

1 **Beyond brain size**

2

3 Corina J Logan^{1*}, Shahar Avin², Neeltje Boogert³, Andrew Buskell⁴, Fiona R. Cross^{5,6}, Adrian
4 Currie², Sarah Jelbert⁷, Dieter Lukas¹, Rafael Mares^{8,9}, Ana F Navarrete¹⁰, Shuichi Shigeno¹¹,
5 Stephen H Montgomery¹

6

7 ¹Department of Zoology, University of Cambridge, Downing Street, Cambridge, UK, CB2 3EJ,

8 ²Center for the Study of Existential Risk, University of Cambridge, David Attenborough
9 Building, Pembroke Street, CB23QY

10 ³Centre for Ecology and Conservation, Daphne du Maurier building, University of Exeter,
11 Penryn Campus, Penryn, Cornwall, TR10 9FE

12 ⁴Department of History and Philosophy of Science, University of Cambridge, Free School Lane,
13 Cambridge, UK, CB2 3RH

14 ⁵School of Biological Sciences, University of Canterbury, Private Bag 4800, Christchurch, New
15 Zealand

16 ⁶International Centre of Insect Physiology and Ecology, Thomas Odhiambo Campus, P.O. Box
17 30, Mbita Point, Kenya

18 ⁷Department of Psychology, University of Cambridge, Downing Street, Cambridge, UK, CB2
19 3EB

20 ⁸Department of Anthropology, University of California, Davis, CA 95616, USA

21 ⁹Smithsonian Tropical Research Institute, Balboa, Ancon, Panama

22 ¹⁰Centre for Biodiversity, School of Biology, University of St Andrews, Harold Mitchell
23 Building, St Andrews, KY16 9TH

1

- 24 ¹¹Department of Biology and Evolution of Marine Organisms, Stazione Zoologica Anton Dohrn,
25 Villa Comunale, 80121 Naples, Italy
26 * Corresponding author
27 Keywords: Brain evolution, cognition, comparative method, neuroethology, intelligence

28 ABSTRACT

29 Despite prolonged interest in comparing brain size and behavioral proxies of ‘intelligence’ across
30 taxa, the adaptive and cognitive significance of brain size variation remains elusive. Central to
31 this problem is the continued focus on hominid cognition as a benchmark, and the assumption
32 that behavioral complexity has a simple relationship with brain size. Although comparative
33 studies of brain size have been criticized for not reflecting how evolution actually operates, and
34 for producing spurious, inconsistent results, the causes of these limitations have received little
35 discussion. We show how these issues arise from implicit assumptions about what brain size
36 measures and how it correlates with behavioral and cognitive traits. We explore how
37 inconsistencies can arise through heterogeneity in evolutionary trajectories and selection
38 pressures on neuroanatomy or neurophysiology across taxa. We examine how interference from
39 ecological and life history variables complicates interpretations of brain-behavior correlations,
40 and point out how this problem is exacerbated by the limitations of brain and cognitive measures.
41 These considerations, and the diversity of brain morphologies and behavioral capacities, suggest
42 that comparative brain-behavior research can make greater progress by focusing on specific
43 neuroanatomical and behavioral traits within relevant ecological and evolutionary contexts. We
44 suggest that a synergistic combination of the ‘bottom up’ approach of classical neuroethology
45 and the ‘top down’ approach of comparative biology/psychology within closely related but
46 behaviorally diverse clades can limit the effects of heterogeneity, interference, and noise. We
47 argue this shift away from broad-scale analyses of superficial phenotypes will provide deeper,
48 more robust insights into brain evolution.

49 1) Motivation

50 In 1856, unusual bones were unearthed in a German limestone quarry in the Neander valley.
51 Though later identified as the first Neanderthal specimen, the initial identification of the bones as
52 a distinct species was highly contested (Madison, 2016). Indeed, the unusual size and features of
53 the skull led to a range of questions: might the bones be from a pathological human or a new
54 variety of fossil ape? One of the key novel methods employed to answer these questions was
55 ‘craniometry’: the measurement of morphological features of the skull and their relationships
56 (Goodrum, 2016; Madison, 2016). Employing these methods, Schaaffhausen (1858) found that
57 the skull likely housed a large brain, falling within the range of contemporary humans. To many
58 scientists at the time (though not all), the fossil’s brain size was evidence enough for the human,
59 or human-like, status of the fossils (cf. Madison, 2016).

60 The distinctive brain morphology—and associated behavioral features—of hominins, and
61 *Homo sapiens* in particular, continue to fascinate. Indeed, comparisons between human and non-
62 human brains remain a central investigative target in contemporary comparative research. Such
63 investigations have expanded the brain measurement tool-kit: in addition to craniometry, metrics
64 such as brain size relative to body size (e.g., encephalization quotients), absolute brain size, and
65 neuronal density are now common (Montgomery, 2017).

66 One of the central motivations for the continued research into brain measurement is its
67 potential to reveal links between neuroanatomical structures and cognitive capabilities. Yet, just
68 as debates about the special status of Neanderthals hinged upon the size of its cranium relative to
69 humans, contemporary debates on the evolution of brain size and complex behavior have tended
70 to privilege measures where humans come out on top. This bias has been built into a number of
71 ‘monolithic’ general hypotheses (Barton, 2012) that claim links between measures of brain size

72 and a diverse range of proxy-measures of complex behavior, such as ‘social’ intelligence
73 (Dunbar & Shultz, 2007a, 2007b), ‘cultural’ intelligence (Tomasello, 1999; van Schaik &
74 Burkart, 2011; van Schaik, Isler, & Burkart, 2012), ‘general’ intelligence (Burkart, Schubiger, &
75 van Schaik, 2016; Reader, Hager, & Laland, 2011), and behavioral drive (Navarrete, Reader,
76 Street, Whalen, & Laland, 2016; Wyles, Kunkel, & Wilson, 1983). In each of these cases, *Homo*
77 *sapiens* emerges as the presumed pinnacle of a trajectory of brain evolution that correlates with
78 increasing behavioral flexibility, intelligence, or socialization. Yet, both the significance of brain
79 size and the interpretation of the correlated behaviors as more ‘complex’ or ‘cognitive’ remain
80 poorly elucidated.

81 Here, we build on arguments made by Healy and Rowe (2007), unpack problems
82 associated with using proxies, bring recent evidence from molecular techniques into the debate,
83 and develop a framework that incorporates bottom-up and top-down approaches to advance the
84 field. We argue that a more fruitful approach to linking brain measures and cognition is to de-
85 emphasize coarse-grained notions of ‘intelligence’ and whole-brain measurements in favor of i)
86 taxa-specific measurements of brains and ecologically meaningful behaviors, and ii) ‘bottom-up’
87 extrapolation of intraspecies measures based on phylogenetic context. This means measuring
88 ecologically relevant features of brains and behaviors directly, rather than using proxies, at the
89 within-species level. Then comparing closely related species to understand the relationships
90 between traits, which will inform the even broader taxonomic scaling to make cross-species
91 generalizations based on these validated correlations. Central to this is a movement away from
92 *Homo sapiens* as the measuring stick for evaluating the neuroanatomical features and behavioral
93 capabilities of other animals.

94 Below, we introduce a wide variety of research that examines brains and behavior across
95 various phyla and discuss how lessons learned from disparate taxa can inform the way we
96 interpret brain evolution, even among more familiar taxa such as vertebrates. Our aim is to
97 emphasize the advantages and disadvantages of the different metrics, methods, and assumptions
98 in this field. We review criticisms leveled against comparative studies of brain size, but go
99 further by establishing why the recognized limitations arise. By doing so, we show why broad-
100 stroke narratives struggle to capture the wide diversity of neuroanatomical features and
101 behavioral capacities in animals. As a result, we argue that a more targeted ‘bottom-up’ approach
102 that measures brains and behaviors at the intraspecies level to investigate cognitive,
103 neuroanatomical, and behavioral diversity is needed to fully understand how behavioral
104 complexity emerges from neural systems, and how well, or poorly, brain size reflects this
105 variation.

106 **2) Limitations of research on brain size and cognition**

107 Interpreting how variation in brain size might be related to variation in cognition involves a set
108 of assumptions that are frequently made in comparative studies:

- 109 • Brain size can be measured with negligible error
- 110 • Investing in a larger brain comes at a cost of investing in other tissues and/or life history
111 traits
- 112 • Scaling relationships between brain size and body size are conserved within and across
113 species
- 114 • Brain regions scale uniformly with total brain size
- 115 • Brain size scales with neuron number
- 116 • Cognitive abilities are discretely coded in the brain
- 117 • Cognitive abilities can be unambiguously ascertained by measuring behavior

- 118 • Brain size is directly and linearly associated with variation in cognition
- 119 • Selection on cognitive abilities and brain measures acts uniformly across species

120 These assumptions are applied uniformly both across and within species. The validity of these
121 assumptions has previously been challenged by Snell (1892) and Healy and Rowe (2007), and
122 we provide additional arguments in this section. First, the use of brain size as a trait makes
123 implicit assumptions about how brains develop and evolve (see §2.1). Second, when correlating
124 brain size and a measure of cognition we make assumptions about how selection acts on, or for,
125 either trait (see §2.2). Finally, measuring cognition inevitably requires making some assumptions
126 about the nature of behavioral complexity and what we view as a ‘cognitive’ trait (see §2.3). In
127 each case, the lack of data supporting the validity of these assumptions directly limits our
128 capacity to make reliable inferences about the link between brain size and cognition.

129

130 **2.1 Assumptions and limitations of what brain size measures**

131 Brain size may seem an easy neuroanatomical trait to measure, and the ease of obtaining a data
132 point for a species, using one to a few specimens, renders it a historically useful starting point for
133 many studies (Healy & Rowe, 2007; Jerison, 1985). However, brain size has also become the end
134 point for many studies, with the variability of this trait becoming a target for evolutionary
135 explanation. Large databases are populated by both individual measures and species’ brain size
136 averages, which are used to examine cross-species correlations between brain size and a number
137 of other traits. Researchers look to these databases for answers to questions such as *What is the*
138 *significance of a large brain? What are the costs, and what are the benefits?* (e.g., Aiello &
139 Wheeler, 1995; Armstrong, 1983; Harvey & Bennett, 1983; Isler & van Schaik, 2009; Nyberg,
140 1971). Cross-species correlations reveal that relative brain size (brain size relative to body size)
141 is putatively associated with a range of life history and ecological traits. For example, relative

142 brain size may correlate positively with longevity (a benefit) and negatively with fecundity (a
143 cost) in mammals (Allman, McLaughlin, & Hakeem, 1993; Deaner, Barton, & van Schaik, 2003;
144 González-Lagos, Sol, & Reader, 2010; Isler, 2011; Isler & van Schaik, 2009; Sol, Székely, Liker,
145 & Lefebvre, 2007). Crucially, however, these correlations are not necessarily independent nor
146 consistent across taxa; for example, relative brain size and longevity do not significantly
147 correlate in strepsirrhine primates (lemurs and lorises; Allman et al., 1993). Other analyses
148 suggest the relationship may be a consequence of developmental costs rather than an adaptive
149 relationship (Barton & Capellini, 2011). Such inconsistencies in applicability and explanation
150 raise the question: are we failing to accurately measure and explain brain size and associated
151 traits?

