

# Using macroevolutionary patterns to distinguish primary from secondary cognitive modules in primate cross-species performance data on five cognitive ability measures

Michael A. Woodley of Menie<sup>a,\*</sup>, Mateo Peñaherrera-Aguirre<sup>b</sup>, JohnMichael Jurgensen<sup>c</sup>

<sup>a</sup> Vrije Universiteit Brussel, Center Leo Apostel for Transdisciplinary Studies, Brussels, Belgium

<sup>b</sup> University of Arizona, Department of Psychology, Tucson, USA

<sup>c</sup> Boston University, Department of Philosophy, Boston, USA

## ARTICLE INFO

### Keywords:

Big-G  
Macroevolution  
Phylogenetic methods  
Primary modules  
Selection regimes  
Secondary modules

## ABSTRACT

Species-level data on five cognitive ability measures from 69 primate species are used in conjunction with comparative phylogenetic methods to test for the existence of primary and secondary modules. The former are 'hard wired', and solve phylogenetically recurrent problems, whereas the latter are a function of domain general problem-solving mechanisms being applied to solving narrower problems, which yields the ability to spontaneously solve those problems once the solutions are learned. It is found that these abilities exhibit affinities for different macroevolutionary patterns relative to 'Big G', and positive associations with dietary breadth and brain size. The analyses were also conducted using each ability residualised for *G*. It was found that the Ornstein-Uhlenbeck (OU) model best captured the macroevolution of residual tactical deception, and White Noise (WN) best fit the remainder. Residual tactical deception positively associates with brain volume, whereas the extractive foraging and innovation residuals negatively associate with this and the innovation residual negatively associates with social group size. The affinity of residual tactical deception for the OU model indicates that it may be a primary module under adaptive optimization selection. The predominance of WN in characterizing the macroevolution of the remaining residuals indicates that they may be secondary modules, under the influence of developmental and ecological (rather than phylogenetic) factors. Negative associations involving brain size (in two cases) and social group size (in one) suggest that the optimal conditions for cultivating these modules exist when these parameters are low.

## 1. Introduction

Species differences in the means of observational counts of problem-solving behaviors sampled with respect to a wide array of domains have been found to give rise to a 'general intelligence' or *g*-like general factor. As this pattern emerges from comparisons involving different species (as opposed to individuals) the term Big-*G* (or just *G*) is used to differentiate it from the 'little-*g*' that exists at the level of individual differences (Burkart, Schubiger, & van Schaik, 2017a; Burkart, Schubiger, & van Schaik, 2017b; Fernandes, Woodley, & te Nijenhuis, 2014). The phenomenon of *G* was first described by Deaner, van Schaik, and Johnson (2006). These researchers performed a meta-analysis of studies examining cognition in non-human primates, grouping these based on subject attributes and experimental design. A Bayesian latent variable model

found that primate genera differed in their overall performance across tasks. The presence of *G* was strongly corroborated by the finding that across 229 genus-by-genus comparisons, genera could be consistently ranked with respect to task performance in 85% of cases. The researchers were able to rule out the influence of species-by-domain effects (or modules). These researchers also noted that Great Apes consistently outperformed other primates across tasks.

Reader, Hager, and Laland (2011) were able to identify the existence of *G* among primate species using a more straightforward factor analytic approach. They synthesized data from over 4000 articles, spanning 75 years, and were able to obtain observational frequency counts of performance with respect to five cognitive domains (*innovation, social learning, tool use, extractive foraging, and tactical deception*) for a total of 62 primate species. The use of Principal Components Analysis yielded a

\* Corresponding author.

E-mail address: [michael.woodley@vub.be](mailto:michael.woodley@vub.be) (M.A. Woodley of Menie).

$G$  factor accounting for 65% of the variance in performance across tasks, with loadings ranging from 0.88 (in the case of *tool use*) to 0.74 (in the case of *tactical deception*). Significant positive associations were found between  $G$  and log-transformed body and brain size, progression index, log-transformed neocortex volume, the ratio of the neocortex to the rest of the brain, and the residual of the log-transformed brain size on log-transformed body mass. As with Deaner et al. (2006), Reader et al. (2011) also noted that the Great Apes outscore all other primates in terms of level of  $G$ .

Fernandes et al. (2014) expanded Reader et al.'s (2011) dataset to include an additional seven species, bringing the total up to 69. The use of Principal Axis and Unit-Weighted factor analytic approaches confirmed the presence of a robust  $G$  factor in these data, even after controls for research effort. It was found that the  $G$  loadings onto each of the five abilities also strongly, positively moderated the magnitude of the average pair-wise differences between species and between-species variance associated with each ability, meaning that species differences were most pronounced on the more  $G$  loaded measures. It was also found that the evolutionary lability (specifically the rate of evolutionary change) increased with the ability's  $G$  loading, indicating that  $G$  is the primary variance component on which macroevolutionary pressures have operated in shaping primate cognition.

A more recent analysis involving the same dataset (Fernandes, Peñaherrera-Aguirre, Woodley of Menie, & Figueredo, 2020) found that  $G$  exhibited greater evolutionary lability relative to a variety of neuro-anatomical volume measures commonly used as proxies for intelligence in comparative studies. These included both absolute and (body size) residualised brain size, neocortex size, neocortex ratio, cerebellum size, and absolute body mass. Among these measures,  $G$  exhibited the greatest evolutionary rate (0.15 standard deviations [SDs] per million years). Brain size (one of the most commonly used proxies for intelligence) exhibited the lowest lability (close to 0.00 SDs). The only neuroanatomical measures that came close to matching  $G$  in terms of lability were residualised and unresidualised cerebellum (both in the region of 0.10 SDs).

### 1.1. The frequency count method

The collection of behavioral data from the literature (including information based on naturalistic observations in captive and field settings) for comparative phylogenetic examinations is a fairly common practice across an array of scientific disciplines including evolutionary behavioral ecology, evolutionary anthropology, evolutionary psychology, and comparative psychology (e.g., Allen, Street, & Capellini, 2017; Capellini, Baker, Allen, Street, & Venditti, 2015; Creighton, Greenberg, Reader, & Mooers, 2021; Ducatez, Sol, Sayol, & Lefebvre, 2020; Heldstab, Isler, Burkart, & van Schaik, 2019; Kamilar & Cooper, 2013; Nunn & Samson, 2018; Opie, Atkinson, Dunbar, & Shultz, 2013; Plavcan & van Schaik, 1997; Shultz, Opie, & Atkinson, 2011; Sol, Duncan, Blackburn, Cassey, & Lefebvre, 2005; Varricchio et al., 2008; Wich & Nunn, 2002). It is appreciated that this method might be somewhat unfamiliar to (human) intelligence researchers and differential psychologists, who are the primary audience for this paper. This method will therefore be explained in some detail here, in order to avoid confusions concerning the nature of the analyses that are to be conducted subsequently.

It should first be noted that data gathering for the purposes of comparative phylogenetic examinations is not strictly analogous to a literature review leading to traditional meta-analysis wherein the main goal is to determine the magnitude and consistency of a reported parameter estimate across publications. Instead, researchers using this sampling method collect information on the presence of a particular behavioral phenotype as described by the literature and subsequently incorporate this information into phylogenetic models that allow for various hypotheses to be tested.

The present study is based on a sampling method developed by Lefebvre, Whittle, Lascaris, and Finkelstein (1997) and Reader & Laland

(2002), whereby researchers gather data on the frequency of a particular behavior for different species (Fernandes and colleagues coined the term Frequency Count Method to refer to this sampling procedure; Fernandes et al., 2014). In particular, the database analyzed in this study comes from an online data repository uploaded by Reader et al. (2011) and subsequently expanded by Fernandes et al. (2014). An advantage of the Frequency Count Method (when conducting comparative phylogenetic examinations) is that since this procedure collects data on the prevalence of behavioral phenotypes, rather than on the performance of individuals on a particular cognitive test, it remains ecologically valid at the species-level. In contrast, cross-species comparisons based on experimental conditions at the individual-level may provide inaccurate results if the various cognitive tests are not specifically adapted to be used across species (e.g., floor and ceiling effects may provide some indication of a test's limited ecological validity; see discussion in Woodley of Menie, Fernandes, te Nijenhuis, Peñaherrera Aguirre, & Figueredo, 2017; Woodley of Menie & Peñaherrera-Aguirre, 2022). This procedure does not consider the number of individuals observed per study either, treating the species or higher taxonomic rank (e.g., genus) instead as the unit of analysis, however previous publications have found that demographic information (e.g., group size or population size) does not bias the final frequency count of the various behavioral phenotypes (Lefebvre, 2011).

In constructing their 2011 database, Reader and colleagues surveyed articles gathered primarily from four primatological journals: *Primates*, the *American Journal of Primatology*, *Folia Primatologica*, and the *International Journal of Primatology*, in addition to other pertinent publications. According to the authors, these outlets were selected because they publish the largest volume of primate behavioral research. As part of their search strategy, the authors used several keywords, including "extract," "traditional," "novel," used to classify behavioral patterns such as "innovation," "social learning," or "extractive foraging." This strategy limited the classification of a behavioral pattern to the descriptions provided by the authors of the articles. Hence, this procedure was designed to circumvent potential subjective bias on the part of Reader and colleagues when collecting the data.

Reader et al. (2011) computed the observation frequencies as the overall number of reported instances per behavioral category. Behavioral patterns such as tool use, social learning, and innovation occurred across environmental and behavioral situations (e.g., in response to predators, during locomotion, as part of social displays, or in association with foraging behavior). Although Reader et al. (2011) initially collected information on all instances in which the behavior occurred (i.e., either reported in a captive or wild setting, as part of an experimental intervention or in response to human activity including habitat degradation or food provisioning), their analyses were restricted to a limited dataset comprised of instances of naturalistic observations conducted in a field setting without human intervention (Reader & Laland, 2002). The authors also conducted inter-observer reliability analyses on the reports of the aforementioned behavioral patterns to determine whether any of the authors introduced a subjective bias in the course of coding the behaviors. The analyses revealed high concordance values for tool use (internal consistency [IC] = 0.94), social learning (IC = 0.95), and innovation (IC = 0.83) across all raters. The authors also collected data on tactical deception from Byrne and Whiten (1990), who in turn surveyed primatological experts to determine the prevalence of tactical deception across species of nonhuman primates. These reports were also based on naturalistic observations rather than the results of experimental procedures conducted either in captivity or in the wild.

