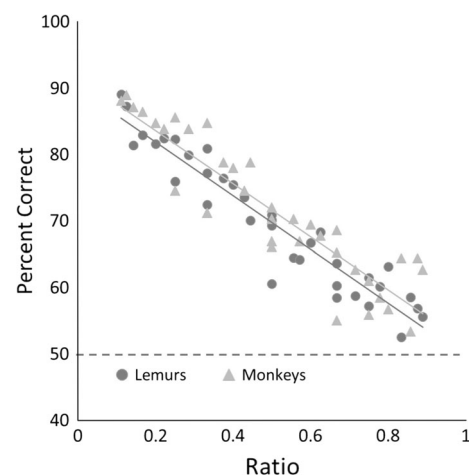

Fig. 2 Accuracy as a function of numerical ratio from Experiment 1. Unlike the results of Merritt et al. (2011), we found no prominent difference in accuracy between lemurs and monkeys

numerosity by the smaller numerosity (values ranging from 0.11 to 0.89). We performed Pearson correlations on the combined data for each species to test for a linear relationship between the 27 unique ratio values and accuracy. Accuracy decreased as the ratio between the two numerosities being compared approached one for all species: mongoose lemurs [$r(25) = -0.94$, $P < 0.001$], blue-eyed black lemurs [$r(25) = -0.94$, $P < 0.001$], ring-tailed lemurs [$r(25) = -0.96$, $P < 0.001$], and rhesus macaques [$r(25) = -0.96$, $P < 0.001$]. In contrast, we found no systematic relationship between ratio and median response time for correct trials. Median RT decreased as the ratio between the two numerosities approached one only for blue-eyed black lemurs [$r(25) = -0.39$, $P < 0.05$]. For the remaining species, there was no significant linear relationship between median RT and ratio value mongoose lemurs [$r(25) = -0.01$, NS], ring-tailed lemurs [$r(25) = -0.32$, NS], and rhesus macaques [$r(25) = -0.21$, NS].

To test for ratio dependence within the small number range, we separately examined ratio effects for pairs of values less than or equal to four (Fig. 4). This resulted in 5 unique numerical ratio values (0.25, 0.33, 0.50, 0.67, and 0.75). We performed Pearson correlations on the combined data for each species to test for a linear relationship between the 5 unique ratios in this range and accuracy. We found a significant effect of ratio, such that accuracy decreased as the ratio between the two numerosities being compared approached one for all species: mongoose lemurs [$r(3) = -0.99$, $P < 0.01$], blue-eyed black lemurs [$r(3) = -0.91$, $P < 0.05$], ring-tailed lemurs [$r(3) = -0.99$, $P < 0.01$], and rhesus macaques [$r(3) = -0.92$, $P < 0.05$].

Quantifying numerical sensitivity There was a significant positive linear relationship between w_{Linear} and w_{Log}, $r(14) = 1.00$, $P < 0.001$. To quantify uncertainty in our fits of w_{Linear} and w_{Log} and perform model comparisons, we performed a hierarchical Bayesian analysis (Wagenmakers et al. 2010). In our case, Bayesian analysis has the advantage of incorporating our prior intuitions about the size and distribution of w's for different individuals in a principled way and potentially offering support for (not mere rejection of) the null hypothesis. The hierarchical Bayesian analysis provides a quantification of the strength of the evidence for the hypothesis of no species difference. Table 1 shows log Bayes Factors, a measure of strength of evidence for each of the models. Species-specific models are less preferred than those positing a single distribution of w's across all species for each choice model, with the two linear models outperforming the two log models. Thus, the Bayesian analysis strongly supports the hypothesis that there are no species differences and that instead the 16 data

points were drawn from a single distribution of w values (see Fig. 3a for individual Weber fractions by species). The analysis also favors the linear model as a description of numerical encoding and choice. When calculating small Weber fractions within a relatively small range, the linear and log models perform equally well. However, when there is a wider range of larger Weber fractions, such as those seen in Experiment 1, the linear model has a higher log likelihood than the log model (Table 1).