152 Despite Healy and Rowe's (2007) warning, studies reporting cross-species correlations
153 between brain size measures and various behavioral and life-history traits continue to
154 accumulate. This is in spite of recent evidence falsifying many of the assumptions listed in §2
155 (see Montgomery, 2017 for a review). For example, brain size does not scale linearly with body
156 size within (Rubinstein, 1936) or across species (e.g., Fitzpatrick et al., 2012; Montgomery et al.,
157 2013; Montgomery, Capellini, Barton, & Mundy, 2010), brain regions do not scale uniformly
158 with total brain size across species (§4.1; e.g., Barton & Harvey, 2000; Farris & Schulmeister,
159 2011; Gonzalez-Voyer, Winberg, & Kolm, 2009), brain size does not uniformly scale with
160 neuron number across taxa (§4.1; Herculano-Houzel, Catania, Manger, & Kaas, 2015; Olkowicz
161 et al., 2016), brain size does not necessarily translate into cognitive ability (see §2.3, §3.2, Box
162 1), and brain size is not consistently related to variables of interest even within species (see §2.2;
163 e.g., there are sex differences with regard to brain size and its relationship with cognition
164 [Kotrschal et al., 2013, 2014]; and fitness and longevity [Logan, Kruuk, Stanley, Thompson, &

165 Clutton-Brock, 2016]). Therefore, a research program that relies on one or more of these
166 assumptions is limited in its ability to make reliable inferences about what brain size measures
167 and what it means when it correlates (or not) with other traits.

168

169 **2.2 Does selection act on brain size?**

170 Attempts to explain variation in brain size often implicitly assume that natural selection acts
171 directly on brain size. In vertebrates this assumption has been given added traction from models
172 exploring how brain development may shape patterns of evolution that place greater emphasis on
173 the conservation of brain architecture. This renders brain size a potent target of selection, in
174 contrast to selective adaptation of particular brain regions (see §4.2). Artificial selection
175 experiments further highlight the capacity for selection to directly act on brain size (e.g., Atchley,
176 1984; Kotrschal et al., 2013). For example, artificial selection for small and large brain size in
177 guppies (*Poecilia reticulata*) produced a grade-shift in the scaling relationship between brain and
178 body size, resulting in ~15% differences in relative brain size between selection lines (Kotrschal
179 et al., 2013). While the resulting large- and small-brained guppies differed in several traits,
180 including performance in learning tasks (Kotrschal et al., 2013, 2014) and survival (Kotrschal et
181 al., 2015), almost all of these correlations between behavioral performance and brain size were
182 either test context- or sex-dependent (Kotrschal et al., 2013, 2014, 2015; van der Bijl, Thyselius,
183 Kotrschal, & Kolm, 2015).

184 These various trade-offs and sex-specific effects suggest the selection landscape in
185 natural populations may routinely be more complex than under laboratory conditions. Several
186 recent studies of variation in brain composition among closely related populations or species that
187 are isolated by habitat reveal heritable divergence in particular brain components rather than

188 overall size (Gonda, Herczeg, & Merilä, 2011; Montgomery & Merrill, 2017; Park & Bell,
189 2010). Indeed, a recent analysis of brain morphology in wild guppies suggests selection may
190 frequently favor changes in the size of specific brain regions, although in this case a role for
191 plasticity has not been ruled out (Kotrschal, Deacon, Magurran, & Kolm, 2017). Focusing solely
192 on overall brain size, as in the artificial selection experiments, might mask the co-occurring
193 changes within the brain that underlie the observed differences in behavior. Accordingly,
194 adaptive responses to ecological change may involve alterations in specific components of neural
195 systems, presumably in response to selection on particular behaviors. This latter distinction is
196 important. It is unlikely that selection ever acts ‘on’ any neuroanatomical trait because what
197 selection ‘sees’ is variation in the phenotypes produced by neural systems (i.e., behavior), and
198 the energetic and physiological costs associated with their production.

199 Understanding how brain size relates to selection for behavioral complexity or cognition
200 is therefore a two-step process. First, we must understand how behavioral variation emerges from
201 variation in neural systems. Second, we must understand how this variation in neural systems
202 relates to overall brain size. Currently, our ability to take these steps is limited by a paucity of
203 well understood examples of behavioral variation in natural populations. However, existing
204 examples provide some insight into the limitations of total brain size as a unitary trait. Recent
205 studies of the proximate basis of schooling behavior in fish (Greenwood, Wark, Yoshida, &
206 Peichel, 2013; Kowalko et al., 2013), and burrowing (Weber, Peterson, & Hoekstra, 2013) and
207 parental behaviors in *Peromyscus* mice (Bendesky et al., 2017), suggest outwardly unitary
208 ‘behaviors’ may often be composites of genetically discrete behavioral phenotypes whose
209 variation is determined by independent neural mechanisms. For example, in a recent study of
210 parental care in *Peromyscus*, Bendesky and colleagues (2017) used a quantitative genetics

211 approach to map variation in the propensity to care between two species. They revealed a highly
212 modular genetic architecture, with some loci affecting general care behavior and other loci
213 affecting specific traits such as nest building, and a high propensity for sex-specific effects.
214 Variation in specific traits such as nest building in males can be linked with particular brain
215 regions, in this case, the hypothalamus.

216 The role of *FOXP2*, a transcription factor, in language development and evolution
217 provides a second informative example. *FOXP2* is generally highly conserved across mammals,
218 but has two human-specific amino acid substitutions which were likely fixed by positive
219 selection (Enard et al., 2002). Disruption of this gene in humans severely impacts language
220 acquisition (Lai, Fisher, Hurst, Vargha-Khadem, & Monaco, 2001), suggesting it plays a key role
221 in vocal learning. Insertion of the human version of the protein into the mouse genome affects
222 the development of particular cell types in the basal ganglia without gross effects on brain size or
223 morphology (Enard et al., 2009), yet leads to improved performance on certain learning tasks
224 and may have a broader role in motor learning (Schreiweis et al., 2014).

225 These examples illustrate how variation in behaviors that are considered by many
226 comparative studies to be correlated with whole brain size may in fact arise from localized
227 changes in brain development that do not affect total size. This may be the kind of incremental
228 variation selection plays with over small evolutionary time scales, and it is reasonable to assume
229 that the accumulation of this kind of change makes a significant contribution to species
230 differences in total brain size. While there is some evidence that genetic pleiotropy (i.e., genetic
231 variation in loci that cause phenotypic variation in multiple traits) can drive shifts in multiple
232 behaviors, in many cases selection may be able to shape specific behavioral traits independently
233 of other behaviors. Global measures of brain size and cognition both suffer from a lack of

234 support for the underlying assumption that the correlated variation in their component parts
235 stems from a shared proximate basis.

236

237 **2.3 Assumptions and limitations about what brains mean for cognition**

238 While ongoing efforts attempt to discover and characterize specific links between particular
239 behavioral and cognitive traits and brain size, this work exists alongside a highly visible thread
240 running through the literature that takes brain size itself as a proxy for ‘intelligence’ (e.g.,
241 Jerison, 1969; Table 1). For example, Jerison (1973, 1985) hypothesized that species showing
242 behaviors assumed to require increased neural processing required the evolution of a larger brain
243 relative to their body size to create ‘extra neurons’ for those seemingly complex behaviors.

244

245 **Table 1:** Examples of cross-species comparisons that link cognition and brain size, and a
246 description of the caveats about the ability to draw inferences due to the limitations involved in
247 measuring both traits.

Study	Cognitive measure	Brain size measure that correlates with the cognitive measure	Caveats
MacLean et al. (2014)	Attempts to measure ‘self control’ in 36 species of birds and mammals, using the A-not-B test and the cylinder test	Absolute and relative brain size	It is unclear whether self control was measured See Box 1
Benson-Amram et al. (2016)	Problem solving in 39 species of carnivores, using a	Relative brain size Relative brain size + regional brain	Problem solving is an ambiguous concept, with success being heavily influenced by differences

	puzzle-box that could be opened to obtain food.	volume	<p>in motivation, neophobia and animals' typical behavioral repertoires, amongst other things. Additionally, here, the puzzle boxes could be opened in multiple ways: i.e. by sliding a latch on the box open, or rolling the box over (which could cause the latch to slide open without being manipulated). It is unclear what success on this task really means in terms of underlying cognitive ability.</p> <p>Captive animals in zoos were tested, which likely have variable rearing histories and experiences with enrichment or solving problems</p> <p>Motivation identified as a confounding variable</p> <p>Given the large variation among individuals within species on cognitive tests (see Table 1), a sample size of one to a few individuals is not likely to be representative of the species (mean=4.9)</p>
Deaner et al. (2007; 2006)	General cognition as indicated by successful performance of 24 primate genera on many different tests: string pulling, detouring, invisible displacement, object discrimination, reversal learning, oddity learning, sorting, delayed	Absolute brain size	<p>Mean brain and body mass were collected per genus rather than per species because cognition data were only available per genus. This scale is likely too broad (see §3.2).</p> <p>Cognitive test data were pooled when the same tests had been conducted on different species. However, inter-lab noise has been shown to mask any differences or similarities even among individuals of the same species (Thornton & Lukas, 2012)</p>

	response		Did not account for ecological differences among genera
			Many tasks draw on similar sensory-motor capacities, which questions whether ‘general cognition’ can be inferred
Reader et al. (2011)	General cognition as indicated by successful performance of 62 primate species on many different tests: extractive foraging, innovation frequency per species, social learning, tactical deception, tool use	Ratio of the neocortex and the rest of brain Ratio of neocortex + striatum and the brainstem Neocortex size	Some of the cognitive measures are proxies of the behavior in question (see §3.2) The method of correcting for research effort does not account for biases in the publication of the reports on which the data are based (see §3.2)
Herculano-Houzel (2017)	Tasks in Deaner et al. (2007) (above in this table), and cylinder and A not B tasks in MacLean et al. (2014) and Kabadayi et al. (2016)	Number of cortical or pallial neurons	See row above on Deaner et al. (2007) and Box 1

248

249 In discussing indicators of ‘cognition’, we first need to know when a behavior is

250 ‘cognitive’ or indicative of ‘complex cognitive abilities’ (sometimes referred to as ‘intelligence’

251 and often invoking the term ‘behavioral flexibility’ [Mikhalevich, Powell, & Logan, 2017];

252 Table 2). This is problematic because these terms are not defined well enough to test empirically

253 or even to properly operationalize, and therefore cannot be measured in a systematic way.

254 Appeals to ‘neural processing’ likewise suffer from ill definition and an inability for accurate

255 quantification in most contexts. Researchers studying animal behavior tend to avoid using the
256 term ‘intelligence’ due to its anthropocentric connotations, and instead often adopt Shettleworth’s
257 definition of cognition as “the mechanisms by which animals acquire, process, store, and act on
258 information from the environment. These include perception, learning, memory, and decision-
259 making” (Shettleworth, 2010, p. 4). However, this all-encompassing definition still does not
260 allow us to answer basic questions about the proximate machinery underlying ‘cognitive’ traits:
261 Is a behavior more ‘cognitively complex’ if it engages more neurons, or certain networks of
262 neurons, or neurons only in particular brain regions that are responsible for learning and
263 memory? Or should we think of neural processing in dynamic terms, such as the ‘flexibility’ of
264 neurons to abandon old connections and form new ones as task demands change? Is behavior
265 only considered to rely on complex cognition if it is flexible? There are no clear answers to these
266 questions because data are greatly lacking.

267 Indeed, it is nearly impossible to determine which behaviors require increased neural
268 processing when they are observed in isolation from real-time brain activity. Creative studies
269 using imaging technology can now measure behavior and brain activity at the same time, but
270 only in species that can be trained to remain stationary in an fMRI scanner (e.g., dogs: Andics et
271 al., 2016; pigeons: De Groof et al., 2013; see also Mars et al., 2014). However, without *a priori*
272 predictions about which neural measures indicate complex cognition, this will remain a process
273 of *post hoc* explanations and goal-post moving based on anthropocentric biases about which
274 species should be ‘intelligent’ (see Mikhalevich et al., 2017).