Concerning extractive foraging, the authors operationalized this variable as follows:

"Feeding on foods that must first be extracted from matrices in which they are embedded or encased, including nutmeat, shellfish, snails,

eggs, brains, bone marrow, roots, tubers, and ant and termite mounds” (Reader et al., 2011; supplementary material, p. 5).

After reviewing the relevant articles, the authors validated their extractive foraging data by comparing their results with Gibson’s (1986) description of this behavioral pattern across species of nonhuman primates. Their examination provided additional evidence concerning their measures’ consistency with previously assembled behavioral catalogs. Reader and colleagues also mentioned that reports of extractive foraging in the literature (as naturalistic observations) predict primate performance in cognitive tasks requiring extractive foraging in experimental settings. After extracting *G* from the aforementioned behavioral patterns, Reader et al. (2011) correlated the factor scores with a global performance index calculated based on an array of cognitive subtests collected from several primate species in a captive setting (Deaner et al., 2006). Reader and collaborators found a sizeable and significant association between *G* and the laboratory global performance scores.

Similar data collection procedures have also been used with other clades. Over the past two decades, (Lefebvre et al., 1997; Lefebvre, Juretic, Nicolakakis and Timmermans, 2001; Lefebvre, Reader and Sol, 2004) have gathered innovation data from the ornithological literature. The authors demonstrated that this type of procedure is not affected by: a) type of journal; b) the journal’s editorial policy; c) the associated research effort; d) observation bias; e) population size; f) developmental processes; g) or common ancestry. Moreover, Lefebvre and collaborators found that their innovation metric correlated well with various measures of learning performance in a laboratory setting. Inter-observer agreement estimates have been found to be consistently high, ranging from 0.82 to 0.95 (Lefebvre, 2011; Nicolakakis & Lefebvre, 2000). Consequently, the aforementioned publications provide strong evidence concerning the validity of this method, whereby the results remained relatively unaltered even after accounting for potential sampling biases.

## 1.2. Primary and secondary modules

Moving now to the main research question, while much research has been conducted on the question of *G* and its phylogenetic, ecological, and neuroanatomical correlates in comparative psychological research, little research by contrast has been conducted into the no less important question of the nature of modularity at this level of analysis. The existence of two basic kinds of modules have been proposed in comparative psychology, primary and secondary. Primary modules are a function of selection favoring the emergence of specialized and dedicated problem-solving systems with respect to which members of a species are optimized (i.e., there exists very little variance among individuals). Such modules are experience expectant, meaning that their functioning is elicited by exposure to very specific phylogenetically recurrent signals. They are also inflexible, meaning that they can only be used to solve narrowly compatible problems. Burkart et al. (2017a) give as an example of such a module the tendency for small moving objects to elicit hunting-repertoire derived behaviors from young felids (the example of a house cat pawing at a piece of string dangling in front of it is likely familiar to the reader). Secondary modules are a function of learned skills, acquired in ontogenetic time, via recurrent exposure to a stimulus. These modules once acquired allow for automatic solutions to be derived for a given problem, but only after a period of learning has taken place. These modules reflect underlying behavioral flexibility and likely stem from the action of executive functioning (i.e., attention and control) applied to the process of learning. One way to think of these is to imagine domain general processes (such as those related to *g*) building, via the allocation of cortical real estate, secondary dedicated cognitive structures capable of managing domain specific problems once the problem has become sufficiently recurrent, so that the solution no longer needs to be relearned with each iterative exposure. Burkart et al. (2017a) give as an example of this the ability for humans to learn algorithms with which they can then solve algebraic equations efficiently

with enough practice, and without the need to relearn the algorithm once it has been ‘captured’ in the form of a secondary module.

In the present work, a comparative phylogenetic analysis will be conducted in order to identify primary and secondary modules via reanalysis of the extended primate *G* dataset utilized in (Fernandes et al., 2014; Fernandes et al., 2020). Each of the five abilities will be considered independently in relation to their affinities for different macroevolutionary modes, and their association with a small set of neuroanatomical and ecological correlates. These analyses will be conducted on the abilities in both unresidualised and residualised (for *G*) form. The latter are likely to be especially informative as to the evolutionary dynamics influencing the emergence of different forms of modularity, as these residuals will be *wholly independent* of phylogenetic processes acting at the level of *G*, and their associated macroevolutionary patterns will necessarily result from *domain specific* macroevolutionary pressures acting on these. Two macroevolutionary regimes in particular are likely to allow for primary and secondary models to be clearly distinguished; Ornstein-Uhlenbeck and White Noise. The former relates to forms of selection that optimize populations for particular adaptations (i.e., via centripetal, or variance-reducing stabilizing selection) (Butler & King, 2004), and is expected to be the preferred macroevolutionary mode for primary modules. The latter relates to a role for ecological and ontogenetic effects operating outside of phylogenetic processes in conditioning trait levels (Pennell et al., 2014). This macroevolutionary mode is expected to predominate in the case of secondary modules, as these arise in ontogenetic (as opposed to phylogenetic) time, and are conditional upon the presence of specific ecological problems that may vary in terms of their presence/absence and significance substantially between species. In the present analysis a variety of alternative phylogenetic models will be fitted to the data in order to thoroughly exhaust the space of likely macroevolutionary modes. Both the residualised and unresidualised forms of the abilities will also be examined in relation to between-species variation with respect to three variables (diet breadth, brain mass, and social group size) in order to determine whether there are any notable associations.

## 2. Methods

### 2.1. Measures

The current analyses used data collected and examined by Fernandes et al. (2020) and Reader et al. (2011) on the following list of cognitive abilities in a sample of 69 nonhuman primate species. The number of studies per behavioral pattern can be found in Fernandes et al. (2020):

#### 2.1.1. Extractive foraging

The cognitive and behavioral ability to extract food resources that have been concealed, cached, or covered by an external barrier (Fernandes et al., 2014; Reader et al., 2011).

#### 2.1.2. Tool use

The creation and use of artifacts for addressing physical or social challenges. This cognitive indicator is often viewed as a proxy for an organism’s ability to control and alter nearby surroundings (Gibson, Gibson, & Ingold, 1994; Wynn, 1988). Several studies have found that tool use is strongly correlated with other cognitive indicators in non-human species (Fernandes et al., 2014; Lefebvre et al., 2004; Reader & Laland, 2002).

#### 2.1.3. Innovation

An organism’s capacity to develop novel solutions to new, complex, social, and physical challenges. More specifically, this cognitive indicator entails the developmental capacity to adjust to novel ecological conditions. Consequently, innovation is regarded as a strong indicator of general cognitive ability (Fernandes et al., 2014; Lefebvre et al., 2004; Reader & Laland, 2002; Sol et al., 2005).

#### 2.1.4. Tactical deception

An organism's ability to redirect the attention of others in order to mislead, damage, or benefit from others or via their distraction (Byrne & Whiten, 1990). Tactical deception does not require the individual to accurately interpret or "read" the mental states of the "dupe," but can be achieved by predicting the behavior of the dupe in a particular social context (Byrne and Whiten, 1990).

#### 2.1.5. Social learning

The ability to gain information, skills, and knowledge from kin, peers, and other conspecific adults. Although previous publications considered social learning an essential part of Machiavellian intelligence (Byrne & Whiten, 1990; Whiten & Byrne, 1988), more recent examinations also view this ability as a central capacity for the evolution of traditions and culture (Reader et al., 2011).

### 2.2. Unit-weighted factor estimations and phylogenetic generalized least squares models

Traditional statistical analyses, such as regression-based models, assume that the data's residuals are independent. As closely related species have a higher probability of resembling each other than more distantly related lineages, the data's residuals collected across several species are not statistically independent. Ignoring the underlying pseudoreplication issues associated with this type of data increased the probability of Type I errors (Nunn, 2011; Peñaherrera & Fernandes, 2021). Over the past three decades, various comparative phylogenetic methods have been developed to account for these statistical conditions (Nunn, 2011). For example, diagnostic statistics such as Pagel's  $\lambda$  provides a quantitative estimate of the extent to which the macroevolutionary pattern observed in the phylogenetic tree matches the phenotypic similarity across species in the dataset (Pagel, 1999). Values of Pagel's  $\lambda$  closer to 1.0 indicate that the rate of macroevolutionary change is attributable to the passage of time. In contrast, when Pagel's  $\lambda$  is close to zero, it indicates that the phenotype across species is less similar to what would be expected from the underlying phylogeny. In addition to reducing the probability of Type I errors, comparative phylogenetic methods have also been found to increase the model's statistical power and reduce the probability of Type II errors (Maclean & Nunn, 2017; Peñaherrera & Fernandes, 2021). Although some statistical procedures such as Linear Mixed Models may account for pseudoreplication issues by including cladistic-related factors as random effects (e.g., genus, tribe, family, subfamily, infraorder, suborder, order, or class), comparative phylogenetic methods, such as phylogenetic generalized least square regression, incorporate the trait's phylogenetic signal into the model providing a more accurate computation of the various macroevolutionary parameter estimates.