A major limitation of our method for species comparisons is that we were only able to test four individuals per species, which, although a typical sample size for comparative primate research, is unlikely to provide enough power to detect small species differences. We directly tested this by performing two types of power analysis based on a simple t test. First we asked how many individuals would be required for the observed differences in mean w's across species to be significant at the $P < 0.05$ level. These numbers ranged from as few as 36 (between *L. catta* vs. *E. mongoz*) to as large as 3,200 (*L. catta* vs. *M. mulatta*). This analysis assumes that the larger sample would maintain the same mean and variance observed in the small samples that we obtained for each species. Our second approach was to ask how large the difference in w between species would have to be, given our sample sizes, to be detected with a 10 % chance of Type II error (false negative). This analysis suggested that the difference in w would need to be larger than 0.53 in the best case (*L. catta* vs. *M. mulatta*) and larger than 1.43 in the worst case
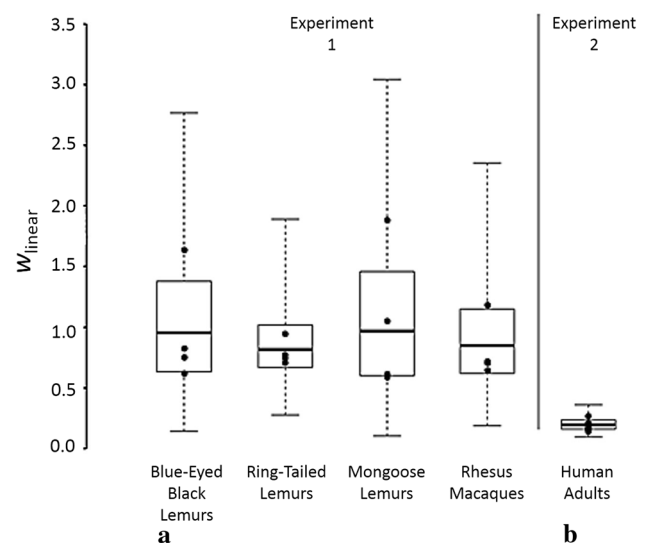

Fig. 3 Weber fractions (w_{Linear}) for **a** the three lemur species and rhesus macaques from Experiment 1, **b** adult humans from Experiment 2. *Black circles* represent medians of posterior distributions for individual Weber fractions. Boxplots show quantiles of posterior distributions for species-specific Weber fractions in the linear model. Distance between whiskers represents the 95 % credible interval for the species-specific w

(*L. catta* vs. *E. mongoz*) to be detectable with our sample size. To summarize, our results suggest that the hypothesis of no species difference is strongly supported given the current data, but also suggest that much larger samples may be needed to provide evidence of small differences, should they exist.

Surface area   To determine whether subjects were using number or cumulative surface area to order the stimuli, we conducted separate analyses on trials when 1) cumulative surface area was equated across both stimuli and 2) dot size was equated across both stimuli. Overall, we found that accuracy was above chance regardless of whether the cumulative surface area of the stimuli was equated [$M = 70.04$ %, $t(15) = 14.01$, $P < 0.0001$] or the size of the elements in each stimulus was equated [$M = 70.45$ %, $t(15) = 13.76$, $P < 0.0001$]. We calculated a surface area difference score for each animal (w_{Linear} for trials in which cumulative surface area was equated across both stimuli subtracted from w_{Linear} for trials in which dot size was equated). On average, difference scores did not differ significantly from zero, $t(15) = 0.57$, NS, indicating no consistent effect of cumulative surface area congruence on Weber fraction across subjects.

To confirm that this pattern did not differ by species, we conducted a one-way ANOVA. Results revealed no significant main effect of species, $F(3, 12) = 0.425$, NS. One macaque and one mongoose lemur, however, had large positive difference scores (0.37 and 0.29, respectively) indicating that their performance was superior when cumulative surface area was congruent with numerosity. In addition, two macaques and one blue-eyed black lemur had large negative difference scores (macaques: -0.28, -0.26; blue-eyed black lemur: -0.49), indicating that their performance was superior when surface area for the two numerosities compared was equal.

To assess the changing effect of surface area over time, we calculated a difference score for the first and last block of 360 trials of testing for each individual. A Wilcoxon signed-rank test revealed that there was no significant difference between difference scores for the first and last testing block, $z = 0.71$, NS, indicating that the effect of surface area congruence did not change over the course of testing.

The effect of training on numerical acuity   To assess the effect of experience on numerical sensitivity, we compared w_{Linear} for the first and last block of the three 360 trial-test blocks. A Wilcoxon signed-rank test revealed that there was a significant difference between w_{Linear} for the first and last testing block; $z = -2.07$, $P < 0.05$. Thirteen of the 16 individuals showed a decrease in Weber fraction from the first block of testing to the last block of testing, indicating

an improvement in numerical discrimination over repeated testing.