275 Burgeoning research in artificial intelligence and machine learning suggests the
276 correlation between raw computing power (‘brain size’) and ‘intelligence’ is unlikely to be
277 straightforward. For example, a machine-learning algorithm designed to solve a specific task

278 may indeed get a performance boost from a ‘bigger brain’ (i.e., utilizing more hardware, for
 279 example, when playing Go [Silver et al., 2016]). However, algorithmic improvements that create
 280 more efficient ways of forming ‘neuronal’ connections based on input data may account for even
 281 greater performance or speed improvements given fixed hardware. The effective utilization of
 282 hardware resources is itself an active research field within machine learning (e.g., Nair et al.,
 283 2015), hinting that a ‘bigger brain’ does not straightforwardly translate into greater speed or
 284 better performance.

285
 286 **Table 2:** Examples of experiments attempting to test cognition. We note that there may be
 287 additional confounds in such studies that are likely to have affected test performance; however,
 288 these cannot be ruled out until explicitly quantified and taken into account in analyses (see also
 289 Macphail, 1982).

Cognitive Tests	Attempting to test	Assumed cognitive measures might be confounded by	Examples of studies that investigate or review these cognitive tests
String pulling	Insight Learning speed Means-end understanding	Responses to perceptual-motor feedback Motivation to obtain reward Age Attention Rearing effects Dexterity Object permanence Exploration Neophobia/neophilia Side biases Visual acuity Salience of the stimuli	Jacobs & Osvath (2015)
Aesop’s Fable tube tests	Physical cognition Causal reasoning	Learning speed Motivation to obtain reward Object biases	Jelbert et al. (2014; 2015), Logan et al.

		Responses to perceptual-motor feedback Salience of the stimuli	(2014), Miller et al. (2016)
Cylinder and A not B tasks	Self control / inhibition	Neophobia/neophilia Exploration Dexterity Perseveration Salience of the stimuli Visual acuity/tracking (see Box 1)	MacLean et al. (2014), Jelbert et al. (2016)
Puzzle-boxes	Innovativeness Problem solving Creativity	Neophobia/neophilia Exploration Dexterity Motivation to obtain reward Behavioral repertoire size Perseveration Operant conditioning (likelihood of repeating actions that led to obtaining reward) Salience of the stimuli Age Dominance rank Sex	Benson-Amram & Holekamp (2012), Boogert et al. (2008), Thornton & Samson (2012)
Reversal learning	Learning speed Behavioral flexibility	Neophobia/neophilia Exploration Perseveration Learning speed Motivation to obtain reward Body condition Age Sex Reproductive hormonal state Habituation to captivity Salience of the stimuli	Boogert et al. (2011; 2010), O'Hara et al. (2015)
Trap-tube tasks	Tool-use Physical cognition Causal reasoning	Operant conditioning Learning speed Dexterity Neophobia Motivation to obtain the reward Motivation to avoid incorrect responses Inhibitory control Salience of the stimuli	Seed et al. (2006), Mulcahy & Call (2006)

291 Theoretical reflection within the field of artificial intelligence has provided alternative
292 definitions of intelligence that highlight the difficulties faced by cognitive ethologists. For
293 example, Legg and Hutter (2007) aim to provide a universal definition that could apply to
294 machine intelligence as well as human and non-human animal intelligence. Informally, their
295 definition suggests: “*Intelligence measures an agent’s ability to achieve goals in a wide range of*
296 *environments*”. Following Legg and Hutter’s definition (without committing to whether theirs is
297 definitive) clarifies several difficulties with the current approach to evaluating intelligence in
298 non-human animals, and subsequently our ability to relate it to brain size. More specifically:

- 299 1. Intelligence is goal-dependent. A behavior, no matter how complex, cannot be counted as
300 intelligent if it does not serve a clear goal. Yet, interpreting goal-orientation in non-
301 humans is inherently difficult, even under strict experimental conditions.
- 302 2. Intelligence is environment-dependent. Problematically, behavioral features often
303 associated with complex cognition such as innovation, planning and tool use may have
304 varying degrees of availability or relevance in different environments, which may affect
305 whether they are displayed or not, irrespective of the organism’s ability to display them.
- 306 3. Intelligence of an organism is displayed across a range of environments. The few
307 experimental setups usually used to quantify ‘intelligence’ in captive animals may
308 therefore be minimally informative; instead, the ability of an organism to achieve its
309 goals should be evaluated across the range of environments it is likely to encounter
310 within its lifetime.

311 Regardless of the validity of the definition, these three features - goal orientedness, environment
312 dependency and utility across heterogeneous conditions - highlight the practical limitations of
313 assessing cognition in animals. The focus on utility further illustrates why selection may favor

314 ‘simple’ behavioral solutions to a task, or why the expression of simple behavior does not
315 preclude the ability of an organism to identify and carry out more complex solutions in
316 alternative contexts. If cognition is something akin to problem-solving capacity, then we should
317 develop measures that pay careful attention to the **range** of problems animals face in their
318 **natural environments**, rather than transferring proxies of intelligence in humans that are
319 relevant to the problems humans face in human environments (see Box 2 for an example).

320 Nevertheless, many comparative studies do find associations between gross measures of
321 brain size and broadly descriptive behaviors. In what follows, we focus on two factors that
322 explain when and why the results of such comparative studies should be treated with caution:
323 biological heterogeneity, and statistical noise and interference.

324 **3) Why do these limitations of brain-behavior comparative studies arise? Noise and** 325 **interference**

326 The lack of consistency in results from comparative studies (see Healy & Rowe, 2007 for an
327 overview) strongly suggests some underlying variability in the relationship between brain size
328 and complex cognition. In attempting to understand the properties of a particular system, it is
329 useful to distinguish between *noise* (exogenous source) and *interference* (endogenous source;
330 Currie & Walsh, n.d. in review) as distinct kinds of confounds in brain-behavior correlations.
331 *Noise* limits our ability to accurately determine and measure the co-evolving brain structures and
332 cognitive abilities (this section), while interference results from the endogenous features of
333 systems interacting with one another (section 4). One major factor shaping limitations of
334 comparative studies of brain size and cognition is the influence of noise and the numerous
335 covariates influencing the reliability and power to detect true associations. Noise results from
336 exogeneous factors that undermine our capacity to extrapolate across data-points. Measurement

337 error is an inevitable source of noise in these studies because behavior is noisy (§3.1), the
338 behaviors observed might not directly reflect single specific cognitive abilities (§3.2), and the
339 feasibility of obtaining brain measures differs across species, thus limiting comparison (§3.3).
340 Reducing noise requires different experimental and analytical approaches.

341

342 **3.1 Measuring behavior is noisy because behavior is noisy**

343 Animal behavior depends on the integration of internal motivational states and external
344 environmental cues. Although many behaviors are largely stereotyped, the kinds of behavioral
345 traits routinely studied in comparative studies of cognition are not. The expression of behaviors
346 that we might interpret as ‘social cognition’ (such as theory of mind) or ‘physical cognition’
347 (such as tool use) depend on an individual’s internal state and perception of the external
348 environment, factors that are not readily assessed. This introduces a degree of stochasticity in an
349 animal’s behavioral expression and noise in our behavioral measurements.

350 For example, Japanese macaques (*Macaca fuscata*) provide a famous case of innovation.
351 In this case, two novel behaviors involving washing sweet potatoes before eating them and
352 separating grains from dirt by throwing them in the water, were innovated by a single female
353 (called Imo) and spread through a wild population via social transmission (cf. Allritz, Tennie, &
354 Call, 2013; Kawai, 1965). At a population level, the high rate of social transmission may be
355 impressive, but at an individual level does the innovativeness of Imo suggest some
356 neuroanatomical variation that supports more complex cognition and increased innovation?
357 Imo’s brain may be no more innovative than her peers: she may simply have been in the right
358 place at the right time or more receptive to reward stimuli. Whether inferred behavioral
359 categories such as innovation reflect population-level variation in cognition is therefore unclear.

360 The assumption that innovation and behavioral flexibility reflect similar cognitive processes is
361 extrapolated from anthropocentric concepts and experiences of innovation, but this clearly
362 requires empirical validation.

363 Developments in artificial intelligence and machine learning, especially in the field of
364 reinforcement learning, suggest one reason to be wary of the possible conclusion that Imo's
365 innovative ability is due to neuroanatomical variation at the intraspecies level: high levels of task
366 performance can be achieved by systems that combine trial-and-error learning with feedback on
367 their performance in the form of a reward (similar to reward-based associative learning). As
368 reported by Mnih and colleagues (2015), a system trained from raw pixel inputs and reinforced
369 using an environment-provided performance metric (game score) was able to achieve human-
370 level performance in Atari game play by iteratively searching for the patterns that maximize
371 game score. Silver and colleagues (2016) report another achievement of artificial intelligence,
372 namely of human-level performance on the game of Go, with gameplay that has been described
373 as 'creative' and 'innovative', by the artificial agent first learning to predict expert moves
374 (supervised learning) and then by improving performance through self-play (reinforcement
375 learning). These engineering achievements suggest that the combination of chance, feedback,
376 and repeated iterations (possibly over generations) could yield the same behavioral performance
377 by artificial intelligence, at least in narrow domains, as organisms associated with having
378 'complex cognition'.

379

380 **3.2 Measuring cognition through behavior is noisy because we use unvalidated proxies**

381 Cognition is unobservable and must be inferred from behavior (Box 1). Many of the 50+ traits
382 that have been correlated with brain size across species (Healy & Rowe, 2007) are proxy

383 measures of the actual trait of interest (e.g., the number of novel foraging innovations at the
384 species level is a proxy for individual-level behavioral flexibility). This would not be a problem
385 if proxies were validated by directly testing the link between the trait of interest and its
386 correlational or causal relationship with its proxy within a population. However, the proxies used
387 are generally not validated, which contributes noise and uncertainty about what the correlations,
388 or lack thereof, between these trait-proxies and brain size actually mean. An in-depth discussion
389 of innovation illustrates why and how unvalidated proxies are an issue for this field.

390 The hypothetical link between innovation frequency per species and their relative brain
391 size was originally proposed by Wyles and colleagues (1983). Lefebvre and colleagues (1997)
392 operationalized the term *innovation* to make it measurable and comparable across bird species,
393 defining it as the number of novel food items eaten and the number of novel foraging techniques
394 used per species as anecdotally reported in the literature (see also Overington, Morand-Ferron,
395 Boogert, & Lefebvre, 2009). Innovation is assumed to represent a species' ability to modify its
396 behavior in response to a change in its environment, and is therefore a trait-proxy for *behavioral*
397 *flexibility* (e.g., Overington et al., 2009; Reader & Laland, 2002; Sol & Lefebvre, 2000; Sol,
398 Timmermans, & Lefebvre, 2002). Behavioral flexibility is defined here as modifying behavior in
399 response to changes in the environment based on learning from previous experience
400 (Mikhalevich et al., 2017; Swaddle, 2016). Two challenges emerge from this conceptualization:
401 i) how to measure innovation, and ii) how to validate that innovation frequency per species really
402 is an accurate reflection of behavioral flexibility.