It is also worth noting that the probability of reporting a particular behavioral phenotype is expected to vary depending on how extensively a nonhuman species has been described in the literature. Consequently, Reader et al. (2011) operationalized research effort (i.e., the degree to which some species have been studied more than others) as the number of publications per taxon. The authors collected this information from the *Zoological Record*, an online data repository featuring scientometric data on a large number of nonhuman taxa, and proceeded to use the number of publications as a predictor of the various behavioral patterns (social learning, innovation, tactical deception, tool use, and extractive foraging) in several regression analyses. The authors then extracted the models' residuals, yielding a set of values that provided a better estimate of the prevalence of these behavioral phenotypes across species of nonhuman primates net of effort. Subsequent publications, including Fernandes et al. (2014, 2020), have also used this analytical procedure to conduct macroevolutionary examinations with the aforementioned behavioral patterns.

A series of Linear Models were conducted using research effort (number of publications per species in the *Zoological Record*) as a

predictor of the various cognitive abilities. The pertinent residuals were then extracted and used to compute a unit-weighted (UW) *G* factor. This process involves the standardization of all relevant indicators and the calculation of an average across the pertinent *z*-indicators. In contrast to other factor scoring procedures, UW does not suffer from sample-specific analytic limitations (error estimations associated with small sample sizes or case numbers; Gorsuch, 1983). Consequently, factor loadings under UW scoring are equivalent to the part-whole correlations between the standardized indicators and the standardized UW factor. Information was also gathered on the species' diet breadth, adult brain mass (in grams), and social group size from the online data repository *Pantheria* (Jones et al., 2009). The *mice* (Zhang, 2016) and *sjmisc* (Lüdecke, 2018) packages were used to compute MCMC imputation across 100 samples and 50 iterations for ecological indicators featuring missing data. A consolidated ecological dataset was later integrated into several Phylogenetic Generalized Least Squares Models (PGLS), predicting the various unresidualized and *G* residualized cognitive abilities.

These analyses were conducted with the phylogenetic comparative package *Phytools* (Revell, 2012) in R v. 4.0.1.

### 2.3. Macroevolutionary selection regimes

The present study compared each cognitive indicator across nine different macroevolutionary models: Brownian Motion, Pagel's  $\lambda$ , Ornstein-Uhlenbeck, Early Burst, Acceleration-Deceleration, Kappa, Mean Trend, Rate Trend, and White Noise. This approach allowed for comparative examination of the macroevolutionary patterns characteristic of each trait's level of phylogenetic signal and its degree of preservation, as well as corresponding evolutionary rates across the primate phylogeny. The various macroevolutionary models were estimated with the *Geiger* package (Pennell et al., 2014) in R v. 4.0.1.

**Brownian Motion (BM)** acts as a null macroevolutionary model depending on the phylogenetic trees' branch length. Thus, correlation matrices are assumed to be proportional to the shared ancestry among species (Felsenstein, 1973). The dimension of the branches represents how much time has passed since the species emerged. As a result, under BM, a trait is more likely to change in older lineages compared to those in more recent ones. This model does not account for the direction of the trait's change, i.e., whether it increases or decreases over time, or the speed at which the feature evolved (deceleration or acceleration). The model assumes the phenotypic attribute evolved at a stable rate following a random walk (Nunn, 2011).

**Pagel's  $\lambda$**  alters the length of the branches in the phylogenetic tree based on the trait's phylogenetic signal (Pagel, 1999). This measure ranges from 0 to 1, where values closer to 1 indicate the trait evolved under BM, whereas values closer to 0 suggest high evolutionary lability. This macroevolutionary model does not give any indication as to the reasons for a trait's preservation level. The model operates as an all-encompassing alternative to all other models in the analysis. Pagel's  $\lambda$  operates under the logic of rejecting a BM model but acknowledging that the trait did not evolve under any of the competing macroevolutionary models under consideration, as reflected by the model's statistical fit.

**Ornstein-Uhlenbeck (OU)** builds upon the BM model by including the parameter  $\alpha$  that represents the magnitude of a constraint force, considering whether the selective pressure orients all lineages in the phylogeny toward a specific value. Even though the Ornstein-Uhlenbeck model is generally viewed as a measure of stabilizing selection (Butler & King, 2004), this macroevolutionary model considers the alternative hypothesis that the trait evolved toward an optimum, regardless of whether it occurred due to stabilizing or directional selection. In contrast to other models, such as Pagel's  $\lambda$ , the values computed by an Ornstein-Uhlenbeck model do not have an upper range (i.e., they span from zero to infinity).

**Early Burst** evaluates any exponential increase or decrease of the trait's evolutionary rate across time (Harmon et al., 2010). This estimation derives from the addition of a statistical parameter. Whereas a

**Table 1**  
Phylogenetic signal of non-residualized and residualized cognitive abilities.

Non-residualized Cognitive Abilities /GCA			Residualized Cognitive Abilities		
Cognitive ability	Pagel's $\lambda$	p-value LRT (0)	Cognitive ability	Pagel's $\lambda$	p-value LRT (0)
Z- T. Deception	0.69	0.0000	Z-res. T. Deception	0.46	0.0086
Z- Ex. Foraging	0.44	0.6426	Z-res. Ex. Foraging	0.21	0.3191
Z- Innovation	0.24	1.0000	Z-res. Innovation	0.07	0.3623
Z- Social Learning	0.00	1.0000	Z-res. Social Learning	0.00	1.0000
Z- Tool Use	0.69	0.0020	Z-res. Tool Use	0.08	0.5685
Z-GCA	0.54	0.0179			

value equal to 0 is analogous to a BM evolutionary model, an estimate below 0 reflects a fast niche-filling process accompanied by an exponential decline in evolutionary rates. Such a dynamic is dependent on novel ecological niches becoming available and eventually reaching a point of saturation.

**Delta** considers the speed at which the trait evolved over time (Pagel, 1999). Similar to previous models, this model incorporates an additional statistical parameter, whereby values above 1 indicate an accelerated macroevolutionary process acting particularly at the tip of the phylogenetic tree. Consequently, larger values reflect a linear increase in evolutionary rates over time. As a result, this model evaluates whether adaptations are lineage-specific.

**Kappa** assumes that the probability of trait change increases with the speciation events occurring across the phylogenetic tree (Pagel, 1999). Thus, this analysis evaluates whether the trait evolved following a punctuational dynamic. It is worth noting that it is difficult to interpret the results of this model if multiple speciation events are missing in the phylogenetic tree, either due to inadequate sampling or extinction events.

**Mean Trend** examines whether the evolution of a trait occurred following a directional drift or toward a general trend, as reflected by a tendency to display either smaller or larger values over time (Pennell et al., 2014). Although this model is generally recommended for non-

ultrametric trees (in ultrametric trees, the likelihood surface is usually flat relative to the trend's slope), it was considered reasonable here to include this as part of the overall macroevolutionary model comparisons.

**Rate Trend** acts as a diffusion model examining whether the rate of trait evolution followed a linear trend over time, either increasing or decreasing (Pennell et al., 2014).

**White-Noise (WN)** is a non-phylogenetic mode that assumes the information was collected from a normal distribution lacking a covariance structure across the lineages in the dataset (Pennell et al., 2014). This model allows for the implementation of the variance parameter  $s^2$  values, and employs similar bounds to those associated with a BM model.

### 3. Results

#### 3.1. Phylogenetic signal of non-residualized and residualized cognitive abilities

Table 1 shows the values of Pagel's  $\lambda$ , a measure of phylogenetic preservation, for the various cognitive abilities as well as the results of a likelihood ratio test, assuming the phylogenetic signal to be identical to zero. Tactical deception and tool use exhibited strong phylogenetic signals. A likelihood ratio test revealed that these values were significantly different from zero, thus suggesting that the traits are phylogenetically conserved. Alternatively, extractive foraging, innovation, and social learning featured a phylogenetic signal that was not significantly different from zero. The G factor was also evolutionarily conserved, with its phylogenetic signal significantly differing from zero. Residualizing the five cognitive abilities for G reduced the degree to which the traits appeared to be evolutionarily conserved. Whereas the residualized values for tactical deception remained significantly different from zero, extractive foraging, innovation, social learning, and tool use had relatively small values of Pagel's  $\lambda$  that were not statistically different from zero.

**Table 2**

Macroevolutionary model comparison evaluating the effect of nine selection regimes upon five cognitive abilities and G. Bolded values indicate those associated with the best fitting model(s).

Akaike Information Criteria (AIC weight)						
	Z-T. Deception	Z- Ex. Foraging	Z- Innovation	Z-Tool Use	Z- Social Learning	Z-G
Delta AIC	188.51 (0.01)	195.21 (0.33)	197.81 (0.18)	194.49 (0.04)	198.28 (0.15)	195.86 (0.12)
BM AIC	219.18 (0.00)	240.87 (0.00)	254.81 (0.00)	225.64 (0.00)	261.99 (0.00)	239.87 (0.00)
OU AIC	188.16 (0.01)	<b>195.18 (0.34)</b>	197.81 (0.18)	194.36 (0.04)	197.97 (0.18)	195.83 (0.15)
Pagel's $\lambda$ AIC	<b>179.56 (0.83)</b>	197.75 (0.09)	197.97 (0.17)	<b>188.45 (0.86)</b>	197.97 (0.18)	<b>192.37 (0.66)</b>
Early Burst AIC	221.18 (0.00)	242.87 (0.00)	256.81 (0.00)	227.64 (0.00)	263.99 (0.00)	241.87 (0.00)
Kappa AIC	183.02 (0.15)	203.59 (0.01)	214.39 (0.00)	195.15 (0.03)	228.26 (0.00)	202.37 (0.00)
Mean Trend AIC	221.18 (0.00)	242.87 (0.00)	256.81 (0.00)	227.64 (0.00)	263.99 (0.00)	241.87 (0.00)
Rate Trend AIC	214.24 (0.00)	235.76 (0.00)	249.43 (0.00)	220.90 (0.00)	256.62 (0.00)	234.68 (0.00)
WN AIC	195.97 (0.00)	195.97 (0.23)	<b>195.97 (0.46)</b>	195.97 (0.02)	<b>195.97 (0.49)</b>	195.97 (0.11)