Discussion

In summary, in Experiment 1, we found no species difference between macaques and lemurs in numerical acuity as measured by w derived from a numerosity ordering task. This differs from a previous finding by Merritt et al. (2011), in which lemurs performed less accurately than previous reports for macaques tested on the same type of task. The discrepancy highlights the importance of using identical task parameters and training procedures in comparative research. A second finding from Experiment 1 was that the linear model outperformed the log model. This appeared to be due to the wide range of Weber fractions across individuals. A third finding was that ratio dependence held across the full range of numerical values including small values often thought to be in a separate subitizing range. A fourth finding was that we found no systematic effect of surface area in training or test and no change in reliance on surface over the course of testing. A caveat is that we did not control for other continuous stimulus properties such as density and future work should incorporate this additional control. Finally, we found a small but systematic improvement in Weber fraction over the course of testing suggesting that the approximate number system may be malleable (DeWind and Brannon, 2012).

Experiment 2

The goal of Experiment 2 was to provide a comparison between humans and nonhuman primates using the same task, range of numerosities, and stimulus controls from Experiment 1. While a number of papers have reported Weber fractions for adult humans in similar tasks (e.g., Cantlon and Brannon 2006; DeWind and Brannon 2012; Gilmore et al. 2011; Halberda and Feigenson 2008; Piazza et al. 2004; Pica et al. 2004), we conducted Experiment 2 to assess Weber fraction with a similar procedure as that used in Experiment 1 and with the identical stimuli to allow a comparison between humans and nonhuman primates.

Subjects

Ten adult human participants were tested with the same descending ordinal task used in Experiment 1 (mean age 24.51 years, SD 10.58, 4 males). Data for each participant were collected in a single 1-h session. All participants were compensated $10 for completion of the task.

Task stimuli and procedure

Pilot data revealed that when humans were tested with unlimited time to make a response, their performance was at ceiling and thus w could not be calculated. Consequently, stimulus pairs were presented for 100 ms. The task was otherwise identical to that described in Experiment 1 with the following minor exceptions: (a) no auditory feedback or food rewards were provided, (b) Stimuli were presented for 100 ms and then replaced with gray 200×200 pixel placeholders, (c) responses were made via a mouse click on the gray square that replaced that array. These modifications were necessary to make the task difficult enough for humans to avoid a ceiling effect. Finally, (d) participants were tested on a single session of 406–648 trials with pairs of the numerosities 1 through 9, and they were verbally instructed to select the larger numerical value.

Results

All participants selected the larger value more often than predicted by chance (all P values <0.001). Overall accuracy ranged from 87 to 95 % correct ($M = 92$ %, SD = 0.03). There was a clear effect of ratio on accuracy, consistent with Weber's law. Accuracy decreased as the ratio between the two numerosities being compared approached one, $R^2 = -0.84$, $P < 0.001$. In contrast to the results for nonhuman primates in Experiment 1, we found a significant positive correlation between ratio and response time for correct trials, $R^2 = 0.92$, $P < 0.001$.

Quantifying numerical sensitivity

Participants' w_{Linear} ranged from 0.16 to 0.33 ($M = 0.21$, SD = 0.05) (Fig. 3b). Using a hierarchical Bayesian analysis to quantify the strength of the evidence for the hypothesis of no species difference between the nonhuman primates in Experiment 1 and the humans in Experiment 2, we found that the species-specific models dramatically outperformed models that posit a single distribution of w's across both human and nonhuman primates. In addition, the linear model outperformed the log model. Thus, the Bayesian analysis strongly supports the hypothesis that the human w values were drawn from a different distribution of w values than the nonhuman primates tested in Experiment 1.

Surface area

Accuracy was above chance regardless of whether the cumulative surface area of the stimuli was equated [$M = 91.40$ %, $t(9) = 37.81$, $P < 0.0001$] or the size of the elements in each stimulus was equated [$M = 93.33$ %, $t(9) = 63.93$, $P < 0.0001$]. A surface area difference score was calculated for each participant as in Experiment 1 (w_{Linear} for trials in which cumulative surface area was equated across both stimuli subtracted from w_{Linear} for trials in which dot size was equated). On average, difference scores differed significantly from zero, $t(9) = 2.37$, $P < 0.05$, indicating that participants had lower Weber fractions when cumulative surface area was congruent with numerosity than when cumulative area was equated.

Discussion

Comparing the results of Experiments 1 and 2 demonstrates a clear difference between humans and nonhuman primates, indicating significantly better numerical acuity in adult humans compared to macaques or lemurs. This was true despite the modifications to the procedure that we used to avoid a ceiling effect in humans, indicating that the difference between the numerical acuity of adult humans and the numerical acuity of macaques and lemurs is likely greater than that reported here.

