Animal Cognition (2019) 22:1197–1202
https://doi.org/10.1007/s10071-019-01316-8

COMMENTARY

Leveraging brain–body scaling relationships for comparative studies

Daniel J. Horschler[1,2] · **Evan L. MacLean**[1,2,3]

Received: 30 July 2019 / Revised: 15 September 2019 / Accepted: 30 September 2019 / Published online: 11 October 2019

Abstract

In Horschler et al. (Anim Cognit 22(2):187–198, 2019), we found that two components of executive function (short-term memory and self-control) were strongly associated with estimated absolute brain weight across dog breeds, and argued that dogs present a powerful model for studying evolutionary links between cognition and neuroanatomy due to their extraordinary degree of intraspecific morphological variation. In a commentary on this work, Montgomery (Anim Cognit, 2019) raises concerns about the practice of estimating brain weights from brain–body scaling relationships. Montgomery explores the practical significance of this approach, ultimately concluding that such estimations should be avoided. In this response, we point out some limitations of the analyses presented by Montgomery and consider his conclusions in light of these issues. We then explore the extent to which body weight serves as a valid proxy for brain weight under varying conditions. Through simulations, we show that the consequences of using body weight as a proxy for brain weight depend on parameters including effect size, the correlation between brain and body weight, and the variance in brain and body weight within a sample. Under conditions approximating those in Horschler et al. (Anim Cognit 22(2):187–198, 2019), we find that body weight is a reliable proxy for brain weight, and that statistical results from models using either brain weight or body weight as predictor variables are highly convergent. Nonetheless, we wholeheartedly agree with Montgomery that empirical data on brain weight, structure, and cellular composition will be critical for creating new opportunities to investigate the relationships between neuroanatomy and cognition in dogs.

Keywords Allometry · Body size · Brain size · Cognition · Dogs · Breed differences

In his commentary on Horschler et al. (2019), Montgomery (2019) raises concerns about using brain weight estimates derived from brain–body scaling relationships in comparative studies. He notes the potential for error arising from this approach and ultimately concludes that it should be avoided. While we agree with many of the points in Montgomery's commentary, we differ in our interpretation of some critical issues, and argue that estimates of brain weight based on body weight can be effectively employed when certain conditions are met. Below, we first address some specific points from Montgomery's commentary, and then present analyses using empirical and simulated datasets to assess the conditions under which body weight can serve as a reliable proxy for brain weight in comparative studies.

In Horschler et al. (2019), we used the brain–body weight scaling relationship in dogs identified by Bronson (1979) to derive breed-average brain weight estimates for 74 breeds with known breed-average body weights. We then explored the relationships between estimated breed-average brain weight and a range of cognitive measures from the citizen science project *Dognition*. In his commentary, Montgomery notes that despite the high percentage of variance in brain weight explained by body weight (R^2 = ~ 0.92), the error observed in brain weight estimates for some breeds in Bronson's original dataset is still high (mean = 5.3%, max = 13.9%). First, it is important to interpret this degree of error within the context of the range of body weights and estimated brain weights in our sample. Within our sample, body weights ranged from 3.4 kg to 59.8 kg (an over 17-fold difference) and estimated brain weights ranged from 55.1 to 121.7 g (an over 2-fold difference). Although the percentage error in these brain weight estimates is non-negligible, the

✉ Daniel J. Horschler
horschler@email.arizona.edu

[1] School of Anthropology, University of Arizona, Tucson, AZ 85719, USA

[2] Cognitive Science Program, University of Arizona, Tucson, AZ 85719, USA

[3] Department of Psychology, University of Arizona, Tucson, AZ 85719, USA

effects of this error are greatly minimized when there is high variance in brain weight (an issue we illustrate empirically below). Second, Carreira (2016) showed that the scaling relationship identified in Bronson (1979) accurately predicted brain weight from body weight in an independent sample of 69 dogs including brachycephalic, mesocephalic, and dolichocephalic individuals representing a similarly wide range of brain weights (41.2–116.2 g). Therefore, we believe it is unlikely that error induced from Bronson's scaling equation had a significant influence on the relationships we observed.