Corrected Akaike Information Criteria (AICc)						
	Z-T. Deception	Z- Ex. Foraging	Z- Innovation	Z-Tool Use	Z- Social Learning	Z-G
Delta AICc	188.88 (0.01)	<b>195.58 (0.33)</b>	198.18 (0.18)	194.87 (0.04)	198.65 (0.17)	196.23 (0.11)
BM AICc	219.36 (0.00)	241.05 (0.00)	254.99 (0.00)	225.83 (0.00)	262.18 (0.00)	240.05 (0.00)
OU AICc	188.53 (0.01)	<b>195.56 (0.33)</b>	198.19 (0.18)	194.74 (0.04)	198.34 (0.17)	196.21 (0.12)
Pagel's $\lambda$ AICc	<b>179.94 (0.83)</b>	198.13 (0.09)	198.34 (0.16)	<b>188.83 (0.86)</b>	198.34 (0.17)	<b>192.72 (0.65)</b>
Early Burst AICc	221.55 (0.00)	243.24 (0.00)	257.18 (0.00)	228.02 (0.00)	264.37 (0.00)	242.25 (0.00)
Kappa AICc	183.39 (0.15)	203.97 (0.00)	214.77 (0.00)	195.52 (0.03)	228.64 (0.00)	202.74 (0.00)
Mean Trend AICc	221.55 (0.00)	243.24 (0.00)	257.18 (0.00)	228.02 (0.00)	264.37 (0.00)	242.25 (0.00)
Rate Trend AICc	214.61 (0.00)	236.13 (0.00)	249.81 (0.00)	221.28 (0.00)	257.00 (0.00)	235.06 (0.00)
WN AICc	196.15 (0.00)	196.15 (0.25)	<b>196.15 (0.49)</b>	196.15 (0.02)	<b>196.15 (0.50)</b>	196.15 (0.12)

**Table 3**

Macroevolutionary model comparison evaluating the effect of nine selection regimes upon five *G*-residualized cognitive abilities. Bolded values indicate those associated with the best fitting model(s).

Akaike Information Criteria (AIC weight)					
	Z-res. T. Deception	Z-res. Ex. Foraging	Z-res. Innovation	Z-res. Tool Use	Z-res. Social Learning
Delta AIC	189.93 (0.37)	197.97 (0.16)	199.67 (0.08)	197.70 (0.18)	197.96 (0.18)
BM AIC	217.13 (0.00)	246.08 (0.00)	281.14 (0.00)	270.02 (0.00)	253.35 (0.00)
OU AIC	<b>189.74</b> ( <b>0.40</b> )	197.90 (0.16)	197.97 (0.18)	197.63 (0.19)	197.96 (0.18)
Pagel's λ AIC	191.06 (0.21)	196.98 (0.26)	197.14 (0.27)	197.64 (0.19)	197.97 (0.18)
Early Burst AIC	219.13 (0.00)	248.08 (0.00)	283.14 (0.00)	269.97 (0.00)	255.35 (0.00)
Kappa AIC	200.50 (0.00)	212.94 (0.00)	229.09 (0.00)	220.44 (0.00)	227.86 (0.00)
Mean Trend AIC	219.13 (0.00)	248.08 (0.00)	283.14 (0.00)	272.02 (0.00)	255.35 (0.00)
Rate Trend AIC	212.78 (0.00)	241.30 (0.00)	275.31 (0.00)	264.26 (0.00)	248.34 (0.00)
WN AIC	195.97 (0.02)	<b>195.97</b> ( <b>0.43</b> )	<b>195.97</b> ( <b>0.48</b> )	<b>195.97</b> ( <b>0.44</b> )	<b>195.97</b> ( <b>0.48</b> )

Corrected Akaike Information Criteria (AICc)					
	Z-res. T. Deception	Z-res. Ex. Foraging	Z-res. Innovation	Z-res. Tool Use	Z-res. Social Learning
Delta AICc	190.30 (0.37)	198.34 (0.15)	200.04 (0.07)	198.07 (0.18)	198.33 (0.17)
BM AICc	217.31 (0.00)	246.27 (0.00)	281.32 (0.00)	270.21 (0.00)	253.53 (0.00)
OU AICc	<b>190.11</b> ( <b>0.40</b> )	198.27 (0.16)	198.34 (0.17)	198.01 (0.18)	198.33 (0.17)
Pagel's λ AICc	191.44 (0.21)	197.35 (0.25)	197.51 (0.26)	198.02 (0.18)	198.34 (0.17)
Early Burst AICc	219.50 (0.00)	248.46 (0.00)	283.52 (0.00)	270.35 (0.00)	255.72 (0.00)
Kappa AICc	200.88 (0.00)	213.31 (0.00)	229.46 (0.00)	220.82 (0.00)	228.24 (0.00)
Mean Trend AICc	219.50 (0.00)	248.46 (0.00)	283.52 (0.00)	272.40 (0.00)	255.72 (0.00)
Rate Trend AICc	213.16 (0.00)	241.68 (0.00)	275.68 (0.00)	264.63 (0.00)	248.72 (0.00)
WN AICc	196.15 (0.02)	<b>196.15</b> ( <b>0.45</b> )	<b>196.15</b> ( <b>0.50</b> )	<b>196.15</b> ( <b>0.46</b> )	<b>196.15</b> ( <b>0.50</b> )

**3.2. Macroevolutionary model comparison with Nonresidualized cognitive abilities**

Based on both AIC and AICc weights, *tactical deception* was found to have evolved under an OU selection regime. This result suggests that either via directional or stabilizing selection, the evolution of non-residualized tactical deception gravitated toward an optimal point on the phenotypic landscape. Although the AIC weights supported the OU macroevolutionary model for *extractive foraging*, the model comparison based on the AICc weights equally supported both the Delta and OU models. Under the Delta model, *extractive foraging* evolved at an initially accelerated rate and subsequently experienced a decline in its evolutionary rate. This model also suggests *extractive foraging* evolved as a lineage-specific adaptation.

Alternatively, the OU model revealed that this cognitive ability evolved toward a phenotypic optimum, either via stabilizing or

directional selection. In contrast to the previous cognitive abilities, the macroevolutionary model comparison identified white noise (WN) as the most likely model for *innovation*, based on both AIC and AICc weights. This result is not unexpected due to the trait's low phylogenetic signal, an outcome often associated with developmental or ecological effects (Pennell et al., 2014). Across both AIC and AIC weights, Pagel's λ was found to fit the evolution of *tool use* the best. This model follows the notion of rejecting a BM model but acknowledging that other macroevolutionary models not currently considered in the model comparison could have a better statistical fit. Due to its limited phylogenetic signal, WN fitted *social learning* as the best macroevolutionary model. Given its reliance on social transmission of information and knowledge, ecological and ontogenetic effects are expected to contribute substantially to the persistence of these traits over time. Although this model assumes that all lineages in the phylogenetic tree share the same starting value at the root of the tree, over time, social learning appears to have evolved independently within each species. Based on both AIC and AICc weights, the model comparison revealed that Pagel's λ fitted *G* the best. Although previous examinations have found that both the OU and Delta models had an adequate fit (Fernandes et al., 2020), by expanding the number of alternative models, the present study found strong support for Pagel's λ as the most likely selection regime characterizing the macroevolution of *G* (Table 2).

**3.3. Macroevolutionary model comparison with Residualized cognitive abilities**

Removing the *G* factor variance from *tactical deception* yielded indications that the residuals of this cognitive ability evolved toward a phenotypic optimum either via stabilizing or directional selection. These results (in Table 3) strongly suggest that tactical deception residuals, net of *G*, were subject to a unique macroevolutionary selection regime, distinct from those characterizing *G* and other residualized cognitive abilities (Fig. 1). Furthermore, the better fit of the OU model relative to the WN model indicates that a specialized aspect of tactical deception likely has an adaptive dimension, evolving in a modular fashion. This pattern remained consistent across both AIC and AICc weights. Since controlling for the *G* variance diminished the intensity of the cognitive abilities' phylogenetic signal, the analyses reveal that the *extractive foraging* and *tool use* residuals fit the WN model better than other competing alternatives. These differences were also captured by the ancestral character reconstructions evaluating the macroevolutionary process of these two cognitive abilities net of *G* (Figs. 2 and 3). Removing the *G* factor variance from *innovation* and *social learning* did not alter the results found using the unresidualized indicators, whereby the model comparison supported the WN model over the alternative models. The ancestral character reconstruction also evidenced these changes after residualizing the latter cognitive indicators for *G* variance (Figs. 4 and 5). These results are expected given that these two cognitive abilities exhibited low phylogenetic signals, even before residualizing them for *G*. It is worth noting that the decrease of the residuals' phylogenetic signal, relative to the unresidualized cognitive abilities, did not differ across cognitive subdomains such as physical or social cognition. Except for social learning, most traits were influenced by the residualization procedure, with an average loss of 0.248 points across all five cognitive abilities. The consistent support for the WN model in analyses using the latter four ability residuals revealed that other factors, such as ecological or ontogenetic effects, may be primarily responsible for the persistence of these traits.