In addition, the participants in Experiments 1 and 2 differed in the effect that surface area congruence had on performance. While nonhuman primates showed no systematic bias, human participants showed better numerical acuity when cumulative surface area was congruent with numerosity than when cumulative area was equated. Gebuis and Reynvoet (2012) have also reported that adult humans' performance on a numerical estimation task was influenced by the visual properties of the dot arrays. Thus, the results presented in Experiment 1 are especially surprising given the differences in the visual systems of lemurs and monkeys. Macaques, like apes and humans, are trichromats. All three lemur species have photopigments consistent with dichromacy (Jacobs and Deegan 1993, 2003; Leonhardt et al. 2008), although there is behavioral evidence for color discrimination consistent with trichromacy in ring-tailed lemurs (Blakeslee and Jacobs 1985) and black lemurs (Gosset and Roeder 2000). Lemurs have also been shown to have lower visual acuity than Old World monkeys (Veilleux and Kirk 2009). Despite these differences, we found no numerical acuity differences between lemurs and macaques using visual stimuli. This is likely due to the high contrast used in our visual stimuli.

General discussion

Our first finding was that prosimian primates and rhesus monkeys showed quantitative equivalence in numerical discrimination despite approximately 50 million years of evolutionary divergence and significant differences in their

visual system and visual acuity. Thus, our results suggest that previous findings of inferior numerical discrimination in lemurs relative to macaques were due to differences in training or task parameters and did not reflect ability differences between macaques and lemurs (Merritt et al. 2011). These findings highlight the need to conduct comparative experiments with common methods and conditions (Agrillo and Miletto Petrazzini 2012). A similar debate regarding the numerical abilities of elephants has arisen and may be attributed to differences in testing procedures. Irie-Sugimoto et al. (2009) and Irie and Hasegawa (2012) reported that elephants' performance on a quantity discrimination task is not affected by numerical ratio. In contrast, Perdue et al. (2012) recently found ratio-dependent performance in the quantity discriminations of elephants.

These findings should also be interpreted in light of a recent study that suggests that there are no qualitative differences in the numerical discrimination abilities of rhesus macaques and pigeons (Scarf et al. 2011). In that study, pigeons successfully transferred an ordinal numerical rule from the values 1–4 to the values 5–9 much like rhesus monkeys (Brannon and Terrace 1998). Pigeons showed ratio-dependent performance, and although the authors did not model Weber fractions, performance looked very similar across these two species despite over 200 million years of divergent evolutionary history.

Agrillo et al. (2012a, b) also reported qualitative similarities in numerical representation systems between phylogenetically distant species. They compared the numerical discrimination ability of humans and guppies using small and large value and found suggestive evidence that two distinct systems underlie quantity discrimination in both humans and fish: For both adult humans and guppies, accuracy at selecting the larger numerical value was affected by ratio for large values, but not for the numbers 1–4. Evidence for separate systems for representing large and small numerical values has been reported for human infants (Feigenson and Carey 2003, Feigenson 2005; Feigenson et al. 2002) and for a small set of nonhuman species (Agrillo et al. 2008, 2012a, b Hauser et al. 2000; Uller et al. 2003; Uller and Lewis 2009; Wood et al. 2008). In contrast, we found no evidence for separate systems for representing large and small numerical values: Accuracy at selecting the larger numerical value was affected by ratio for all values, including the numbers 1–4 (Fig. 4).

Although many previous studies have reported ratio-dependent number discrimination in different animal species in a wide variety of tasks, rarely have multiple species been tested with the same task and stimuli to allow for a quantitative comparison. Rather than simply failing to reject the null hypothesis that the four species did not differ in numerical acuity, our Bayesian analytic approach
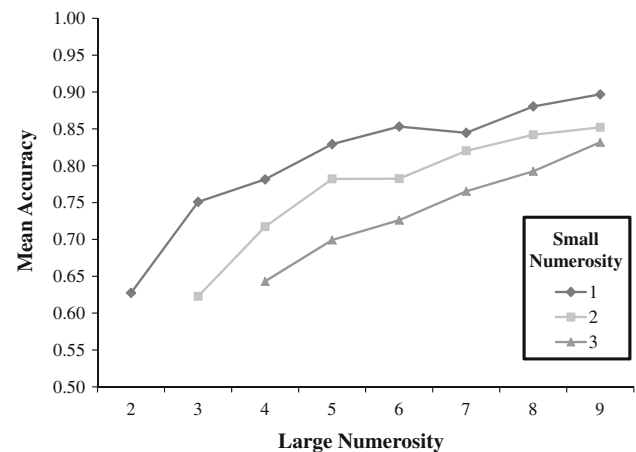