Montgomery also presents new analyses using known brain weights and body weights from Bronson (1979) to predict breed-average cognitive measures from the *Dognition* dataset (Stewart et al. 2015). As he notes, it is possible that Bronson's observed brain and body weights may not be representative of true breed averages because of the small sample sizes within some breeds. We share this concern and did not include these or similar analyses in the original paper for this reason. In contrast to Bronson's dataset, the dataset we used to obtain breed-average body weights [Canine Behavioral Assessment & Research Questionnaire (C-BARQ); Hsu and Serpell 2003] includes measures from thousands of dogs and is more likely to be representative at the breed level. For example, genome-wide association studies using breed-average body weight reported in the C-BARQ identify the same genetic variants associated with body size in several other cross-breed studies (MacLean et al. 2019). Comparing breed-average body weights between the Bronson and C-BARQ datasets, we find large discrepancies in many cases (mean difference = 14.8%, max difference = 49.3%). Note, however, that this limitation does not preclude using Bronson's data to describe the scaling relationship between brain and body weight across dogs of different sizes, as Bronson measured both brain weight and body weight from the same individuals to quantify the relationship between these variables (and the scaling relationship has been validated in independent datasets).

Montgomery ultimately argues that estimating brain weight from a scaling relationship adds no additional information to the analyses because estimated brain weights are merely linear transformations of body weight. We agree that the transformation of body weight to estimated brain weight does not do any additional explanatory work, nor did we intend it to. However, there are some benefits of using brain weight estimates rather than body weights in the statistical analyses. For example, the beta coefficient for the predictor variable can be interpreted in units of brain weight rather than body weight. In turn, this facilitates predictions about how cognitive measures may vary across a range of brain weights (rather than body weights) and sets up testable hypotheses for future datasets involving measures of brain weight (which may be obtained independently from measures of body weight). By reporting and analyzing estimated brain weights rather than only body weights, it is also easier to relate our findings to previous work linking brain size to cognition within and across other non-primate mammals [which generally adhere to the same interspecific neural scaling rules (Herculano-Houzel 2012)]. Therefore, while we agree with Montgomery that purely from a modeling perspective estimating brain weight from body weight does little explanatory work, this approach is still valuable in guiding hypotheses for future studies and situating our results within the context of a larger body of literature on relationships between brain size and cognition.

As Montgomery points out, we reported that models using estimated brain weight rather than body weight as the predictor variable generally had lower AIC values, but noted that the differences in AIC were very small (i.e., less than 2). We agree that these comparisons cannot assess whether brain weight or body weight is a better predictor of a dependent measure. However, disentangling the associations between body weight, brain weight, and cognitive measures was not the main aim of our work, and would have been impossible given the data available. Rather, our primary aim was to exploit the strong scaling relationship between brain and body weight to estimate associations between brain weight and the cognitive measures. Nonetheless, even in cases where both brain and body weight are known, disentangling associations with these variables can be challenging due to their strong collinearity (MacLean et al. 2014). Therefore, when making causal inferences, rather than blindly relying on statistical results, researchers must also consider plausibility at a biological level. Given that cognition is critically dependent on the brain, and innumerable studies reveal links between neuroanatomy and cognition (Benson-Amram et al. 2016; Buechel et al. 2017; Deaner et al. 2007; Healy and Krebs 1992; Kotrschal et al. 2013, 2015; Krebs et al. 1989; MacLean et al. 2014; Overington et al. 2009; Reader and Laland 2002; Sol et al. 2008; Sol et al. 2005), there is a reasonable basis for inferring that these associations are driven by brains more so than bodies. Nonetheless, we acknowledge the possibility of other confounding variables (many of which we attempted to control for, including genetic relatedness, training history, perceptual factors due to differences in skull shape, and breed group), as it is always possible that an association is mediated by an unidentified variable not included in the analysis.

Setting aside the issue of causal inference, we now return to the question of whether body weight can serve as a useful proxy for brain weight in comparative studies with dogs, which we address through a series of simulations. The main aim of these simulations was to assess concordance between the results of models using either brain weight, or body weight as a proxy variable for brain weight, in regressions where the dependent measure is associated with brain weight