**3.4. Phylogenetic generalized least squares models**

Table 4 summarizes the results of several Phylogenetic Generalized Least Squares (PGLS) models exploring the effect of socioecological and neuroanatomical indicators on the unresidualized and *G* residualized cognitive abilities. The omnibus test evaluating the influence of diet

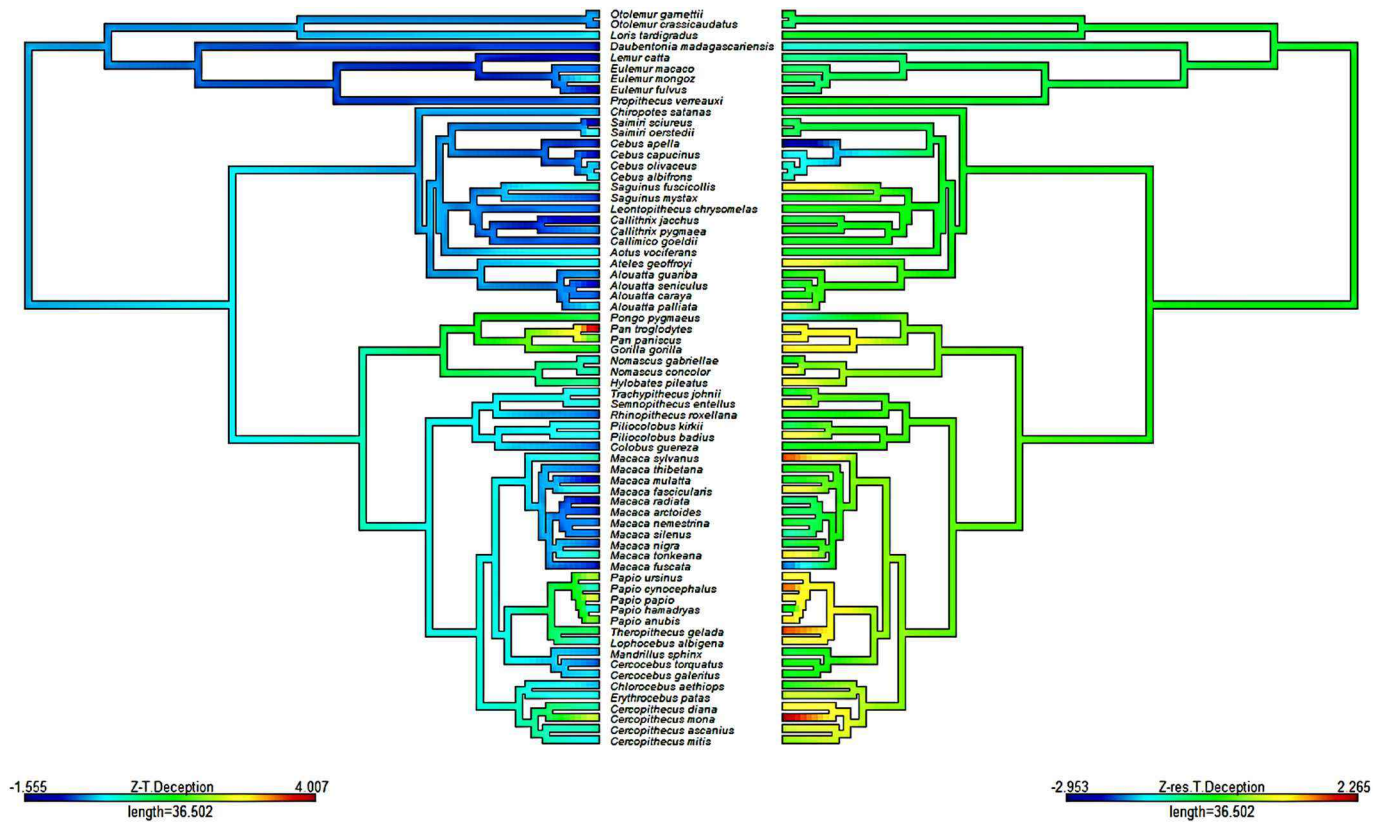


Fig. 1. Ancestral character reconstructions illustrating the phylogenetic difference between unresidualized and G residualized cognitive abilities: Standardized tactical deception (left) relative to standardized tactical deception residuals (right).

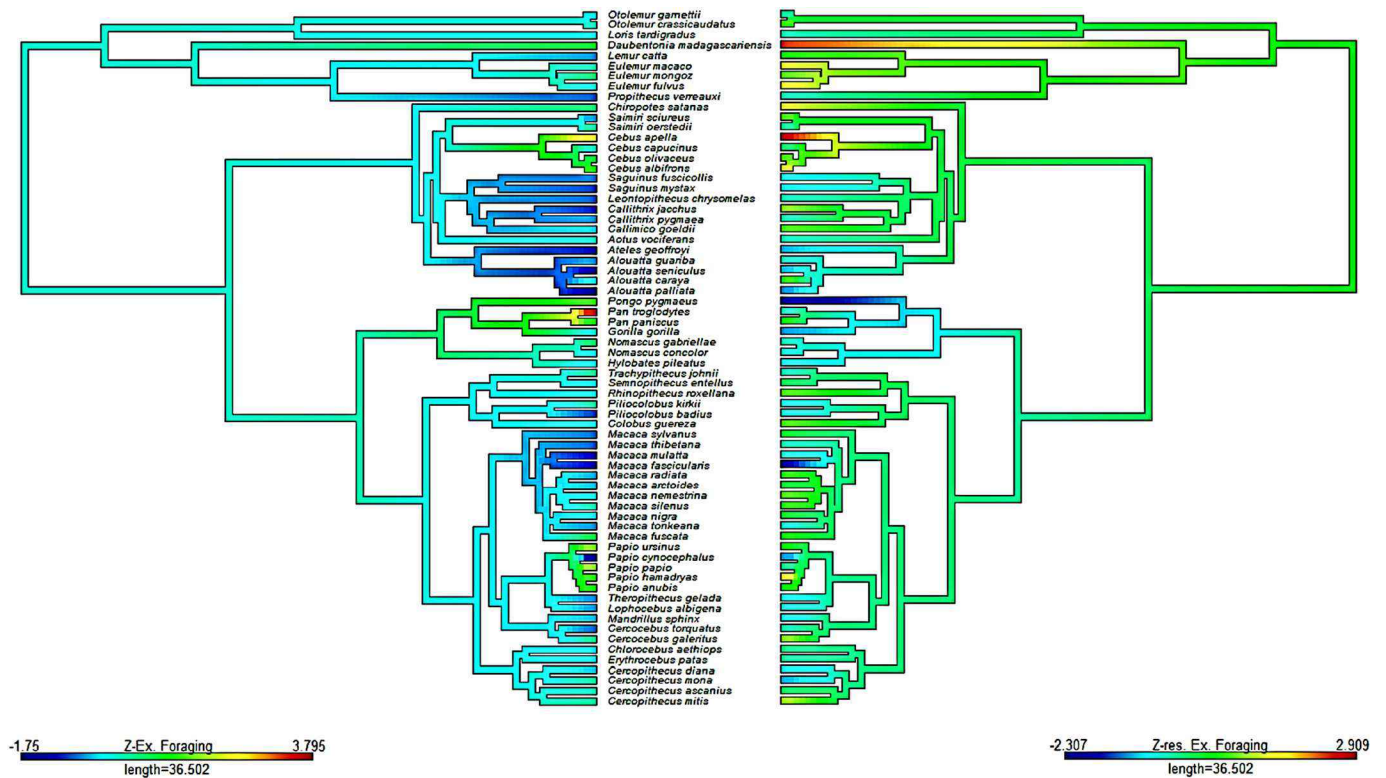


Fig. 2. Ancestral character reconstructions illustrating the phylogenetic difference between unresidualized and G residualized cognitive abilities) Standardized extractive foraging (left) relative to standardized extractive foraging residuals (right).

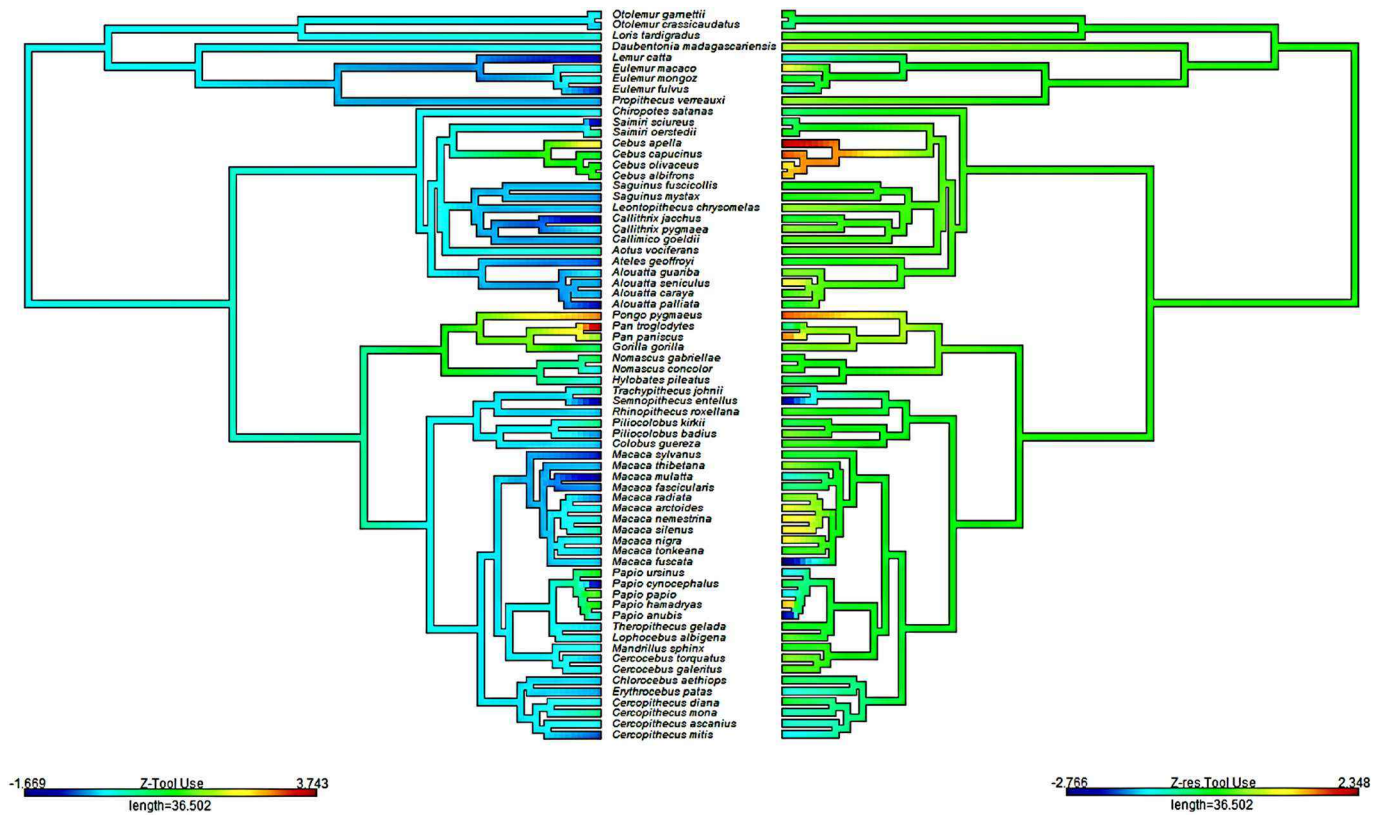


Fig. 3. Ancestral character reconstructions illustrating the phylogenetic difference between unresidualized and G residualized cognitive abilities: Standardized tool use (left) relative to standardized tool use residuals (right).