Fig. 4 Accuracy for pairs involving n_S 1, 2, and 3 as a function of n_L from Experiment 1, where n_S and n_L are the small and large values being compared, respectively. We found no evidence of a separate system of numerical discrimination for small numerical values

provides strong evidence against pronounced species differences. We compared the strength of two different models which posit that (1) the Weber fractions of all subjects, regardless of species, were drawn from the same distribution and (2) that there are species differences in Weber fraction for numerical acuity. Our analyses revealed strong support for models that assume a single distribution of ws across all nonhuman species compared to models in which each non-human primate species was assumed to have a unique distribution of ws.

In addition to quantitatively comparing the numerical abilities of lemurs and monkeys, a second goal for the present study was to compare numerical acuity in three different lemur species to assess whether numerical acuity might have been subject to selection pressures exerted by variations in social structure or feeding ecology. Our results suggest that despite the large variation in social structure, home range size, and diet in these three lemur species, there is much more variability within a species than between species. The lack of species differences here is striking in light of other studies that compared closely related species and found differences that were consistent with evolutionary predictions (Bond et al. 2003; MacLean et al. 2008; Platt et al. 1996; Rosati et al. 2007; Tomasello et al. 2001). In prior research, ring-tailed lemurs have often exceeded the performance of other lemur species and performed more similarly to macaques than other lemurs. For example, Sandel et al. (2011) showed that ring-tailed lemurs, but not mongoose lemurs, black lemurs, or ruffed lemurs, spontaneously exploited social cues regarding a competitor's visual orientation. These similarities may reflect cognitive adaptations for living in large social groups rather than a common ancestral trait. Similarly, MacLean et al. (2008) demonstrated that ring-

tailed lemurs show advanced transitive reasoning skills relative to mongoose lemurs, consistent with the prediction that animals that live in complex social structures have evolved cognitive adaptations for tracking dominance relationships.

In contrast, MacLean et al. (2013) recently reported that differences in group size in 6 prosimian species predicted performance on a measure of social cognition, but not a nonsocial inhibitory control task. Our finding is consistent with that data given that our task was quantitative and did not tap social cognition. It must be emphasized, however, that power analyses indicated that our sample sizes may not have been large enough to detect species differences in Weber fraction, were they to exist, due to large within-species variability. Obviously, samples of 3,200 or even 36 individuals per species are not feasible for comparative cognitive research with prosimian primates. Thus, a caution is that small sample sizes with comparative research will only be useful insofar as they will uncover relatively large species differences that surpass individual variability within a species.

Another notable aspect of our results is that individual variability in numerical acuity is marked in nonhuman primates. These findings are consistent with the large variability in w seen in adult human samples (see Piazza and Izard 2009 for a review). For example, w ranges in human samples have been as wide as 0.18–0.76 (DeWind and Brannon 2012) or even 0.22–1.50 (Gilmore et al. 2011). Our own study (Experiment 2), using limited stimulus exposure but otherwise identical parameters as that used to test lemurs and macaques in Experiment 1, also revealed individual variability with a range of 0.16–0.33. Future studies should explore whether this individual variation reflects differences in general cognitive ability or more specialized quantitative abilities by correlating performance across a wide range of tasks.

In some sense, the lack of species differences coupled with the large amount of individual variability poses a conundrum. On the one hand, selective pressures for the existence of the ANS seem to have been so extreme and ubiquitous that there is little variation across species. On the other hand, there is considerable variability across individuals within each species. If the ANS is a cognitive mechanism evolved under strong selective pressure and closely related to fitness, this variability may be surprising. Comparisons of heritability indicate that traits under stronger directional selection (e.g., fitness-related traits) have reduced genetic variance compared with traits with more distant connections to fitness (e.g., Fisher 1930; Mousseau and Roff 1987; Roff and Mousseau 1987). However, an alternative hypothesis that may resolve this discrepancy is that traits closely related to fitness are expected to have higher additive genetic and nongenetic

variability due to the number of genetic and environmental events that affect complex, behavioral traits (e.g., Houle 1991, 1992). The considerable variability of acuity in the ANS within species, along with the improvement across training, may suggest that environmental factors and experience within the lifespan play an important role in numerical sensitivity.

In conclusion, while research over the last century has revealed numerical abilities in a wide range of animal species, rarely has a single task been used to allow quantitative comparisons between species. By employing a common set of task parameters and stimuli, our experiments revealed that there is a surprising amount of overlap in numerical acuity between nonhuman primate species, while some aspects of nonhuman primate cognition appear to have been shaped by the socioecological milieu in which the species has evolved the approximate number system appears to be insensitive to these variations.