403 It is unclear how to calculate innovation frequency per species, or what its biological
404 significance is to the species in question. For example, Logan (unpublished data) tried to follow
405 standard methods (from Lefebvre et al., 1997; Overington et al., 2009) to quantify the number of

406 innovations in New Caledonian crows (*Corvus moneduloides*). Innovations were extracted from
407 anecdotal reports “if authors used terms like ‘opportunistic’, ‘novel’, ‘first description’,
408 ‘unusual’, ‘not noted before’, and ‘no previous mention in the literature’” (Lefebvre et al., 1997,
409 pp. 550–551). Technical innovations were also extracted from anecdotal reports and defined as
410 falling into one of these categories: “novel technique, novel technique in an anthropogenic
411 context, novel parasitic behaviour, novel commensal behaviour, novel mutualistic behaviour,
412 novel proto-tool behaviour, novel true tool behaviour and novel caching behaviour” (Overington
413 et al., 2009, p. 1002). Logan found two food-type innovations and 10 technical innovations.
414 However, it was unclear how distinct each technical innovation was (e.g., three involve
415 manipulation of *Pandanus* leaves). It was also unclear whether tools that were used in a similar
416 way, but made of different materials should count as separate innovations or the same innovation
417 (e.g., using non-stick materials in a stick-like manner). Finally, and most importantly, it became
418 clear that these innovations were only novel or unusual to the humans who saw crows
419 performing these behaviors; these behaviors are commonly performed by New Caledonian crows
420 across their natural habitat (e.g., Hunt & Gray, 2002) and are certainly not novel to them,
421 suggesting that innovation frequency databases (e.g., Overington et al., 2009) may contain many
422 similar cases of species-typical behaviors that had gone unnoticed to the human observer.
423 Therefore, it is also unclear what innovation frequency per species means to that species, which
424 further confounds the significance of innovation frequency per species.

425 Evidence has only recently become available to test the hypothesis that innovation
426 frequency per species is a reliable proxy for behavioral flexibility. A small number of
427 comparative studies have tested individuals from different species that vary in brain size and
428 innovation frequency (both are species-level measures) on the same test of behavioral flexibility.

429 Results showed that innovation frequency per species did not correlate with measures of
430 behavioral flexibility in individuals (Auersperg, Bayern, Gajdon, Huber, & Kacelnik, 2011;
431 Bond, Kamil, & Balda, 2007; Ducatez, Clavel, & Lefebvre, 2015; Jelbert et al., 2015; Logan,
432 2016a, 2016b; Logan, Harvey, Schlinger, & Rensel, 2016; Logan et al., 2014; Manrique, Völter,
433 & Call, 2013; Reader et al., 2011; Tebbich, Sterelny, & Teschke, 2010) or with species level
434 estimates of brain size (Cnotka, Güntürkün, Rehkämper, Gray, & Hunt, 2008; Ducatez et al.,
435 2015; Emery & Clayton, 2004; Isler et al., 2008; Iwaniuk & Nelson, 2003; Pravosudov & de
436 Kort, 2006) in predictable ways. The absence of consistent associations between intraspecific
437 measures of behavioral flexibility and species-level measures of innovation and brain size erodes
438 the logical basis of comparative studies across species. If behavioral flexibility is to be
439 considered a marker of cognitive ability, it should be measured directly in individuals of each
440 species rather than using unsupported species-level proxies such as the reported frequency of
441 innovation. The continued use of innovation frequency is due solely to convenience and data
442 availability. Although the first comparative studies using this metric provided promising
443 glimpses into brain evolution, the time has now come to descend to the within-species level to
444 understand the proximate origins of individual variation in behavioral flexibility. More generally,
445 despite a lack of validation that they accurately reflect the trait of interest, proxies of behavioral
446 traits are pervasive in the comparative brain size literature and introduce unknown amounts of
447 exogenous noise into cross-species analyses. This noise may generate spurious results, masking
448 ‘true’ patterns in the data and impeding their interpretation.

449

450 **3.3 Measuring brain size is noisy because it is more difficult than it seems**

451 If cross-species correlations indicate relationships that are actually present, these associations
452 should persist within species if we assume a direct relationship between brain size and behavior
453 in a given task. However, these associations often do not persist at the intraspecies level, which
454 may be due to extensive measurement errors in quantifying brain size within species, or to the
455 confounding effects of variation in a brain trait being attributable to multiple functions. Most
456 work on brain evolution has focused on overall brain size or changes in large regions of the
457 brain, such as the forebrain and the cerebellum (see review in Healy & Rowe, 2007; see also
458 Herculano-Houzel, 2012; Reader et al., 2011). However, volumetric measurements are
459 particularly noisy. We use primate brain data to illustrate the difficulties involved in obtaining,
460 preserving, and measuring brain volumes.

461 More is known about brain anatomy in primates than in other orders, yet volumetric
462 measurements of specific brain regions in this group are only available for a few species, and
463 some of these measurements are pooled from only a few individuals per species (Reader &
464 Laland, 2002). This introduces a large amount of noise because a species' average brain, or brain
465 region, volume might be biased due to sexual dimorphism or other variables that differ across
466 individuals (Montgomery & Mundy, 2013). Information on primate brain size is scarce and
467 primarily comes from captive individuals. Further, access to primate brains is limited to only a
468 few brain collections (Zilles, Amunts, & Smaers, 2011).

469 Additional complications arise in determining whether it is appropriate to correlate
470 behavioral data from wild individuals with morphological data (e.g., brain size) obtained from
471 captive individuals. Studies comparing the morphology of wild and captive animals have shown
472 that rearing conditions may influence body composition (e.g., skull shape, brain size, digestive
473 tract) after only a few generations (O'regan & Kitchener, 2005). In primates, brain mass is not

474 generally affected by captivity (Isler et al., 2008), but body mass is: some species become
475 heavier, while others become lighter due to inadequate diets (O'regan & Kitchener, 2005).

476 Furthermore, although brain size might not be affected by captivity, primate populations
477 of the same species that were reared under different captive conditions differ in cortical
478 organization (Bogart, Bennett, Schapiro, Reamer, & Hopkins, 2014). In macaques and humans,
479 there is evidence that individual differences in social network size correlate with amygdala
480 volume and areas related to this structure (Bickart, Wright, Dautoff, Dickerson, & Barrett, 2011;
481 Kanai, Bahrami, Roylance, & Rees, 2011; Sallet et al., 2011). Among individuals of the same
482 species, brain anatomy changes significantly with age (Hopkins, Cantalupo, & Tagliabue,
483 2007). Choosing individuals with closely matched histories can reduce noise in brain measures
484 that are introduced by individual differences in previous experience, but the noise involved in
485 brain volume measurements is most effectively controlled and minimized by obtaining large
486 sample sizes per species to acquire more reliable species averages. This problem is particularly
487 vexing when combining behavioral data sets from observations in the wild with neuroanatomical
488 data from captive populations.

489 Data collection methods can also compromise the quality of the data. Many reported
490 brain weights and brain volumes are actually proxies of these measures, obtained instead by
491 calculating endocranial volume from skulls, which are much easier data to collect (e.g., Isler et
492 al., 2008; Iwaniuk & Nelson, 2002). While endocranial volume has been shown to reliably
493 approximate brain volume across species of primates (Isler et al., 2008) and birds (Iwaniuk &
494 Nelson, 2002) and within species of birds (Iwaniuk & Nelson, 2002), this might not always be
495 the case. For example, Ridgway and colleagues (2016) suggest endocranial vascular networks
496 and other peripheral appendages can account for 8-65% of endocranial volume in cetaceans,

497 leading to a consistent overestimation of brain size that is more severe in some species than
498 others.

499 Because brains are valuable tissues, non-invasive methods such as magnetic resonance
500 imaging (MRI) are preferred for obtaining data on brain anatomy and function. Yet high-
501 resolution, high-quality MRIs from primate brains are difficult to obtain from live individuals.
502 Images obtained using in vivo techniques, where the animal is sedated for a short period of time
503 while scanning their brains, might be more accessible, but image quality and resolution is poorer
504 than in images obtained post-mortem (K. L. Miller et al., 2011). Post-mortem MRIs can have a
505 higher resolution and are therefore more suited to calculating volumes. However, even MRIs are
506 problematic because of other sources of noise that arise from brain extraction methods, including
507 the post-mortem delay between death and extraction and preservation, and the ‘age’ of the
508 preserved brain (i.e., how long a brain has been stored for; Grinberg et al., 2008; K. L. Miller et
509 al., 2011). While post-mortem MRI is the best method available for calculating brain volumes,
510 brain volume in itself is a noisy measure because of its unclear, and usually untested, relationship
511 with other variables of interest (see §3.2).

512 Given that volumetric brain measurements suffer from many additional sources of noise,
513 it may be more productive to focus instead on non-volumetric features of brain composition
514 (e.g., neuron density, grey matter as a measure of local connectivity, white matter as a measure of
515 long-distance connectivity). For example, transcranial magnetic stimulation is increasingly used
516 in humans to temporarily ‘knock out’ particular brain areas to understand their functionality and
517 relationship with behavior and cognition (e.g., Zatorre, Chen, & Penhune, 2007). This type of
518 (non-volumetric) technique allows one to elucidate causal relationships, which generates data of
519 a much higher quality because it is validated (i.e., not a proxy) and directly connected with the

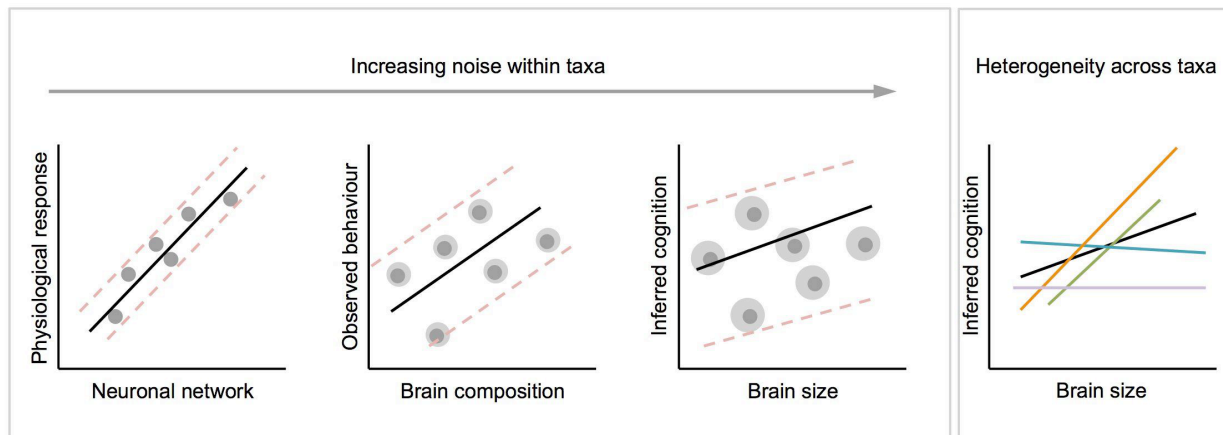
520 behavior under study, which greatly reduces measurement noise. Ideally, multiple methods
521 should be applied to the same system to determine whether different types of evidence arrive at
522 the same conclusions regarding brain-behavior relationships.

523 **4) Why do limitations in brain-behavior comparative studies arise? Evidence of** 524 **interference**

525 Interference occurs when systems consist of multiple interacting parts whose interactions tend to
526 be complex. A potentially useful way of understanding some critiques of brain size-cognition
527 comparative studies is to consider the ramifications of heterogeneity within and across species in
528 terms of their brain architectures and associated traits (e.g., behavior, cognition, life history;
529 Figure 1). If parts of the brain evolve in concert due to developmental coupling, for instance,
530 then interference from those components makes it difficult to isolate the evolutionary causes of
531 changes in brain size, or any of its components, over time. Similarly, if many ecological and life
532 history traits covary, identifying which factors drive changes in brain size is complicated by
533 autocorrelation between independent variables. Philosophers distinguish *heterogeneity* within
534 and between systems as a useful concept for framing the validity of comparisons (Elliott-Graves,
535 2016; Matthewson, 2011). Heterogeneity arises as a confounding factor in comparisons among
536 individuals and/or species when the components of a system (e.g., brain structures) differ (§4.1),
537 or when similar components exist but differ in scaling relationships or patterns of connectivity
538 (e.g., neuron density, neural network; §4.2). Treating brain size as a unitary trait assumes either
539 that the brain is a unitary trait or that any signal from a brain-behavior association is sufficient to
540 overpower the influence of heterogeneity on either trait. Comparisons of taxonomically diverse
541 neural systems can identify where similar brain architectures exist and where heterogeneity in
542 brain composition is masked by comparisons of brain size (§4.3). Interference in the form of

543 heterogeneity between systems occurs because of the complex interactions among life-history
544 and ecological factors that shape the co-evolution of cognitive abilities and particular brain
545 measures (§4.4).