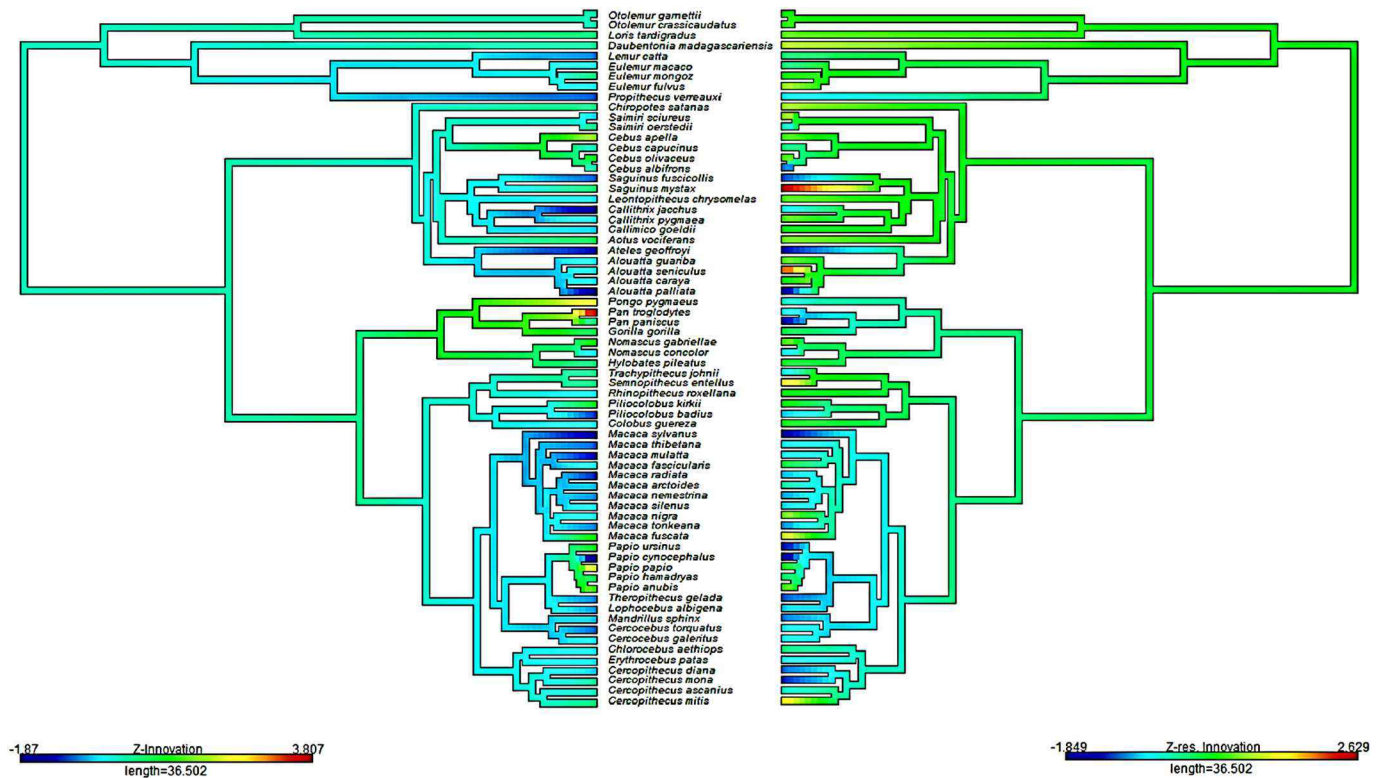


Fig. 4. Ancestral character reconstructions illustrating the phylogenetic difference between unresidualized and G residualized cognitive abilities: Standardized innovation (left) relative to standardized innovation residuals (right).



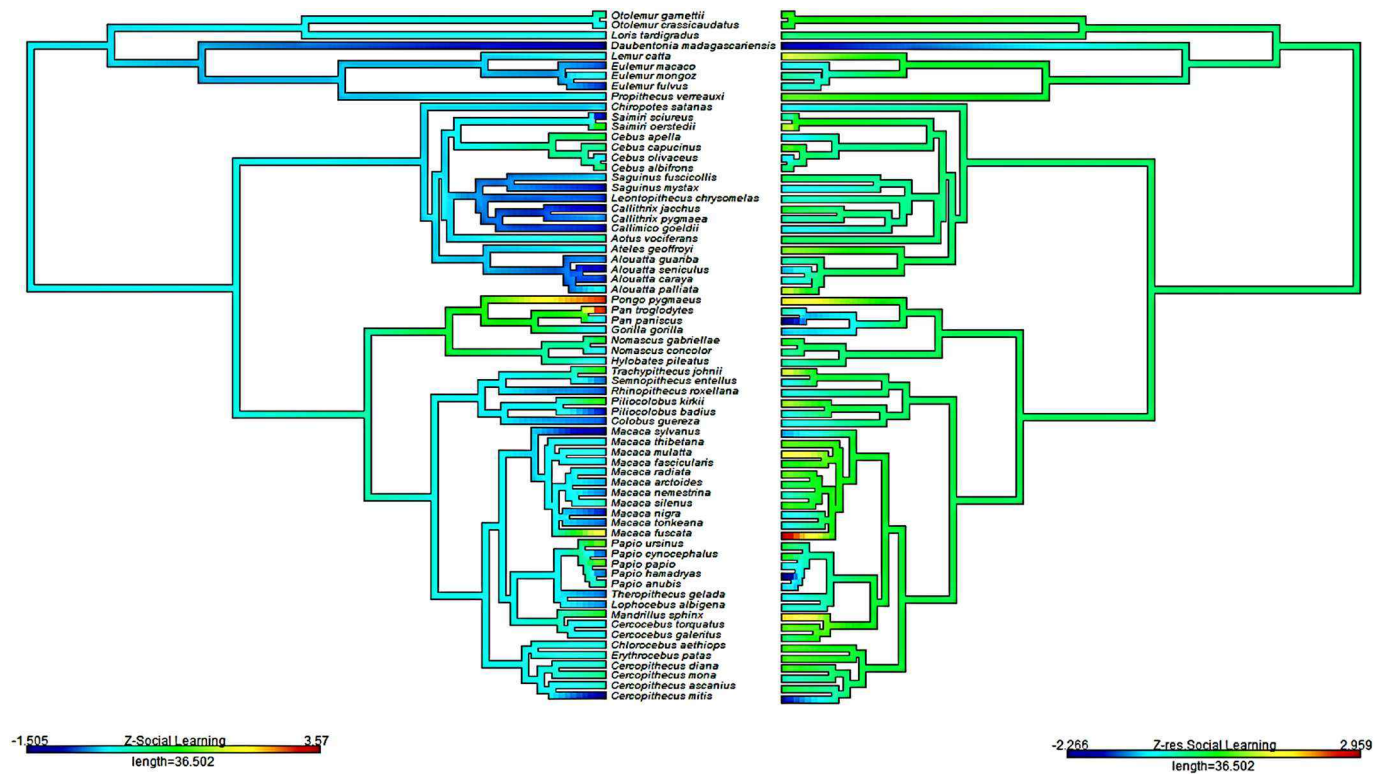


Fig. 5. Ancestral character reconstructions illustrating the phylogenetic difference between unresidualized and  $G$  residualized cognitive abilities: Standardized social learning (left) relative to standardized social learning residuals (right).

breadth, brain mass, and social group size on *tactical deception* reached statistical significance ( $F_{2, 55} = 14.22$ ;  $R^2 = 0.341$ ;  $p < .0001$ ). Brain mass had a positive and sizeable influence on this cognitive ability. The PGLS analysis with the *residualized tactical deception* values remained statistically significant ( $F_{2, 55} = 3.98$ ;  $R^2 = 0.126$ ;  $p = .0243$ ), with brain mass having a positive and significant influence. These results suggest that tactical deception net of  $G$  has an adaptive dimension that is positively predicted by brain mass. The omnibus test evaluating the influence of diet breadth, brain mass, and social group size on *extractive foraging* reached statistical significance ( $F_{2, 55} = 7.13$ ;  $R^2 = 0.206$ ;  $p = .0018$ ). Diet breadth and brain mass positively predicted this cognitive ability. The PGLS analysis with the *residualized extractive foraging* values remained statistically significant ( $F_{2, 55} = 4.00$ ;  $R^2 = 0.127$ ;  $p = .0239$ ), with adult brain mass negatively influencing this cognitive ability. The omnibus test evaluating the influence of diet breadth, brain mass, and social group size on *innovation* reached statistical significance ( $F_{2, 55} = 8.56$ ;  $R^2 = 0.237$ ;  $p = .0006$ ), with diet breadth and brain mass positively predicting this cognitive ability. The PGLS analysis with the *residualized innovation* values remained statistically significant ( $F_{2, 55} = 6.25$ ;  $R^2 = 0.185$ ;  $p = .0036$ ), with adult brain mass and social group size negatively influencing this cognitive ability. The omnibus test evaluating the influence of diet breadth, brain mass, and social group size on *tool use* reached statistical significance ( $F_{2, 55} = 13.22$ ;  $R^2 = 0.326$ ;  $p < .0001$ ) with both diet breadth and brain mass positively predicting this cognitive ability. Alternatively, the omnibus test exploring the influence of the latter predictors on the *residualized tool use* values did not reach statistical significance ( $F_{2, 55} = 1.14$ ;  $R^2 = 0.018$ ;  $p = .6111$ ). The omnibus test evaluating the influence of diet breadth, brain mass, and social group size on *social learning* reached statistical significance ( $F_{2, 55} = 4.91$ ;  $R^2 = 0.152$ ;  $p = .0109$ ) with brain mass positively predicting this cognitive ability. Alternatively, the omnibus test exploring the influence of the latter predictors on the *residualized social learning* values did not reach statistical significance ( $F_{2, 55} = 1.14$ ;  $R^2 = 0.040$ ;  $p = .3276$ ).