Conflict of interest The authors declare that they have no conflict of interest.

Ethical standards The experiments reported here comply with the current laws of the United States of America.

References

Agrillo C, Miletto Petrazzini ME (2012) The importance of replication in comparative psychology: the lesson of elephant quantity judgments. Front Psychol 3:181

Agrillo C, Dadda M, Serena G, Bisazza A (2008) Do fish count? Spontaneous discrimination of quantity in female mosquitofish. Anim Cogn doi. doi:10.1007/s10071-008-0140-9

Agrillo C, Miletto Petrazzini ME, Tagliapietra C, Bisazza A (2012a) Inter-specific differences in numerical abilities among teleost fish. Front Psychol 3:483. doi:10.3389/fpsyg.2012.00483

Agrillo C, Piffer L, Bisazza A, Butterworth B (2012b) Evidence for two numerical systems that are similar in humans and guppies. PLoS ONE 7(2):e31923

Anderson JR, Mitchell RW (1999) Macaques but not lemurs co-orient visually with humans. Folia Primatol 70:17–22

Beran MJ (2008) The evolutionary and developmental foundations of mathematics. PLoS Biol 6:e19

Beran MJ, Johnson-Pynn JS, Ready C (2008) Quantity representation in children and rhesus monkeys: linear versus logarithmic scales. J Exp Child Psychol 100:225–233

Beran MJ, Decker S, Schwartz A, Schultz N (2011a) Monkeys (*Macaca mulatta and Cebus apella*) and human adults and children (*Homo sapiens*) compare subsets of moving stimuli based on numerosity. Front Psychol 2:61. doi:10.3389/fpsyg.2011.00061

Beran MJ, Johnson-Pynn JS, Ready C (2011b) Comparing children's (*Homo sapiens*) and chimpanzees' (*Pan troglodytes*) quantity judgments of sequentially presented sets of items. Curr Zool 57:419–428

Best NG, Cowles MK, Vines K (1995) CODA* convergence diagnosis and output analysis software for Gibbs sampling output Version 0.30. MRC Biostatistics Unit, Cambridge

Blakeslee B, Jacobs GH (1985) Color vision in the ring-tailed lemur (*Lemur catta*). Brain Behav Evol 26:154–166

Bond AB, Kamil AC, Balda RP (2003) Social complexity and transitive inference in corvids. Anim Behav 65:479–487

Brannon EM (2006) The representation of numerical magnitude. Curr Opin Neurobiol 16:222–229

Brannon EM, Terrace HS (1998) Ordering of the numerosities 1 to 9 by monkeys. Science 282(5389):746–749

Brannon EM, Terrace HS (2000) Representation of the numerosities 1–9 by rhesus macaques (*Macaca mulatta*). J Exp Psychol Anim Behav Process 26(1):31–49

Cantlon JF, Brannon EM (2006) Shared system for ordering small and large numbers in monkeys and humans. Psychol Sci 17(5):401–406

Cantlon JF, Brannon EM (2007) Basic math in monkeys and college students. PLoS Biol 5(12):e328

Cowles MK, Carlin BP (1996) Markov chain Monte Carlo convergence diagnostics: a comparative review. J Am Stat Assoc 91(434):883–904

Curtis DJ, Zaramody A (1998) Group size, home range use, and seasonal variation in the ecology of *Eulemur mongoz*. Int J Primatol 19(5):811–835

Deppe AM, Wright PC, Szelistowski WA (2009) Object permanence in lemurs. Anim Cogn 12(2):381–388

DeWind NK, Brannon EM (2012) Malleability of the approximate number system: effects of feedback and training. Front Hum Neurosci 6:68. doi:10.3389/fnhum.2012.00068

Erhart EM, Overdorff DJ (2008) Spatial memory during foraging in prosimian primates: *Propithecus edwardsi* and *Eulemur fulvus rufus*. Folia Primatol 79:185–196

Feigenson L (2005) A double dissociation in infants' representation of object arrays. Cognition 95:B37–B48

Feigenson L, Carey S (2003) Tracking individuals via object-files: evidence from infants' manual search. Dev Sci 6:568–584

Feigenson L, Carey S, Hauser M (2002) The representations underlying infants' choice of more: object-files versus analog magnitudes. Psychol Sci 13:150–156