546



547
548 **Figure 1: Effects of noise and heterogeneity on brain-behavior correlations as measures of a**
549 **biological trait (on both axes) become increasingly crude.** As measurements move away from
550 direct, quantitative data of primary biological processes both axes become increasingly noisy (as
551 indicated by the grey halos around each data point). The interaction between signal, noise, and
552 heterogeneity may result in contrasting correlations between taxonomic groups (indicated by
553 differently colored lines). When correlations are averaged across these groups the resulting
554 associations may retain little information.

555

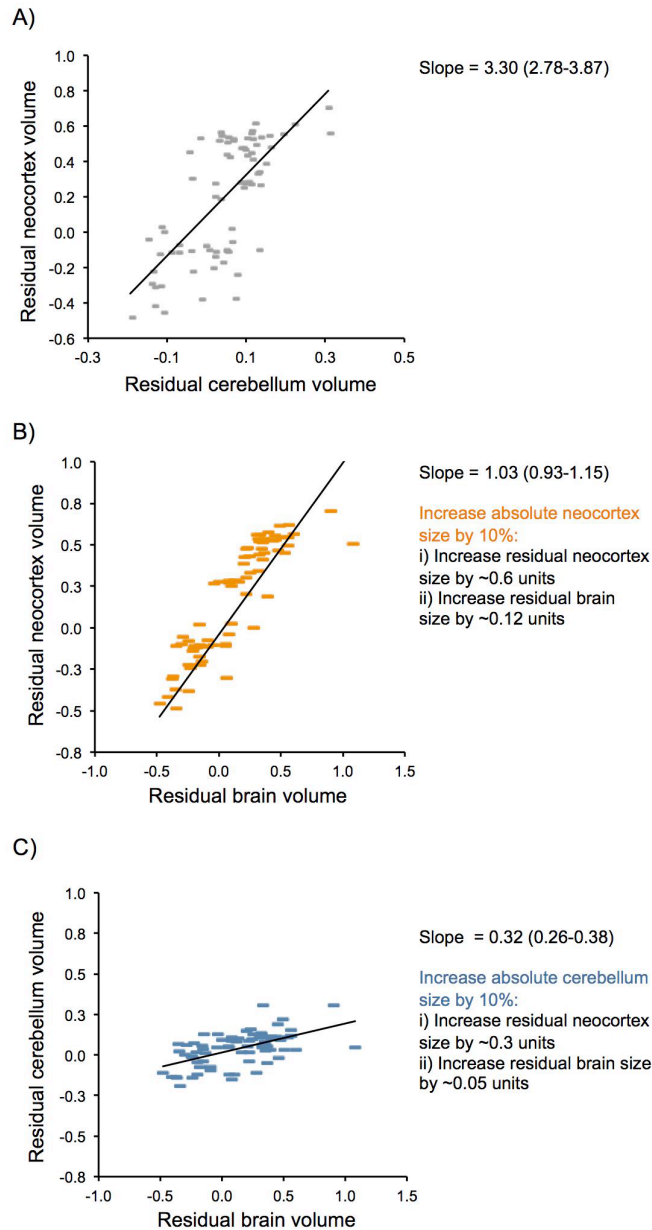
556 **4.1 Heterogeneity in brain composition within taxonomic groups: brains that appear**
557 **similar according to certain measures may actually be different**

558 Broad comparisons across phylogenetically disparate and ancient groups demonstrate how our
559 understanding of the presumed cognitive benefits of large brains is, at best, simplistic. The brain

560 architecture underlying ecologically relevant neural computation will depend on the behavioral
561 requirements of a task, the evolutionary history of the machinery that selection is building on,
562 and the strength of potentially opposing selective forces such as energetic, volumetric, and
563 functional trade-offs and constraints. Even across more closely related species, for example
564 among mammals, heterogeneity between brain structures introduces noise and variation that can
565 complicate brain-behavior relationships.

566 While some authors argue that the major axis of variation in mammalian brains is overall
567 size (e.g., Clancy, Darlington, & Finlay, 2001; Finlay, Darlington, & Nicastro, 2001), there is
568 ample evidence for variation in brain structure across species caused by brain region-specific
569 selection pressures, so called ‘mosaic brain evolution’ (Barton & Harvey, 2000). Accumulating
570 evidence among major taxonomic groups shows differences in brain composition (e.g., Kaas &
571 Collins, 2001; Workman, Charvet, Clancy, Darlington, & Finlay, 2013). When a behavior
572 generated by a specific brain structure is targeted by selection, the effect on total brain size will
573 depend on the scaling relationship between that brain structure and total brain size. For example,
574 one general trend across mammalian brain evolution is a correlated expansion of the neocortex
575 and cerebellum, which occurs independently of total brain size (Barton, 2012; Whiting & Barton,
576 2003). These structures share extensive physical connections and are functionally interdependent
577 (Ramnani, 2006), but, while they tend to co-evolve, both have evolved independently in some
578 evolutionary lineages (Barton & Venditti, 2014; Maseko, Spocter, Haagensen, & Manger, 2012).
579 Independent selection pressure on individual brain components such as the neocortex and
580 cerebellum do not have equal effects on overall brain size or measures of encephalization (Figure
581 2). Neocortex volume scales hyper-allometrically with brain volume (i.e., as brain size increases,
582 the proportion of neocortex tissue increases), while cerebellum volume, and several other major

583 brain components, scale hypo-allometrically with brain volume (Barton, 2012). As a result,
584 increases in neocortex volume have a disproportionate effect on brain volume compared with
585 similar proportionate increases in cerebellum size, largely due to differences in the scaling of
586 neuron density and white matter in the two structures (Barton & Harvey, 2000; Herculano-
587 Houzel, Collins, Wong, & Kaas, 2007). Variations in whole brain size, or measures of brain size
588 relative to body size, such as the encephalization quotient (Jerison, 1973), therefore essentially
589 correspond to variation in neocortex size and mask variation in other brain components, even
590 though the latter may be of great functional significance. The power of the comparative analysis
591 of brain-behavior associations is therefore limited when selection acts on non-cortical structures.
592 Even in vertebrates, mounting evidence suggests this will often be the case. For example, the
593 frequency of tool use in primates (Barton, 2012) and the complexity of nest structure in birds (Z.
594 J. Hall, Street, & Healy, 2013) have been linked with variation in relative cerebellum volume,
595 and spatial memory in birds has been linked with hippocampal volume (e.g., Krebs et al., 1996).
596



597
598 **Figure 2: Effects of brain component scaling on the contributions brain regions make to**
599 **brain expansion.** A) the size of the neocortex and cerebellum, once corrected for the size of the
600 rest of the brain, co-evolve with a positive scaling relationship. Both residual size of the
601 neocortex (B) and cerebellum (C), after correcting for the size of the rest of the brain, correlate
602 with the total brain size corrected for body size indicating both components contribute to
603 encephalization. However, the scaling relationships differ, such that any increase in absolute

604 neocortex volume has a greater influence on residual brain size compared to a similar increase in
605 absolute cerebellum volume (see also Barton, 2012).

606
607 When comparing brain size across species, further heterogeneity is apparent at the level
608 of the cellular composition of brain structures. Recent data on neuron number in brain regions of
609 birds and mammals have revealed extensive variation across taxonomic groups (Herculano-
610 Houzel et al., 2015). For example, primates have significantly higher neuron densities in the
611 neocortex and cerebellum than other closely related terrestrial mammals, while elephants have
612 substantially higher neuron densities in the cerebellum than other Afrotheria (e.g., golden moles,
613 elephants, and sea cows; Herculano-Houzel et al., 2015), and birds pack similar numbers of
614 neurons as found in primates into volumetrically more restricted brains (Olkowicz et al., 2016).
615 Because neurons and their synaptic connections are the basic computational units of any neural
616 system, if variation in brain, or brain region, volume does not consistently reflect variation in
617 neuron number, then any inference made about the cognitive significance of brain size is largely
618 invalid. To illustrate this effect, averaging across brain regions, a 1 gram brain that follows
619 primate neuron number-brain size scaling rules will contain ~26% more neurons than a brain that
620 follows the glire scaling rules (the clade including rodents; Herculano-Houzel et al., 2015). A 1
621 gram brain that follows psittacine (parrots) scaling rules will contain ~100% more neurons than a
622 brain that follows the glire scaling rules, and ~58% more than one that follows the primate
623 scaling rules (Olkowicz et al., 2016). Comparing brain size across these taxa thus erroneously
624 assumes that the computational output (based on neuron number) of these hypothetical brains
625 would be equal.

626 The assumption that brain volume is comparable and meaningful across species is
627 explicitly made in broad phylogenetic studies of cognitive ability (e.g., MacLean et al., 2014;
628 Box 1). Variation in brain structure and cellular composition strongly questions this assumption.
629 The effect of incorporating more fine-grained data, even if they are relatively crude, is apparent
630 in existing studies. For example, in Benson-Amram and colleagues' (2016) analysis of how
631 performance on a puzzle-box test is associated with brain size across 39 species of mammalian
632 carnivore, the addition of data on volumetric variation in brain structure significantly improved
633 their predictive model compared to one containing only brain volume. In a recent opinion piece,
634 Herculano-Houzel (2017) also argued that (cortical) neuron number outperforms total brain size
635 as a predictor of behavioral performance in self-control tests reported by MacLean and
636 colleagues (2014). The power of brain size as a causative predictor of cognitive performance is
637 therefore apparently vulnerable to the addition of only narrowly more fine-grained data. The sole
638 reasons for the continued focus on brain size as a unitary trait are convenience and data
639 availability. If we are to progress beyond a superficial understanding of brain-behavior
640 correlations, this justification must also be set aside now that new forms of more detailed data
641 are becoming increasingly available on how neuron number and connectivity vary across brain
642 regions and species.

643 **4.2 Deep convergence in brain architecture: brains that appear different according to** 644 **certain measures may actually be homologous**

645 At the broadest taxonomic scale, brain composition is remarkably diverse. For example,
646 comparative studies have traditionally focused on linking learning and memory with arachnid
647 protocerebrums (e.g., Meyer & Idel, 1977; Punzo & Ludwig, 2002), insect mushroom bodies
648 (e.g., Snell-Rood, Papaj, & Gronenberg, 2009), cephalopod vertical lobes (e.g., Grasso & Basil,

649 2009), the vertebrate pallium (e.g., Jarvis et al., 2005), and mammalian neocortices (e.g.,
650 Pawłowski, Lowen, & Dunbar, 1998). Despite their independent evolution, some research points
651 toward commonalities in the molecular and neural systems that function in heterogeneous brain
652 organizations across animal phyla. A combinatorial expression pattern of developmental control
653 genes suggests the deep origin of key learning and memory centers, including in the complex
654 sensory centers and cell types of the mushroom bodies of annelids and arthropods, and the
655 pallium of vertebrates (Tomer, Denes, Tessmar-Raible, & Arendt, 2010). Similarly, Pfenning and
656 colleagues (2014) proposed that vocal learning in birds and humans has evolved via convergent
657 modification of brain pathways and molecular mechanisms. Roth (2013, p. 292) proposed that
658 the centers for learning and memory in insect, octopus, avian, and mammalian brains share a
659 comparable associative network that “bring[s] the most diverse kinds of input into the same data
660 format and [integrates] the respective kinds of information.” These broad comparisons suggest
661 that such brain structures in taxonomically and anatomically diverse animals may share a number
662 of features, including high neuron density, and similar organizations with hierarchical
663 connectivity (G. Roth, 2013). Similarly, the vertebrate basal ganglia and insect central complex
664 have been shown to exhibit a deep homology, sharing similar network organizations,
665 neuromodulators, and developmental expression machineries (Strausfeld & Hirth, 2013).
666 Accordingly, divergent structures may have converged on similar architectures and
667 computational solutions to analogous behavioral challenges (Shigeno, 2017). By simplifying
668 brain measures by focusing only on size, we may miss out on opportunities to study how
669 convergences in behavior and complex neural systems can inform how cognition evolves.