#### 4. Discussion

Among the unresidualised abilities, tactical deception and tool-use both mirror  $G$  in terms of their affinity for the Pagel's  $\lambda$  macroevolutionary mode. This suggests a relatively conservative selection regime that is not evolving the trait toward an optimum, and is likely merely proxying selection acting at the level of  $G$ . The other abilities seem to be characterized by divergent macroevolutionary modes. Specifically, extractive foraging is best characterized by the OU model, which indicates selection is acting to optimize the trait level across species consistent with the action of either stabilizing or directional selection. Innovation and social learning on the other hand are best characterized by the WN model. This is a non-phylogenetic source of variation among species, associated with the influence of ontogenetic and ecological factors. Brain mass positively predicted variation in tactical deception, extractive foraging, innovation, and social learning. In addition to this it was found that diet breadth positively predicted variation in extractive foraging, innovation, and tool use.

The results from the analyses employing unresidualised abilities are not especially informative, as many of their associations might be confounded with those of  $G$ . The differential affinity of some of these abilities (considered independently) for different macroevolutionary modes and socioecological and neuroanatomical correlates suggest that when  $G$  is applied to solving problems in specific domains, different patterns of selection might start to act on the trait within those narrow ecological contexts. This might also explain why diet breadth is positively correlated with innovation, tool use, and extractive foraging, but not with social learning and tactical deception. The need to efficiently secure a wide-array of food-stuffs might be expected to favor manifestations of  $G$  in the form of greater innovation and tool-use capability, especially as it pertains to the use of physical object-based innovations in hunting (i.e., sticks for ant-fishing). More social cognition-oriented manifestations of  $G$ , such as tactical deception and social learning, might not be as useful in terms of accessing a wider variety of food-

**Table 4**

Phylogenetic Generalized Least Square Models, with hierarchical partitioning of variance, evaluating the influence of diet breadth, adult brain mass, and social group upon unresidualized and *G* residualized cognitive abilities.

<i>Z-T. Deception</i>					<i>Z-res T. Deception</i>				
Predictors	$\beta$	Std. Error	F-value	Pr(>F)	Predictors	$\beta$	Std. Error	F-value	Pr(>F)
<i>Z-Diet breadth</i>	0.044	0.111	0.80	0.3738	<i>Z-Diet breadth</i>	-0.148	0.128	1.01	0.3189
<i>Z-Adult Brain mass (g)</i>	0.552	0.118	27.33	0.0000	<i>Z-Adult Brain mass (g)</i>	0.273	0.136	6.22	0.0157
<i>Z-Social group size</i>	0.065	0.118	0.30	0.5842	<i>Z-Social group size</i>	0.116	0.136	0.73	0.3971
<i>Z- Ex. Foraging</i>					<i>Z-res Ex. Foraging</i>				
Predictors	$\beta$	Std. Error	F-value	Pr(>F)	Predictors	$\beta$	Std. Error	F-value	Pr(>F)
<i>Z-Diet breadth</i>	0.294	0.122	7.11	0.0100	<i>Z-Diet breadth</i>	0.223	0.128	1.93	0.1700
<i>Z-Adult Brain mass (g)</i>	0.294	0.130	6.91	0.0111	<i>Z-Adult Brain mass (g)</i>	-0.329	0.136	4.28	0.0433
<i>Z-Social group size</i>	0.063	0.130	0.24	0.6289	<i>Z-Social group size</i>	0.182	0.136	1.79	0.1866
<i>Z-Innovation</i>					<i>Z-res Innovation</i>				
Predictors	$\beta$	Std. Error	F-value	Pr(>F)	Predictors	$\beta$	Std. Error	F-value	Pr(>F)
<i>Z-Diet breadth</i>	0.281	0.119	7.97	0.0066	<i>Z-Diet breadth</i>	0.197	0.123	2.76	0.1025
<i>Z-Adult Brain mass (g)</i>	0.385	0.127	7.55	0.0081	<i>Z-Adult Brain mass (g)</i>	-0.148	0.132	4.49	0.0387
<i>Z-Social group size</i>	-0.161	0.127	1.60	0.2109	<i>Z-Social group size</i>	-0.301	0.131	5.26	0.0256
<i>Z-Tool Use</i>					<i>Z-res Tool Use</i>				
Predictors	$\beta$	Std. Error	F-value	Pr(>F)	Predictors	$\beta$	Std. Error	F-value	Pr(>F)
<i>Z-Diet breadth</i>	0.226	0.112	6.52	0.0135	<i>Z-Diet breadth</i>	0.073	0.135	0.43	0.5161
<i>Z-Adult Brain mass (g)</i>	0.512	0.120	2.02	0.0000	<i>Z-Adult Brain mass (g)</i>	0.109	0.145	0.50	0.4824
<i>Z-Social group size</i>	-0.038	0.119	0.10	0.7522	<i>Z-Social group size</i>	-0.037	0.144	0.07	0.7974
<i>Z-Social Learning</i>					<i>Z-res Social Learning</i>				
Predictors	$\beta$	Std. Error	F-value	Pr(>F)	Predictors	$\beta$	Std. Error	F-value	Pr(>F)
<i>Z-Diet breadth</i>	0.053	0.126	0.60	0.4404	<i>Z-Diet breadth</i>	-0.194	0.134	2.18	0.1454
<i>Z-Adult Brain mass (g)</i>	0.391	0.134	9.15	0.0038	<i>Z-Adult Brain mass (g)</i>	-0.023	0.143	0.06	0.8037
<i>Z-Social group size</i>	-0.036	0.134	0.07	0.7870	<i>Z-Social group size</i>	-0.026	0.142	0.03	0.8532

stuffs, however larger brains may well promote these specific manifestations of *G*, in that they may enhance the capacity for primates to capture and model complex social dynamics. The predominance of the WN model in characterizing the narrow macroevolution of innovation and social learning is logical, given the dependence of these abilities upon 'epigenetic' vertical (e.g., cultural and ecological) transmission of acquired knowledge, which might be expected to condition manifestations of *G* in these domains. The affinity of extractive foraging for the OU model suggests that within this narrow ecological domain, *G* is under optimizing selection. This might be because for any given niche there is a consistently strong fitness premium associated with the capacity to efficiently and maximally exploit dietary resources.

More informative, and more interesting are the macroevolutionary modes and associations involving the residualised abilities. As these are free of the confounding influence of *G*, their macroevolutionary properties can be said to better capture the sorts of processes that condition the evolution of distinct cognitive modules. The residual of tactical deception was found to exhibit an affinity for the OU model, which indicates that, unlike in the case of the unresidualised form of this ability, there is a tendency toward adaptive optimization. This suggests that independently of *G* (which shares an affinity for a common macroevolutionary model with unresidualised tactical deception), there is highly domain specific selection acting on tactical deception favoring the formation of a primary module (where individuals within a species are optimized with respect to some domain specific aspect of tactical deception). The positive association of residualised tactical deception with brain size suggests that, wholly independently of *G*, larger brained

primates are better tactical deceivers, possibly because larger brains can accommodate greater cortical real-estate that can be given over to the development of a dedicated primary module. The other ability residuals all show greater affinity for the WN model. This suggests that variation among primate species with respect to these, independently of *G* is largely a function of processes operating via ontogeny, likely in response to ecological contingencies. This pattern is consistent with these being secondary modules, whose domain specific manifestations owe more to non-phylogenetic processes acting in ontogenetic time (such as via vertical transmission pathways). More interesting still are the unusual affinities that these secondary modules exhibit for neuroanatomical and social-ecological factors. The residuals of extractive foraging and innovation both exhibit negative associations with brain mass, and social group size negatively predicts the innovation residual. An implication of these findings is that the positive association of unresidualised extractive foraging and innovation with brain mass is entirely a function of *G*. Independently of this, *smaller* brains seem to better facilitate manifestations of these secondary modules. One possible explanation for this is that among small brained primates, the need to innovate and be efficient extractive foragers might be more contingent upon predictable ecological pressures acting on ontogeny (smaller brained primates tend to be found occupying narrower and less variable habitats characterized by smaller home ranges and more limited activity periods; Powell, Isler, & Barton, 2017). The existence of low environmental variability might therefore allow for these narrow secondary modules to be cultivated in the absence of higher-levels of *G*. A similar logic might obtain in the case of the negative association between social group size and the innovation

residual. Constrains on overall group size might be reflective of the action of ecological processes that facilitate the development of an innovation-specific secondary module independently of the action of *G*.