Fichtel C, Kappeler PM (2010) Human universals and primate symplesiomorphies: Establishing the lemur baseline. In: Kappeler PM, Silk J (eds) Mind the gap: tracing the origins of human universals. Springer, Heidelberg

Fisher RA (1930) The genetical theory of natural selection. Oxford University Press, Oxford

Gebuis T, Reynvoet B (2012) The role of visual information in numerosity estimation. PLoS ONE 7(5):e37426. doi:10.1371/journal.pone.0037426

Gelman A, Hill J (2007) Data analysis using regression and multilevel/hierarchical models. Cambridge University Press, Cambridge

Gelman A, Carlin JB, Stern HS, Rubin DB (2004) Bayesian data analysis. Chapman & Hall/CRC, Boca Raton, FL

Genty E, Palmier C, Roeder J (2004) Learning to suppress responses to the larger of two rewards in two species of lemurs, *Eulemur fulvus* and *E. macaco*. Anim Behav 67(5):925–932

Gilmore C, Attridge N, Inglis M (2011) Measuring the approximate number system. Q J Exp Psychol 64(11):2099–2109

Gosset D, Roeder J (2000) Colour and shape discrimination in black lemurs (*Eulemur macaco*). Folia Primatol 71:173–176

Halberda J, Feigenson L (2008) Developmental change in the acuity of the "Number Sense": the approximate number system in 3-, 4-, 5-, 6-year-olds and adults. Dev Psychol 44(5):1457–1465

Hanus D, Call J (2007) Discrete quantity judgments in the great apes (*Pan paniscus, Pan troglodytes, Gorilla gorilla, Pongo pygmaeus*): the effect of presenting whole sets versus item-by-item. J Comp Psychol 121(3):241–249

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

Horvath JE, Weisrock DW, Embry SL, Fiorentino I, Balhoff JP, Kappeler P, Wray GA, Willard HF, Yoder AD (2008) Development and application of a phylogenomic toolkit: resolving the evolutionary history of Madagascar's lemurs. Genome Res 18:489–499

Houle D (1991) Genetic covariance of fitness correlates: what genetic correlations are made of and why it matters. Evolution 45:630–648

Houle D (1992) Comparing the evolvability and variability of quantitative traits. Genetics 130:195–204

Irie N, Hasegawa T (2012) Summation by Asian elephants (*Elephas maximus*). Behav Sci 2:50–56. doi:10.3390/bs2020050

Irie-Sugimoto N, Kobayashi T, Sato T, Hasegawa T (2009) Relative quantity judgment by Asian elephants (*Elephas maximus*). Anim Cogn 12:193–199. doi:10.1007/s10071-008-0185-9

Jacobs GH, Deegan JF (1993) Photopigments underlying color vision in ringtail lemurs (*Lemur catta*) and brown lemurs (*Eulemur fulvus*). Am J Primatol 30(3):243–256

Jacobs GH, Deegan JF (2003) Diurnality and cone photopigment polymorphism in strepsirrhines: examination of linkage in *Lemur catta*. Am J Phys Anthropol 122(1):66–72

Jones SM, Brannon EM (2012) Prosimian primates show ratio dependence in spontaneous quantity discriminations. Front Psychol 3:550. doi:10.3389/fpsyg.2012.00550

Jordan K, Brannon EM (2006a) A common representational system governed by Weber's law: nonverbal numerical similarity judgments in six-year-old children and rhesus macaques. J Exp Child Psychol 95:215–229

Jordan K, Brannon EM (2006b) Weber's Law influences numerical representations in rhesus macaques (*Macaca mulatta*). Anim Cogn 9(3):159–172

Kass R, Raftery A (1995) Bayes factors. J Am Stat Assoc 90:773–795

Leonhardt SD, Tung J, Camden JB, Leal M, Drea CM (2008) Seeing red: behavioral evidence of trichromatic color vision in strepsirrhine primates. Behav Ecol 20(1):1–12

Lewis KP, Jaffe S, Brannon EM (2005) Analog number representations in mongoose lemurs (*Eulemur mongoz*): evidence from a search task. Anim Cogn 8(4):247–252

Luhrs ML, Dammhahn M, Kappeler PM, Fichtel C (2009) Spatial memory in the grey mouse lemur (*Microcebus murinus*). Anim Cogn 12:599–609

MacLean E, Merritt D, Brannon EM (2008) Social complexity predicts transitive reasoning in prosimian primates. Anim Behav 76:479–486

MacLean E, Sandel A, Bray J, Oldenkamp R, Reddy R, Hare B (2013) Group size predicts social but not nonsocial cognition in lemurs. PLos One 8(6):e66359. doi:10.1371/journal.pone.0066359