670 Nevertheless, the heterogeneity identified by these studies may also provide useful
671 variation that can contribute to our understanding of brain and cognitive evolution. For example,

672 if neuronal density can vary independently of volume, why? And how does this impact the
673 functional properties of the pathways that produce complex behaviors associated with cognitive
674 prowess?

675 **4.3 Effects of size-efficient selection**

676 Compared with vertebrates, arthropods have tiny brains and vastly fewer neurons in their
677 nervous systems (Eberhard & Weislo, 2011), yet many insects and spiders display highly
678 sophisticated motor behaviors, social organizations, and cognitive abilities (Chittka & Niven,
679 2009; Box 2). For example, insects and spiders exhibit numerical cognition (Cross & Jackson,
680 2017; Dacke & Srinivasan, 2008; Rodríguez, Briceño, Briceño-Aguilar, & Höbel, 2015),
681 planning (Cross & Jackson, 2016; Tarsitano & Jackson, 1997), selective attention (Jackson & Li,
682 2004) and working memory (Brown & Sayde, 2013; Cross & Jackson, 2014; Zhang, Bock, Si,
683 Tautz, & Srinivasan, 2005)—all typically studied in vertebrates and considered cognitively
684 demanding (Chittka & Niven, 2009), illustrating that selection has favored highly efficient
685 neuronal systems in these taxa.

686 Although heterogeneity in brain systems limits the scope of comparative studies of brain
687 size, it also provides an opportunity to understand how selection acts on neural systems, and why
688 selection favors particular solutions over others. One key factor may be the role of size-efficient
689 selection and redundancy in nervous systems. Neurons are energetically expensive cells, and
690 their total cost scales predictably with the size of neural systems (Laughlin, de Ruyter van
691 Steveninck, & Anderson, 1998). Selection must therefore constantly trade-off behavioral
692 performance with energetic and computational efficiency. Exploring how these trade-offs are
693 resolved in real and artificial systems has the capacity to greatly inform why some animals invest

694 in larger brains, while others do not (Burns, Foucaud, & Mery, 2010; Chittka & Niven, 2009;
695 Chittka, Rossiter, Skorupski, & Fernando, 2012; Menzel & Giurfa, 2001).

696 While an imperfect analogy, researchers' experience with training artificial neural
697 networks provides an insight into how efficient neural networks can be constructed. Indeed,
698 researchers who aim to create an artificial network that serves as a pattern-learning machine have
699 been largely inspired by the organization of the cerebral cortex in mammals (Mnih et al., 2015).
700 This comparison between artificial networks and cerebral cortex organization was made more
701 notable with recent advances in deep convolutional neural networks (an artificial neural network
702 with a large number of intermediary layers, specialized in identifying patterns in perceptual
703 inputs) such as the deep-Q network (DQN). Beyond mammals, this layer-like organization can
704 also be identified in the brains of for example the common octopus and *Drosophila*, suggesting
705 that a common functionality of information processing patterns may be represented both in
706 artificial and biological neural networks (Shigeno, 2017).

707 One of the key messages from such research is that training large neural networks is still
708 difficult (Bengio, Simard, & Frasconi, 1994; Glorot & Bengio, 2010; Pascanu, Mikolov, &
709 Bengio, 2013). Even when training is successful, it requires a great deal of time and input data,
710 but more importantly, training too-large a network without the right algorithm often simply fails.
711 In artificial systems, this happens when feedback from the environment is used by the neural
712 network to determine certain flexible values of the computational architecture (e.g., connections
713 between artificial neurons). This problem scales up: greater numbers of flexible values (i.e.,
714 network parameters, which grow in tandem with 'brain size'), require greater amounts of input
715 data and increasingly complex algorithms. Such trade-offs are likely also faced by biological
716 organisms. Thus, in addition to the energetic costs of larger brains, there are also informational

717 costs (i.e., a need for more, better, and/or faster inputs) and computational costs (i.e., efficient
718 ways to use inputs, which may be architecturally difficult for natural selection to find) that limit
719 brain size and may channel the response to selection away from simple increases in the total size
720 of the system or brain.

721 The hand of size-efficient selection can also be seen in the network architecture of large
722 brains that display a ‘small-world’ topology (Ahn, Jeong, & Kim, 2006; Chen, Hall, &
723 Chklovskii, 2006), which minimizes energetically costly long-range connections in favor of
724 proportionally high local connectivity (Bullmore & Sporns, 2012; Buzsáki, Geisler, Henze, &
725 Wang, 2004; Watts & Strogatz, 1998). Yet, if network architecture is constrained by energetic
726 costs, then what does the evidence of variation in cellular scaling between brain components and
727 across species tell us about how brains evolve?

728 Variation in the scaling of neuron number with volume likely reflects differences in cell
729 size and patterns of connectivity between neurons. The low neuron density in the neocortex in
730 mammals, compared to that of the cerebellum, reflects the high proportion of the neocortex given
731 over to white matter that consists of mid- to long-range fibres connecting neurons (Ringo, 1991).
732 Variation in the pattern of neuronal connections, and integration between brain regions may help
733 explain variation in cellular scaling. Similar explanations may also apply to scaling differences
734 across taxa, with the high neuronal density of primates being associated with relatively smaller
735 volumes of white matter and connectivity (Ventura-Antunes, Mota, & Herculano-Houzel, 2013).
736 However, these scaling differences could also be driven in part by external influences related to
737 ecology, body size, and morphology. Body size affects many aspects of an animal’s ecology, diet
738 and energy consumption, and physiology (LaBarbera, 1986). It should be no surprise that this
739 may extend to brain composition. For example, the ancestor of extant primates, and most of its

740 descendants, occupied arboreal niches (Cartmill, 1972) and had arboreal locomotor strategies
741 that constrain body size and favor a low center of mass; a strategy that is likely inconsistent with
742 volumetrically expensive modes of brain expansion. Selection pressures that favored the
743 evolution of increased neuron number may therefore have been constrained by the physical
744 demands of occupying an arboreal niche, resulting in changes in neural development that were
745 associated with increased neuron density. Similar, but stronger, selection regimes may also
746 explain the extremely high neuron densities in bird brains (Olkowicz et al., 2016). Conversely,
747 the much lower neuron densities of cetaceans (Eriksen & Pakkenberg, 2007) would be consistent
748 with brain evolution along a trajectory relatively free of size or locomotor constraints.

749 The expectation that brain size should be a simple predictor of cognitive performance
750 ignores the effect of size-related selection pressures (Chittka & Niven, 2009; Chittka et al.,
751 2012). Size-efficiency is most obvious when considering brain function in small invertebrates,
752 but mounting evidence suggests that the same principles may apply even among vertebrates
753 occupying distinct ecological niches that define the range of permissible body sizes and
754 architectures (Olkowicz et al., 2016). Body size is regularly used as a ‘size-correction factor’ on
755 the assumption that residual brain size is more cognitively relevant, but variation in body size
756 itself reflects the presence of wider ecological and physical selection pressures that may render
757 brain composition and function more divergent than size alone (Fitzpatrick et al., 2012;
758 Montgomery et al., 2013, 2010).

759

760 **4.4 Correlations suffer from interference**

761 Problems of noise are compounded by interference from the complex relationships between
762 many behavioral and anatomical traits. This ‘interference’ not only influences our ability to

763 determine whether a mechanistic link exists between specific brain measures and a certain
764 behavior or cognitive ability, but also in determining their functional link and their adaptive
765 evolutionary history. The comparative study of different species can provide insights into how
766 differences in behavior link with differences in brains (Harvey & Pagel, 1991), and phylogenetic
767 comparisons have been the most widely used approach to test hypotheses about adaptation (see
768 §2.2). However, in addition to relying on unvalidated proxies, adaptive stories are frequently
769 based on correlations. It is therefore necessary to identify potential interference from unmeasured
770 variables to gather evidence for causation before we can accept such adaptive accounts as
771 accurate.

772 There are four main ways in which interference limits the potential to interpret whether
773 correlations represent adaptations. First, any association between differences in brain measures
774 and behavior might not be direct, but caused by interfering factors. For example, increases in
775 brain size and group size both appear to occur in species that eat foods with high nutritional
776 value, therefore the correlation between brain size and group size might be the result of noise
777 from dietary changes (Clutton-Brock & Harvey, 1980; DeCasien, Williams, & Higham, 2017).
778 Second, even if population studies indicate that a measure of brain size and a behavior are
779 directly linked, comparisons across species cannot immediately reveal the causal direction of the
780 association. For example, an association between increased brain size and decreased risk of
781 predation might result from large-brained species being better able to avoid predation (Kotrschal
782 et al., 2015), or from species with low predation pressure having the opportunity to invest
783 additional resources into brain growth (Walsh, Broyles, Beston, & Munch, 2016). Third, external
784 factors frequently mediate the expression of any link across taxonomic groups. For example,
785 switching to a frugivorous diet might lead to selection on olfactory ability in nocturnal species

786 and visual abilities in diurnal species, resulting in independent episodes of brain expansion
787 driven by selection on distinct sensory modalities and brain components (Barton, Purvis &
788 Harvey, 1995). Fourth, any current link between brain size and behavior might be the product of
789 co-option, after the initial evolution of that brain aspect, rather than the driving selection pressure
790 itself. For example, abilities such as object permanence (i.e., the ability to recall the presence of
791 an out-of-sight object) might have been selected because individuals need to remember the
792 spatial position and temporal availability of food sources in their home range, but it could
793 subsequently be used to distinguish social neighbors from strangers (Barton, 1998). Similarly,
794 selection for improved visual acuity in foraging primates may have later been co-opted to serve
795 in individual recognition and social cognition (Barton, 1998). Although some attempts have been
796 made to tease apart these relationships using path analysis (Dunbar & Shultz, 2007b), this
797 approach still suffers from the effects of co-linearity among variables and does not provide a
798 mechanistic understanding of causative relationships (Petraitis, Dunham, & Niewiarowski,
799 1996). Recent advances provide some ways to overcome these limitations in the comparative
800 approach (see §5.4), but, as previous authors have pointed out (Garland, Bennett, & Rezende,
801 2005; Gonzalez-Voyer & Hardenberg, 2014; Harvey & Pagel, 1991), interference fundamentally
802 limits our ability to determine past evolutionary processes based on simple observations of
803 species alive today.

804 These effects are likely to be particularly influential in the small data sets that
805 characterize many comparative analyses of cognition and brain measures, due to the difficulty in
806 obtaining data. With small data sets, correlations are unlikely to be stable, unless the effect size is
807 large, or noise and interference are low (Schönbrodt & Perugini, 2013). In the vast majority of
808 studies, accuracy and sample size are directly traded-off against one another due to logistical and

809 cost constraints. While this is inevitable, studies aiming for broad phylogenetic comparisons by
810 relying on crude proxies of cognition supposedly measurable across very divergent taxonomic
811 groups may be futile. Any trade-off that reduces accuracy to increase taxonomic breadth risks
812 relying on invalid measures, resulting in unstable and potentially meaningless correlations.
813 Comparisons across large, diverse taxonomic groups can be helpful to identify and describe
814 patterns of variation; however, key insights into the evolutionary history of traits and their
815 associations will be gained by incorporating detailed population studies (see §5.2). As
816 neuronanatomical, behavioral, and statistical tools become increasingly comprehensive and
817 sophisticated, the solutions to these issues will be reachable in the near future.