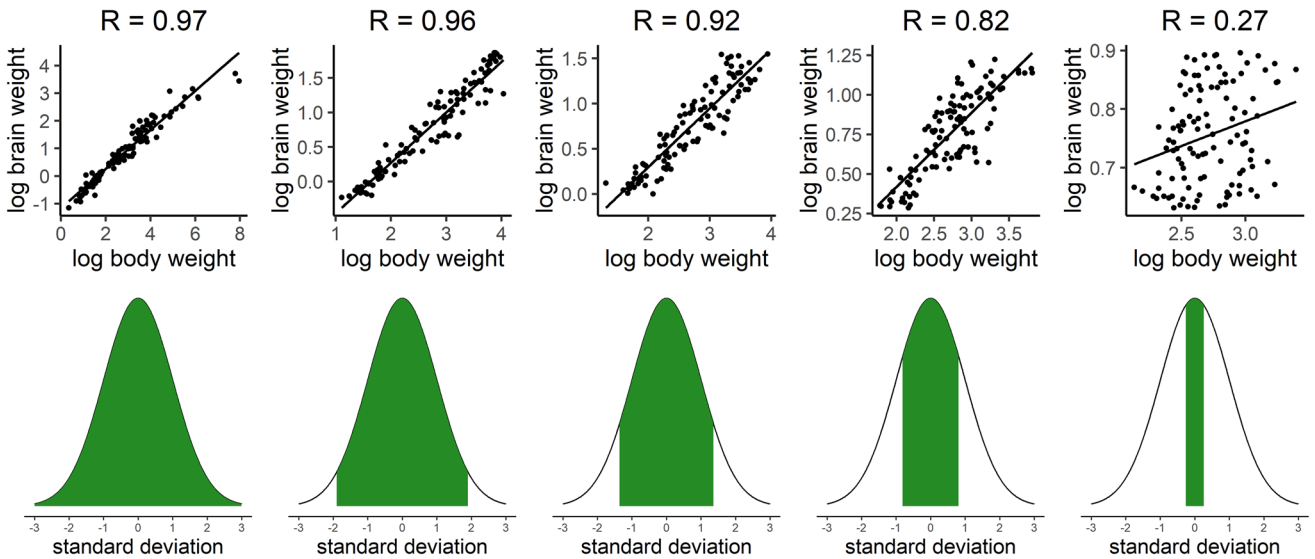

Fig. 1 Correlation between brain and body weight in distributions sampled from increasingly restricted ranges of a parent distribution. The top panels show scatter plots and a linear fit between brain and body weight within each subsample. The bottom panels show the range of the parent distribution (in standard deviation units) from which the data points were sampled

at a given correlation. A key assumption in our original paper was that body weight could serve as a reliable proxy for brain weight if the following two conditions were met: (1) the correlation between brain weight and body weight was high, and (2) the values being predicted were distributed across a sufficiently large range such that the amount of (inevitable) prediction error for individual points would be negligible compared to the variance of the values being estimated. Through the following simulations, we will argue that both conditions were met in the original paper, and, therefore, that body weight could be used as a reliable proxy for brain weight in this dataset.

The first condition is central to the use of any proxy variable because the utility of a proxy variable depends on the strength of its association with the variable it indexes. In the case of brain and body weights, many comparative analyses confirm that these variables obey log-linear scaling relationships, with high R^2 values. But as Montgomery notes, R^2 values are properties of a particular regression, not individual data points. Indeed, the strength of the relationship between brain and body weight—and thus the degree to which body weight can stand in for brain weight—is related to the variance of the distribution. To illustrate this point, we used a recently published dataset on brain–body scaling relationships in 1552 mammalian species (Burger et al. 2019). With this dataset, we divided observations into 11 quantile groups based on brain weight. We then generated five smaller datasets, based on random sampling of the parent dataset, but varying which quantile groups samples were drawn from. The first of these datasets sampled evenly across all quantiles; the second was restricted to observations between the 3rd and 9th quantiles; the third was restricted to observations between the 4th and 8th quantiles; the fourth was restricted to observations between the 5th and 7th quantiles; and the fifth included only observations in the 6th quantile (Fig. 1). Using each of these datasets, we assessed the strength of the correlation between log-transformed brain weight and log-transformed body weight using Pearson correlation. As shown in Fig. 1, the strength of the relationship between brain weight and body weight decreases monotonically across samples drawn from increasingly constrained ranges within the parent distribution, despite the number of observations in each subsample being held constant. Therefore, Fig. 1 illustrates with empirical data that the ranges of variation in samples of brain weight and body weight relate to the strength of the relationship between these variables; brain weight and body weight are more strongly correlated as the ranges get larger. In other words, the reliability of body weight as a proxy for brain weight is contingent on a sufficient range of variation in these variables, a point we return to below.

We next ran a series of simulations to assess the extent to which regressions using brain weight versus body weight as a proxy for brain weight produce convergent results under a range of conditions. Across iterations, we (1) simulated two quantitative variables with the same means and variances, representing brain and body weight.[1] These variables were set to be correlated at R

[1] These variables were simulated to have the same means simply to facilitate comparison of beta coefficients from subsequent models.