Merritt D, MacLean EL, Jaffe S, Brannon EM (2007) A comparative analysis of serial ordering in ring-tailed lemurs (*Lemur catta*). J Comp Psychol 121(4):363–371

Merritt D, MacLean E, Crawford JC, Brannon EM (2011) Numerical rule-learning in ring-tailed lemurs (*Lemur catta*). Front Comp Psychol 2(23):1–9

Mousseau TA, Roff DA (1987) Natural selection and the heritability of fitness components. Heredity 59:181–197

Perdue BM, Talbot CF, Stone A, Beran MJ (2012) Putting the elephant back in the herd: Elephant relative quantity judgments match those of other species. Anim Cogn 15(5):955–961. doi:10.1007/s10071-012-0521-y

Piazza M, Izard V (2009) How humans count: numerosity and the parietal cortex. Neuroscientist 15(3):261–273

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

Pica P, Lemer C, Izard V, Dehaene S (2004) Exact and approximate arithmetic in an Amazonian indigene group. Science 306(5695):499–503

Picq JL (1993) Radial maze performance in young and aged grey mouse lemurs (Microcebus murinus). Primates 24:223–226

Picq JL (2007) Aging affects executive functions and memory in mouse lemur primates. Exp Gerontol 42:223–232

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

Roff DA, Mousseau TA (1987) Quantitative genetics and fitness: lessons from Drosophila. Heredity 58:103–118

Rosati AG, Stevens JR, Hare B, Hauser MD (2007) The evolutionary origins of human patience: temporal preferences in chimpanzees, bonobos, and human adults. Curr Biol 17:1663–1668

Ruiz A, Gomez JC, Roeder JJ, Byrne RW (2009) Gaze following and gaze priming in lemurs. Anim Cogn 12:427–434

Rumbaugh DM, Arnold RC (1971) Learning: a comparative study of lemur and cercopithecus. Folia Primatol 14:154–160

Sandel AA, MacLean E, Hare B (2011) Evidence from four lemur species that ringtailed lemur social cognition converges with that of haplorhine primates. Anim Behav 81:925–931

Santos LR, Barnes JL, Mahajan N (2005a) Expectations about numerical events in four lemur species (Eulemur fulvus, Eulemur mongoz, Lemur catta and Varecia rubra). Anim Cogn 8(4):253–262

Santos LR, Mahajan N, Barnes JL (2005b) How prosimian primates represent tools: experiments with two lemur species (Eulemur fulvus and Lemur catta). J Comp Psychol 119(4):394–403

Sauther ML, Sussman RW, Gould L (1999) The socioecology of the ringtailed lemur: thirty-five years of research. Evol Anthro 8(4):120–132

Savage A, Snowdon CT (1989) Apples and oranges: the pitfalls of comparative intelligence. Behav Brain Sci 12(3):605–606

Scarf D, Hayne H, Colombo M (2011) Pigeons on par with primates in numerical competence. Science 334(6063):1664

Shepherd SV, Platt ML (2008) Spontaneous social orienting and gaze following in ringtailed lemurs (Lemur catta). Anim Cogn 11:13–20

Tattersall I (1982) The primates of Madagascar. Columbia University Press, New York

Tomasello M, Hare B, Fogleman T (2001) The ontogeny of gaze following in chimpanzees (Pan troglodytes) and rhesus macaques (Macaca mulatta). Anim Behav 61:335–343

Uller C, Lewis J (2009) Horses (Equus caballus) select the greater of two quantities in small numerical contrasts. Anim Cogn 12(5):733–738

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

Veilleux CC, Kirk EC (2009) Visual acuity in the cathemeral strepsirrhine Eulemur macaco flavifrons. Am J Primatol 71(4):343–352

Volampeno MSN, Masters JC, Downs CT (2011) Home range size in the blue-eyed black lemur (Eulemur flavifrons): a comparison between dry and wet seasons. Mamm Biol 76(2):157–164

Wagenmakers EJ, Lodewyckx T, Kuriyal H, Grasman R (2010) Bayesian hypothesis testing for psychologists: a tutorial on the Savage–Dickey method. Cogn Psychol 60:158–189

Wood JN, Hauser MD, Glynn DD, Barner D (2008) Free-ranging rhesus monkeys spontaneously individuate and enumerate small numbers of non-solid portions. Cognition 106:207–221

Wynn K (1992) Addition and subtraction by human infants. Nature 358:749–750

Yoder AD (2007) Lemurs: a quick guide. Curr Biol 17(20):866–868