It is clear that measures of cognitive ability across species are a highly meaningful source of information on macroevolutionary processes and correlates. This holds even in cases where the abilities are residualised for their common-factor variance. The use of such residuals furthermore permits an empirically novel distinction to be made between primary and secondary modules via their affinities for different macroevolutionary modes. In keeping with the overarching theme of this special issue, it is therefore demonstrated that in the case of comparative psychology, going beyond (in this case) *G* yields substantially novel information that adds additional depth to the understanding of the evolution of cognition.

## References

- Allen, W. L., Street, S. E., & Capellini, I. (2017). Fast life history traits promote invasion success in amphibians and reptiles. *Ecology Letters*, *20*, 222–230.
- Burkart, J. M., Schubiger, M. N., & van Schaik, C. P. (2017a). The evolution of general intelligence. *Behavioral & Brain Sciences*, *40*, Article e192.
- Burkart, J. M., Schubiger, M. N., & van Schaik, C. P. (2017b). Future directions for studying the evolution of general intelligence. *Behavioral & Brain Sciences*, *40*, Article e224.
- Butler, M. A., & King, A. A. (2004). Phylogenetic comparative analysis: A modeling approach for adaptive evolution. *The American Naturalist*, *164*, 683–695.
- Byrne, R. W., & Whiten, A. (1990). Tactical deception in primates: The 1990 database. *Primate Report*, *27*, 1–101.
- Capellini, I., Baker, J., Allen, W. L., Street, S. E., & Venditti, C. (2015). The role of life history traits in mammalian invasion success. *Ecology Letters*, *18*, 1099–1107.
- Creighton, M. J., Greenberg, D. A., Reader, S. M., & Mooers, A. Ø. (2021). The role of behavioral flexibility in primate diversification. *Animal Behaviour*, *180*, 269–290.
- Deaner, R. O., van Schaik, C. P., & Johnson, V. E. (2006). Do some taxa have better domain-general cognition than others? A meta-analysis of nonhuman primate studies. *Evolutionary Psychology*, *4*, 149–196.
- Ducatez, S., Sol, D., Sayol, F., & Lefebvre, L. (2020). Behavioral plasticity is associated with reduced extinction risk in birds. *Nature Ecology & Evolution*, *4*, 788–793.
- Felsenstein, J. (1973). Maximum-likelihood estimation of evolutionary trees from continuous characters. *American Journal of Human Genetics*, *25*, 471–492.
- Fernandes, H. B. F., Peñaherrera-Aguirre, M., Woodley of Menie, M. A., & Figueredo, A. J. (2020). Macroevolutionary patterns and selection modes for general intelligence (*G*) and for commonly used neuroanatomical volume measures in primates. *Intelligence*, *80*, Article 101456.
- Fernandes, H. B. F., Woodley, M. A., & te Nijenhuis, J. (2014). Differences in cognitive abilities among primates are concentrated on *G*: Phenotypic and phylogenetic comparisons with two meta-analytical databases. *Intelligence*, *46*, 311–322.
- Gibson, K. R. (1986). Cognition, brain size, and the extraction of embedded food resources. In J. G. Else (Ed.), *Primate evolution* (pp. 95–103). New York: Cambridge University Press.
- Gibson, K. R., Gibson, K. R., & Ingold, T. (Eds.). (1994). *Tools, language and cognition in human evolution*. Cambridge: Cambridge University Press.
- Gorsuch, R. L. (1983). *Factor analysis* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.
- Harmon, L. J., Losos, J. B., Jonathan Davies, T., Gillespie, R. G., Gittleman, J. L., Bryan Jennings, W., ... Mooers, A. Ø. (2010). Early bursts of body size and shape evolution are rare in comparative data. *Evolution: International journal of organic. Evolution*, *64*, 2385–2396.
- Heldstab, S. A., Isler, K., Burkart, J. M., & van Schaik, C. P. (2019). Allomaternal care, brains, and fertility in mammals: Who cares matters. *Behavioral Ecology & Sociobiology*, *73*, 1–13.
- Jones, K. E., Bielby, J., Cardillo, M., Fritz, S. A., O'Dell, J., Orme, C. D. L., & Purvis, A. (2009). PanTHERIA: A species-level database of life history, ecology, and geography of extant and recently extinct mammals: Ecological Archives E090-184. *Ecology*, *90*, 2648.
- Kamilar, J. M., & Cooper, N. (2013). Phylogenetic signal in primate behaviour, ecology and life history. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *368*, 20120341.
- Lefebvre, L. (2011). Taxonomic counts of cognition in the wild. *Biology Letters*, *7*, 631–633.
- Lefebvre, L., Juretic, N., Nicolakakis, N., & Timmermans, S. (2001). Is the link between forebrain size and feeding innovations caused by confounding variables? A study of Australian and north American birds. *Animal Cognition*, *4*, 91–97.
- Lefebvre, L., Reader, S. M., & Sol, D. (2004). Brains, innovations, and evolution in birds and primates. *Brain, Behavior and Evolution*, *63*, 233–246.
- Lefebvre, L., Whittle, P., Lascaris, E., & Finkelstein, A. (1997). Feeding innovations and forebrain size in birds. *Animal Behaviour*, *53*, 549–560.
- Liudecke, D. (2018). Sjmisc: Data and variable transformation functions. *Journal of Open Source Software*, *3*, 754.
- Maclean, E. L., & Nunn, C. L. (2017). Phylogenetic approaches for research in comparative cognition. In J. Call (Ed.), *APA handbook of comparative psychology: Vol 1. Basic concepts, methods, neural substrate, and behavior*. Washington, DC: American Psychological Association.
- Nicolakakis, N., & Lefebvre, L. (2000). Innovation rate and forebrain size in birds of western Europe: Feeding, nesting and confounding variables. *Behaviour*, *137*, 1415–1429.
- Nunn, C. L. (2011). *The comparative approach in evolutionary anthropology and biology*. Chicago, IL: University of Chicago Press.
- Nunn, C. L., & Samson, D. R. (2018). Sleep in a comparative context: Investigating how human sleep differs from sleep in other primates. *American Journal of Physical Anthropology*, *166*, 601–612.
- Opie, C., Atkinson, Q. D., Dunbar, R. I., & Shultz, S. (2013). Male infanticide leads to social monogamy in primates. *Proceedings of the National Academy of Sciences*, *110*, 13328–13332.
- Pagel, M. (1999). Inferring the historical patterns of biological evolution. *Nature*, *401*, 877–884.
- Peñaherrera, A. M., & Fernandes, H. B. (2021). Phylogenetic analysis within comparative psychology. In T. K. Shackelford, & V. A. Weekes-Shackelford (Eds.), *Encyclopedia of Evolutionary Psychological Science*. Cham: Springer. [https://doi.org/10.1007/978-3-319-19650-3\\_3605](https://doi.org/10.1007/978-3-319-19650-3_3605).
- Pennell, M. W., Eastmen, J. M., Slater, G., Brown, J. W., Uyeda, J. C., FitzJohn, R. G., ... Harmon, L. J. (2014). Gieger v2.0: An expanded suite of methods for fitting macroevolutionary models to phylogenetic trees. *Bioinformatics*, *30*, 2216–2218.
- Plavcan, J. M., & van Schaik, C. P. (1997). Intrasexual competition and body weight dimorphism in anthropoid primates. *American Journal of Physical Anthropology*, *103*, 37–68.
- Powell, L. E., Isler, K., & Barton, R. A. (2017). Re-evaluating the link between brain size and behavioral ecology in primates. *Proceedings of the Royal Society B: Biological science*, *248*, 1–101, 20171765. Primate report, 27.
- Reader, S. M., Hager, Y., & Laland, K. N. (2011). The evolution of primate general and cultural intelligence. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *366*, 1017–1027.
- Reader, S. M., & Laland, K. N. (2002). Social intelligence, innovation, and enhanced brain size in primates. *Proceedings of the National Academy of Sciences, USA*, *99*, 4436–4441.
- Revell, L. J. (2012). Phytools: An R package for phylogenetic comparative biology (and other things). *Methods in Ecology & Evolution*, *3*, 217–223.
- Shultz, S., Opie, C., & Atkinson, Q. D. (2011). Stepwise evolution of stable sociality in primates. *Nature*, *479*, 219–222.
- Sol, D., Duncan, R. P., Blackburn, T. M., Cassey, P., & Lefebvre, L. (2005). Big brains, enhanced cognition, and response of birds to novel environments. *Proceedings of the National Academy of Sciences*, *102*, 5460–5465.
- Varricchio, D. J., Moore, J. R., Erickson, G. M., Norell, M. A., Jackson, F. D., & Borkowski, J. J. (2008). Avian paternal care had dinosaur origin. *Science*, *322*, 1826–1828.
- Whiten, A., & Byrne, R. W. (1988). Tactical deception in primates. *Behavioral & Brain Sciences*, *11*, 233–244.
- Wich, S. A., & Nunn, C. L. (2002). Do male “long-distance calls” function in mate defense? A comparative study of long-distance calls in primates. *Behavioral Ecology & Sociobiology*, *52*, 474–484.
- Woodley of Menie, M. A., Fernandes, H. B. F., te Nijenhuis, J., Peñaherrera Aguirre, M., & Figueredo, A. J. (2017). General intelligence is a source of individual differences between species: Solving an anomaly. *Behavioral & Brain Sciences*, *40*, Article e223.
- Woodley of Menie, M. A., & Peñaherrera-Aguirre, M. (2022). General intelligence as a major source of cognitive variation among individuals of three species of lemur, uniting *g* with *G*. *Evolutionary Psychological Science*. <https://doi.org/10.1007/s40806-021-00304-x>
- Wynn, T. (1988). Tools and the evolution of human intelligence. In R. W. Byrne, & A. Whiten (Eds.), *Machiavellian intelligence* (pp. 271–284). Oxford: Oxford University Press.
- Zhang, Z. (2016). Multiple imputation with multivariate imputation by chained equation (MICE) package. *Annals of Translational Medicine*, *4*, 30.