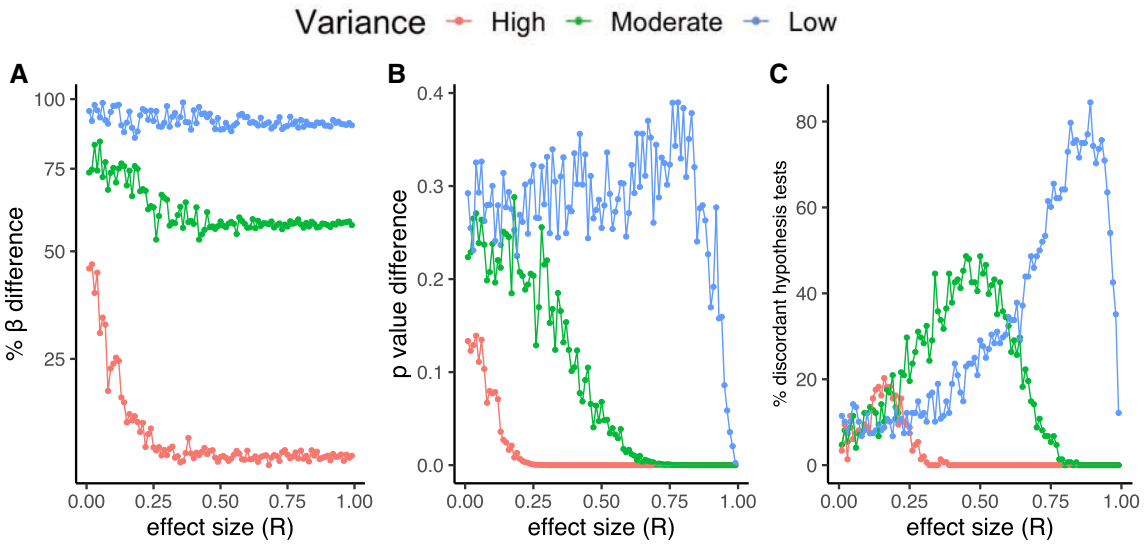

Fig. 2 Agreement between bivariate regressions predicting a dependent measure based on either body weight or brain weight. See text for details

= 0.9, slightly below the correlation coefficient between brain weight and body weight in Bronson's original dataset ($R = 0.96$; Bronson 1979) and between brain volume and body weight in a newly published dataset including 61 dogs representing 33 different breeds after the exclusion of 1 outlier ($R = 0.91$; Hecht et al. 2019). (2) We then generated a third quantitative variable, to be treated as the dependent measure, which was specified to have a fixed correlation with brain weight, without explicitly assigning any correlation between this variable and body weight. Across simulations, we varied the strength with which the dependent measure was correlated with brain weight, testing all values between $R = 0.01$ and $R = 0.99$, in increments of 0.01. (3) We next generated three smaller datasets from this parent distribution using random sampling, but as above, varied which quantile groups these observations were sampled from to generate distributions with high, moderate, and low variance in brain and body weight. For the high-variance distribution, samples were drawn evenly from across 20 quantile groups (1st–100th percentiles). For the moderate-variance distribution, samples were drawn evenly from the 6th through 15th quantiles (26th–75th percentiles). For the low-variance distribution, samples were drawn from the 9th through 12th quantiles (41st–60th percentiles). The total number of observations in each subset was held constant ($N = 200$), regardless of the number of quantile groups samples were drawn from. (4) Lastly, we fit two bivariate regressions. The first model predicted the dependent measure as a function of brain weight; the second predicted the dependent measure as a function of body weight. We compared the results of these models by extracting the following parameters:

- Absolute value of the percentage difference in the β coefficients from the brain weight and body weight models.
- Absolute value of the difference between the p values for the brain weight and body weight β coefficients.
- Whether the significance of null hypothesis tests (Wald test, $a = 0.05$) on the brain and body β coefficients differed between the two models.

In total, we simulated 148,500 regressions, consisting of 500 regressions per dataset ($N = 3$) at each of the different effect sizes ($N = 99$). The results of these simulations are shown in Fig. 2.