818 **5) Beyond brain size**

819 **5.1 Matching the right tool with the right question**

820 In the last two sections we emphasized how heterogeneity in brain composition and
821 behavior/cognition, and the subsequent noise this generates can influence our attempts to
822 measure the relationships between these variables. We think these issues motivate turning from
823 coarse-grained, ‘taxon-neutral’ (or hominid-inspired) measures to more local, taxon-specific
824 studies. This is not to say that heterogeneity on its own undermines existing ‘monolithic’
825 narratives of brain size and behavioral complexity. Rather, these narratives ignore the complexity
826 of links between brain morphology, body morphology, and behavior, and often abstract away
827 from the important ecological and evolutionary drivers of complex behavior that we are trying to
828 understand. We therefore argue against privileging anthropocentric measures or criteria. Instead,
829 we urge a recognition of the multi-dimensional and multi-leveled structure of brains, as well as
830 the disparate and varied ways that brains evolve—in conjunction with bodies, and in response to
831 specific environments—to produce complex behavior. Understanding how brains evolve in

832 response to selection on behavioral complexity or cognition is therefore a two-step process. First,
833 we must understand how behavioral variation emerges from variation in neural systems. Second,
834 we must understand how brains change across species and how this might relate to differences in
835 adaptive regimes.

836 Discovering and probing correlations between properties of brains and behavioral
837 features can be part of a powerful comparative approach, but we should be wary of reification:
838 mistaking an operationalized target of measurement with a ‘real’ object (Whitehead, 1925).
839 There is a difference between something being measurable and it being causally meaningful. We,
840 and others (e.g., Chittka et al., 2012; Healy & Rowe, 2007), have questioned whether coarse-
841 grained, cross-taxa measurements, such as the encephalization quotient, pick out relations that
842 are in fact explanatory of the evolutionary and developmental relationships between brain,
843 cognition, and behavior across lineages. In fact, similar arguments have been made since
844 scientists first started comparing brain measures across species (Snell, 1892). Instead, we argue
845 for an increased focus on a ‘bottom-up’ approach that begins with i) measurements of features
846 that can be validated within particular taxa in *ecologically relevant* experimental contexts, before
847 ii) testing the evolutionary variability in the relationships between brains and behavior across
848 related species. This will help avoid reification by starting with intraspecific, experimentally
849 verifiable causal connections. The first task involves probing how various taxa respond
850 behaviorally to their environments and other stimuli and determining whether those properties
851 correlate with brain measures in revealing ways. These brain measures will frequently be more
852 fine-grained than brain size, concerning particular neuroanatomical and/or neurophysiological
853 features. The second task involves the construction and testing of hypotheses about the ancestral
854 and evolutionary relationships between those taxa, enabling us to expand to broader categories

855 and correlations in a careful, piecemeal fashion. We expect the results of these two tasks to relate
856 in dynamic ways: considerations of evolutionary scenarios are likely to highlight new kinds of
857 experimental tests and hypotheses in local contexts; and these scenarios will depend crucially on
858 information about local taxa.

859

860 **5.2 Bottom-up vs. top-down**

861 The top-down approach uses cross-species correlations between brain measures and a trait of
862 interest and can be useful for generating hypotheses. However, while these are important for
863 motivating research into the links between brains and behavior, we argue that specific hypotheses
864 should then be tested at the within-species level: from the bottom up. The bottom-up approach
865 involves directly testing behavior and cognition in individuals to determine how they relate to
866 brain measures in these particular individuals of a particular species (ideally measured at the
867 same time as behavior/cognition) to build validated, causative correlations (Chittka et al., 2012).
868 When sufficient data on individuals from a variety of species have accumulated, phylogenetic
869 meta-analyses can be conducted to test whether consistent patterns emerge and hold within and
870 across species (see §5.4; Table 3). Correlations within contemporary populations can tell us
871 whether processes are homologous or analogous across species and show the limits of which
872 processes are likely to occur.

873 The contrast between top-down and bottom-up approaches is often presented as a
874 difference in terms of investigating the ultimate (top-down looking at adaptations and fitness)
875 versus proximate (bottom-up looking at mechanisms and development) reasons for the evolution
876 of a trait (Laland, Sterelny, Odling-Smee, Hoppitt, & Uller, 2011; Scott-Phillips, Dickins, &
877 West, 2011). However, the approach we suggest does not necessarily make this potentially

878 problematic distinction (Beatty, 1994; Calcott, 2013; Cauchoix & Chaine, 2016; Laland, Odling-
879 Smee, Hoppitt, & Uller, 2013). Our main argument for a bottom-up approach is to encourage
880 researchers to have a clear understanding of what they are investigating rather than to rely on
881 proxies. Detailed individual-based studies can reveal not only which brain measures are involved
882 in a particular cognitive ability or behavior, but also provide important insights into the
883 ecological correlates and fitness consequences of variation in particular brain measures (Table 3).
884 Further, starting from behaviors in particular species makes ensuring ecological, developmental,
885 and evolutionary relevance significantly more straightforward: it is a strategy for both avoiding
886 reification, and for being sensitive to the heterogeneity of both brains and behavior.

887
888 **Table 3:** Examples of how behavior (directly tested) links with brain measures at the within-
889 species level. These are the kinds of data that can contribute to the bottom-up approach to
890 generate hypotheses based on validated data.

Taxa	Description
Black-capped chickadee (<i>Poecile atricapillus</i>)	Birds in harsher environments (higher latitudes) had larger hippocampus volumes with more neurons (T. C. Roth & Pravosudov, 2009) and more neurogenesis (Chancellor, Roth, LaDage, & Pravosudov, 2011), were more efficient at recovering caches (spatial memory) and better at an associative learning task than conspecifics from less harsh environments (Pravosudov & Clayton, 2002).
Mountain chickadee (<i>Poecile gambeli</i>)	Individuals living at higher elevations had better spatial memory and more hippocampal neurons . Higher elevation environments are more challenging because variables such as day length and temperature vary more annually than they do at lower elevations (Freas, LaDage, Roth II, & Pravosudov, 2012).
Gambel's white-crowned sparrow (<i>Zonotrichia leucophrys gambelii</i>)	Neurogenesis increases in the song control nucleus HVC just prior to the breeding season (Larson et al., 2013). The breeding season is correlated with a higher song quality than in the non-breeding season (Meitzen, Thompson, Choi, Perkel, & Brenowitz, 2009; Tramontin &

Brenowitz, 2000). After the breeding season, as neurons die in the song control center the song structure degrades accordingly (Larson, Thatra, Lee, & Brenowitz, 2014).

Wolf spider
(*Hogna carolinensis*)

Spiderlings that remain in sibling groups with their mother have **larger protocerebrums**, **capture prey more efficiently**, and have better **spatial memory** than spiderlings raised in isolation (Punzo & Ludwig, 2002). Note: there were no differences in the number of neurons between conditions.

Honey bee
(*Apis mellifera*)

Bees with **larger total brain sizes** (due only to an increase in mushroom body calyx size) were better able to **learn** and **remember** to associate a scent with proboscis extensions (Gronenberg & Couvillon, 2010). Mushroom bodies are involved in learning and memory.

Common octopus
(*Octopus vulgaris*)

Octopus and cuttlefish have the **highest brain-to-body mass ratios** of all invertebrates and the ratios exceed that of most fish and reptiles (Packard, 1972). Lesions of the octopus **associative centers**, vertical lobes impair tactile and visual learning and memory (Hochner, Shomrat, & Fiorito, 2006).

891

892 For example, spatial navigation behavior has been directly linked to the hippocampus
893 using the bottom-up approach. Supporting evidence comes from intraspecies behavioral studies
894 in birds with hippocampal lesions, which indicates the causal relationship between location
895 memory and the hippocampus (Hampton & Shettleworth, 1996; Patel, Clayton, & Krebs, 1997).
896 Additionally, ecological correlates were found in black-capped chickadees where individuals
897 living in harsher environments (higher latitudes) were more efficient at recovering caches
898 (spatial memory) and had larger hippocampal volumes with higher neuron densities and more
899 neurogenesis than individuals at lower latitudes (Chancellor et al., 2011; Pravosudov & Clayton,
900 2002; T. C. Roth & Pravosudov, 2009). Further, real-time brain activity has been paired with
901 navigational behavior in rats; when navigating through a maze, particular neurons (place cells)
902 fired at particular locations in the hippocampus (Gupta, van der Meer, Touretzky, & Redish,
903 2010). Later, when the rats were not in the maze, rats mentally ‘ran’ through the maze and even

904 invented novel routes as evidenced by the sequences of the firing of their place cells (Gupta et
905 al., 2010). Place cell research and experimental designs that behaviorally test episodic memory
906 (e.g., Clayton & Dickinson, 1998) provide evidence for brain-behavior causations from the
907 bottom-up.

908 Where functional assays are either unfeasible or unethical, causality can be determined
909 using a quantitative genetics approach to model how multiple measured traits are related to one
910 another. Analyzing brain and behavioral data in pedigrees or full-sibling/half-sibling families
911 allows the estimation of genetic correlations between traits (i.e., demonstrating variation in two
912 traits that share a common genetic basis). If variation in brain size or composition causatively
913 produces variation in behavior we should expect strong genetic correlations between these traits.
914 This approach can be used not only to test brain-behavior relationships (e.g., Kotrschal et al.,
915 2014), but also to help resolve debates about, for example, the relative roles of domain general
916 and domain specific cognition (e.g., Pedersen, Plomin, Nesselroade, & McClearn, 1992), and
917 developmental models of brain evolution (e.g., Hager, Lu, Rosen, & Williams, 2012; Noreikiene
918 et al., 2015).

919

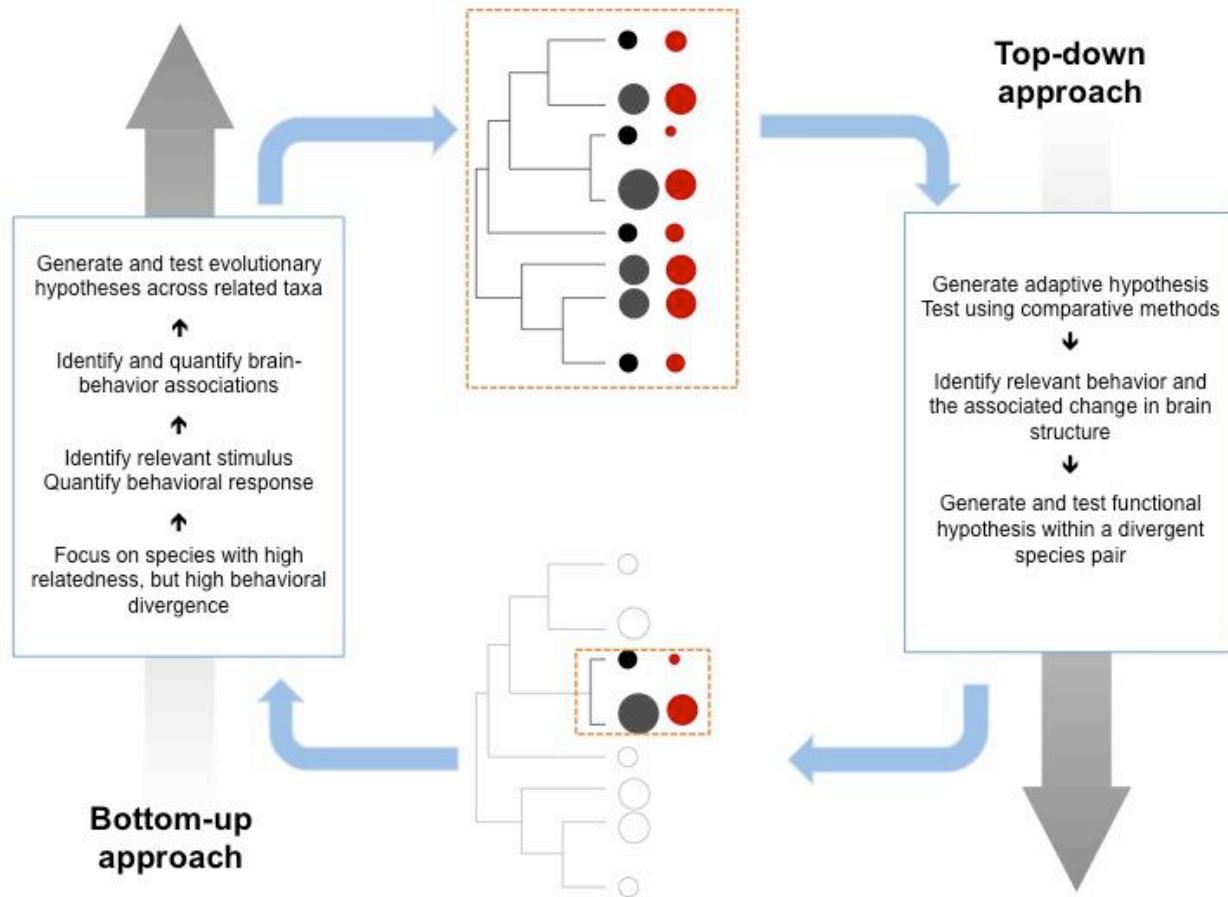
920 **5.3 The comparative approach as a tool for generating hypotheses and testing generality**