Figure 2a shows the mean percentage difference in the β coefficients for brain weight and body weight across datasets with different variances, and across a range of effect sizes (the strength of the correlation between brain weight and the dependent measure in the parent dataset). If both models produced similar results, we expected the β coefficients to be similar, because both the brain and body weight variables were simulated to have the same means and variances. The key takeaways from Fig. 2a are: (1) the differences in β coefficients are dramatically smaller in the high-variance distributions, because brain weight and body weight maintain the highest correlation in this case, as demonstrated with an empirical brain weight and body weight dataset above; (2) for all distributions, the differences in β coefficients are reduced with increasing strength of the association between brain weight and the dependent measure. For the high-variance distribution, the β coefficients for brain weight and body weight become negligibly different once the effect size reaches $R \geq \sim 0.25$. Figure 2b shows a similar effect for the (mean) difference in p values for the β coefficients in the brain weight and body weight models. As

the effect size reaches $R \geq$ ~0.25, the mean difference in p values for the brain weight and body weight β coefficients drops below 0.001. Figure 2c illustrates the percentage of cases in which the null hypothesis would be rejected for one of the two models, but not the other. We can interpret this as a categorical measure of discrepancy between the brain weight and body weight models. Temporarily setting aside differences between the low-, moderate-, and high-variance distributions, some common patterns emerge. First, when the strength of the association between brain weight and the dependent variable is low (e.g. $R <$ ~ 0.1) there is high agreement between the brain weight and body weight models, both of which typically fail to detect an association. As the correlation between brain weight and the dependent measure begins to increase, there is an uptick in discrepant results between the brain weight and body weight models, reflecting cases in which the true signal is detected in the brain weight model, but not the body weight model. Lastly, as the correlation between brain weight and the dependent measure approaches $R = 1$, there is once again high agreement between the brain weight and body weight models, both of which detect the underlying association. Despite these commonalities, there are also clear differences between the datasets characterized by low, moderate, and high variance. Specifically, when there is high variance in the brain weight and body weight data, only a moderate association (R = ~ 0.25) between brain weight and the dependent measure is required for this effect to be reliably detected using either brain weight or body weight as the predictor variable.

Given these results, what inferences might we make about the consequences of using of body weight to estimate brain weight in Horschler et al. (2019)? In Horschler et al.'s paper, the range of variation in body weight was relatively high, with 3.37-fold variation within the interquartile range. In the simulation presented above, the average fold-variation across the interquartile range for the low-, moderate-, and high-variance distributions was 1.22, 1.69, and 3.33, respectively, aligning the data from Horschler et al. closely with the high-variance distribution in our simulation. Second, for the major findings in the Horschler et al. paper, the correlation coefficient between estimated brain weight (derived as a linear transformation of body weight) and the outcome was $R =$ ~ 0.54. Interpolating this value against the functions in Fig. 2, we see that under conditions closely matching the Horschler et al. study, p values for a brain weight versus a body weight model typically differ by less than 0.0001, beta coefficients for the two models are within 10% of one another, and the decision to retain or reject the null hypothesis is the same in all cases across 500 simulated scenarios with these parameters. Therefore, while there are certainly numerous advantages of obtaining direct measures of brain weight, at a practical level, hypothesis tests using either brain weight or body weight as the predictor variable are expected to yield highly similar results. Rather than treating this as an inescapable problem, we view it as an opportunity to probe possible links between absolute brain weight and cognition in dogs, despite the current lack of high-quality data on neuroanatomical variation across breeds (but see Hecht et al. 2019 for new work in this direction).

Of course, no simulation will ever resolve empirical questions about the relationships between brain weight and cognitive variation among dog breeds. However, these simulations do help to determine when and to what extent certain assumptions may be justified, given knowledge about the strength of the association between brain and body weight and an observed effect size. On these grounds, the simulations above suggest that the assumptions in Horschler et al. (2019) were both reasonable and modest. Therefore, while our findings should still be interpreted with caution, we argue that they hold value in guiding new testable hypotheses about how variation in diverse cognitive processes may be linked to aspects of neuroanatomy, both in dogs and across other species. Nonetheless, we wholeheartedly agree with Montgomery that obtaining empirical data on brain weight, structure, and cellular composition across dog breeds will be critical for future work on these questions, and we look forward to progress in this area.

Acknowledgements We are grateful to Jeffrey Katz and Debbie Kelly for the opportunity to write this response, as well as to Brian Hare, Josep Call, Juliane Kaminski, Ádám Miklósi, Laurie Santos, Richard Wrangham, David Ivy, Elliot Cohen, Kip Frey, and all other members of the *Dognition* team. We thank Kevin Doubleday for helpful comments on an earlier draft of the paper, and Stephen Montgomery for valuable discussion about the original paper as well as this commentary.

References

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 113(9):2532–2537. https://doi.org/10.1073/pnas.1505913113

Bronson RT (1979) Brain weight-body weight scaling in breeds of dogs and cats. Brain Behav Evol 16:227–236

Buechel SD, Boussard A, Kotrschal A, van der Bijl W, Kolm N (2017) Brain size affects performance in a reversal-learning test. Proc R Soc Lond Ser B Biol Sci. https://doi.org/10.1098/rspb.2017.2031

Burger JR, George MA, Leadbetter C, Shaikh F (2019) The allometry of brain size in mammals. J Mammol 100(April):276–283. https://doi.org/10.1093/jmammal/gyz043

Carreira L (2016) Using Bronson equation to accurately predict the dog brain weight based on body weight parameter. Vet Sci 3(4):36. https://doi.org/10.3390/vetsci3040036

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(2):115–124. https://doi.org/10.1159/000102973

Healy SD, Krebs JR (1992) Foord storing and the hippocampus in corvids: amount and volume are correlated. Proc R Soc Lond 248:241–245. https://doi.org/10.1098/rspb.1992.0068

Hecht E, Smaers J, Dunn W, Kent M, Preuss T, Gutman D (2019) Significant neuroanatomical variation among domestic dog breeds. J Neurosci 39:303–319

Herculano-Houzel S (2012) Neuronal scaling rules for primate brains: the primate advantage. Progress in brain research. Elsevier, Amsterdam, pp 325–340. https://doi.org/10.1016/B978-0-444-53860-4.00015-5

Horschler DJ, Hare B, Call J, Kaminski J, Miklósi Á, MacLean EL (2019) Absolute brain size predicts dog breed differences in executive function. Anim Cogn 22(2):187–198. https://doi.org/10.1007/s10071-018-01234-1

Hsu Y, Serpell JA (2003) Development and validation of a questionnaire for measuring behavior and temperament traits in pet dogs. J Am Vet Med Assoc 223(9):1293–1300

Kotrschal A, Rogell B, Bundsen A, Svensson B, Zajitschek S, Brannstrom I, Kolm N (2013) Artificial selection on relative brain size in the guppy reveals costs and benefits of evolving a larger brain. Curr Biol 23:168–171

Kotrschal A, Corral-Lopez A, Amcoff M, Kolm N (2015) A larger brain confers a benefit in a spatial mate search learning task in male guppies. Behav Ecol 26(2):527–532. https://doi.org/10.1093/beheco/aru227

Krebs JR, Sherry DF, Healy SD, Perry VH, Vaccarino AL (1989) Hippocampal specialization of food-storing birds. Proceedings of the National Academy of Sciences of the United States of America 86(4):1388–1392. http://www.ncbi.nlm.nih.gov/pubmed/2919184%0Ahttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC286696

MacLean EL, Hare B, Nunn CL, Addessi E, Amici F, Anderson RC, Zhao Y (2014) The evolution of self-control. Proc Natl Acad Sci 111(20):E2140–E2148. https://doi.org/10.1073/pnas.1323533111

MacLean EL, Snyder-Mackler NS, vonHoldt BM, Serpell JA (2019) Highly heritable and functionally relevant breed differences in dog behavior. Proc R Soc B Biol Sci. https://doi.org/10.1098/rspb.2019.0716

Montgomery SH (2019) Are brain weights estimated from scaling relationships suitable for comparative studies of animal cognition? Animal Cognit. https://doi.org/10.1007/s10071-019-01300-2

Overington SE, Morand-Ferron J, Boogert NJ, Lefebvre L (2009) Technical innovations drive the relationship between innovativeness and residual brain size in birds. Anim Behav 78(4):1001–1010. https://doi.org/10.1016/j.anbehav.2009.06.033

Reader SM, Laland KN (2002) Social intelligence, innovation, and enhanced brain size in primates. Proc Natl Acad Sci 99(7):4436–4441. https://doi.org/10.1073/pnas.062041299

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(15):5460–5465. https://doi.org/10.1073/pnas.0408145102

Sol D, Bacher S, Reader SM, Lefebvre L (2008) Brain size predicts the success of mammal species introduced into novel environments. Am Natural 172:S63–S71. https://doi.org/10.1086/588304

Stewart L, MacLean EL, Ivy D, Woods V, Cohen E, Rodriguez K, Hare B (2015) Citizen science as a new tool in dog cognition research. PLoS One 10(9):1–16. https://doi.org/10.1371/journal.pone.0135176