921 Although we argue for increased emphasis on intraspecific studies to validate causative
922 relationships, the comparative approach will remain an integral part of investigations of the
923 evolution of brains and cognitive abilities, though their scope and design might change.
924 Phylogenetic studies extend and inform detailed intraspecific studies, ideally leading to constant
925 feedback that can enhance both (Figure 3). Continuously developing comparative approaches
926 have the potential to reduce noise from small sample sizes, reveal relationships among multiple

927 interfering traits, and indicate the directionality of a causal association—though not all at once.
928 Combining findings from multiple populations can inform mechanistic studies by illustrating the
929 range of possible solutions that might exist, indicating where natural experiments might have
930 shaped evolution in similar ways, revealing potential mediators by indicating in which
931 taxonomic groups established relationships break down, and showing which species to target for
932 further study. In particular, the systematic combination of effect sizes from population studies in
933 phylogenetic meta-analyses reduces noise and can test the robustness of an association between
934 brain measures and behavior while also revealing potential mediators that systematically change
935 the form of the association in some populations or species (Nakagawa & Santos, 2012). For
936 example, they might reveal whether the heritability of brain measures might depend on
937 environmental variability.

938 In turn, the historical component of phylogenetic reconstructions extends population-
939 level studies by revealing whether detected patterns are evolutionarily stable or lineage-specific,
940 and they can contribute to determining causal or adaptive relationships between traits by
941 revealing temporal contingencies (Beaulieu, Jhwueng, Boettiger, & O’Meara, 2012; Pagel, 1999;
942 Pagel & Meade, 2006) in whether a behavior consistently changed prior to or after associated
943 changes in brain measures. The historical component provided by phylogenetic comparisons is
944 necessary to determine whether traits not only occur together, but whether they evolved together.
945 For example, while the enlarged brains (compared to most other reptiles) among birds appear to
946 be linked to cognitive capacities required for flight (Balanoff, Bever, Rowe, & Norell, 2013),
947 evolutionary origins of flying behavior are not associated with particular increases in endocranial
948 volume (Balanoff, Smaers, & Turner, 2016).

949



950

951 **Figure 3: Integrating the top-down and bottom-up approaches.**

952

953 Our discussion of the power of the comparative approach in elucidating the adaptive
954 history of traits indicates the inherent limits in fully explaining traits that supposedly make any
955 species unique. The evolutionary processes themselves are not unique, but the particular
956 combination of processes at play are. As such, understanding how such processes come together
957 in a particular instance is problematic due to a lack of evidence required to confirm these
958 hypotheses (Tucker, 1998). In addition, studies that focus on extraordinary traits in a single
959 species (such as humans) frequently risk misrepresenting evolutionary processes by fixating on
960 the endpoint as an optimal solution, whereas evolution typically progresses by responding to
961 stochastic variation in selection regimes, incrementally adapting to the environment.

962 **5.4 Scaling across taxa to integrate evidence**

963 The bottom-up approach we suggest means that scaling across taxa will initially be more difficult
964 to achieve because studies will have to be designed to take into account the characteristics of the
965 particular species, as well as its phylogenetic and ecological context. Questions, approaches, and
966 methods might need time to converge or to be repeated across a relevant sample of different taxa
967 (Figure 3). However, over an intermediate time frame, the bottom-up approach will be invaluable
968 for comparing and elucidating brain and cognitive evolution across taxa. Although the bottom-up
969 approach initially makes scaling look difficult, we think it has two advantages. First, rather than
970 positing or assuming a coarse-grained, cross-taxa category and applying it across a range of
971 cases (thus losing ecological relevance and increasing the potential for *post-hoc* explanations and
972 reification), the bottom-up approach makes scaling a much more piecemeal, empirically tractable
973 matter. Second, it more easily allows scaling to take place in an evolutionary context.
974 Understanding whether the same genes, genetic pathways, neural regions, neural physiology,
975 and/or neural networks are involved in generating cognitive abilities across taxa will provide us
976 with an understanding of how evolution has shaped the diversity of brains and the behavior they
977 produce. In this sense, phenotypic heterogeneity and taxonomic diversity become a tool for
978 discovery, rather than a source of statistical noise.

979 It is not straightforward to bring together the disparate evidence involved in shifting from
980 local experimental contexts to cross-taxa, evolutionary hypotheses. However, a detailed
981 understanding of the mechanisms underlying brain measures and behavior is crucial to clarify
982 whether traits are homologous, analogous, or completely independent solutions to ecological
983 challenges. To give a sense of the possibilities for integration, we sketch three kinds of

984 approaches to shifting from local (bottom-up) to general (top-down) scales (from Currie, 2013;
985 see also Mikhalevich et al., 2017):

- 986 1. Detecting homologous relationships, where the same brain measures and behavior are
987 related to the same environment across species descended from a recent common
988 ancestor (Currie, 2012), offers opportunities to combine independent findings into one
989 mechanistic pathway. In these instances, inheritance and stabilizing selection have
990 maintained a stable trait, such that findings from one species can be accurately inferred
991 for another. Such investigations will rely on integrated models that bring together
992 disparate evidence to support hypotheses about the evolutionary, developmental, and
993 ecological features of a particular lineage.
- 994 2. Determining whether the same behavior occurs in similar environments across distantly
995 related species can indicate environments most likely relevant for the emergence of the
996 behavior. A bottom-up approach can reveal whether the observed behavior represents
997 analogous re-emergence of a behavior within the same adaptive environment (e.g.,
998 repeated evolution of feathers across dinosaurs; B. K. Hall, 2003; McGhee, 2011). This
999 approach will rely on parallel models that identify brain-behavior correlations within
1000 related taxa for which the main principles of brain evolution are known to be similar. As
1001 closely related taxa will likely share meaningful brain-behavior correlations, such models
1002 are likely to be well-validated, stable, and causally meaningful.
- 1003 3. Observing a similar behavior in similar adaptive environments can reveal whether the
1004 behavior represents a convergent solution to the same environment (e.g., feathered wings
1005 for flight versus bat wings) or whether the relationship is more complex (e.g., wings to
1006 escape into the air versus jumping legs; Currie, 2014; Pearce, 2012; Powell, 2007). This

1007 type of convergent model is similar to the top-down approach; however, convergent
1008 modeling avoids many of the cross-taxa comparison problems by i) being placed in an
1009 explicitly ecological and phylogenetic context, ii) being carried out alongside parallel and
1010 integrated modeling, and iii) avoiding over-interpretation that arises from defining
1011 categories based on superficial similarities because convergent models are inherently
1012 sensitive to the explanatory limits of analogous categories (see Griffiths, 1994 for a
1013 discussion of these limits).

1014 There are a wide variety of different scales at which we may need to infer evolutionary
1015 relationships between brain measures, behavior, and environments across taxa, and the
1016 ecological and evolutionary relevance granted by starting in local contexts is crucial for doing
1017 this.

1018 **Conclusion**

1019 We support a two-pronged strategy for understanding cross-taxa relationships between brain size,
1020 brain composition, behavior, and cognition that focuses on ecologically relevant contexts rather
1021 than attempting broad scale comparisons at gross phenotypic levels. The first prong is an
1022 experimental program examining correlations in closely related species; the second prong
1023 involves the piecemeal identification of correlations at broader taxonomic scales. We have
1024 contrasted our approach with one that has become dominant in recent years. The alternative
1025 approach relies on coarse-grained phenotypes and proxy-measures, typically in anthropocentric
1026 contexts, and attempts to apply these to cross-taxa, correlative contexts. We have highlighted a
1027 number of limitations to this approach. First, applying anthropocentric conceptions of brain
1028 correlates with behavior to disparate taxa comes at the crucial cost of ecological and evolutionary
1029 coherence. Second, the heterogeneity of brain composition and behavior makes coarse-grained

1030 conceptions problematic because cross-taxa comparisons inevitably discount variation that
1031 matters for particular lineages. This variation creates noise in statistical comparisons.
1032 Heterogeneity can also be a source of interference because various interdependencies both
1033 between brain structures (e.g., in development or function) and between multiple behavioral and
1034 ecological traits undermine our capacity to identify selection pressures shaping individual traits
1035 or systems. Third, beginning with ‘general’ measures of intelligence potentially leads to
1036 reification and the establishment of misguided or causally meaningless properties. The top-down
1037 approach has not necessarily been misguided itself: scientific progress is often facilitated by
1038 applying relatively crude measures, highlighting the value of using many investigative
1039 techniques. Indeed, the heterogeneity of these traits have become known to us *because* the top-
1040 down approach has exposed inconsistencies through cross-species correlations. However, it is
1041 time to take the cognitive, behavioral, and brain features of particular lineages seriously, rather
1042 than demand that they be shoehorned into anthropocentric notions, or judged against some
1043 general metric. In doing so, a more general understanding of the nature of cognition and
1044 behavior, and their relationship with brain measures will be built from the bottom up.

1045

1046 **ACKNOWLEDGEMENTS**

1047 We are grateful for manuscript feedback from Nicky Clayton, Tim Clutton-Brock, and Rob
1048 Barton, and we appreciate Christian Rutz for discussions about New Caledonian crow foraging
1049 innovations. We thank our funders: the Isaac Newton Trust and Leverhulme Trust for a
1050 Leverhulme Early Career Fellowship to CJL, which funded the workshop on which this paper is
1051 based; NERC for an Independent Research Fellowship to SHM; the European Research Council
1052 (Grant No. 3399933; SAJ); the Royal Society for a Dorothy Hodgkin Research Fellowship to

1053 NJB; the Royal Society of New Zealand Marsden Fund (UOC1301; FRC); the National Science
1054 Foundation (NSF BCS 1440755; RM); the John Templeton Foundation (AB); and the Templeton
1055 World Charity Foundation (AC; note: the opinions expressed in this publication are those of the
1056 author(s) and do not necessarily reflect the views of Templeton World Charity Foundation).

1057 **AUTHORSHIP CONTRIBUTIONS**

1058 All authors contributed to the concepts in this paper at a workshop organized by Logan in March
1059 2017. Montgomery structured the paper. All authors wrote and edited the paper. Logan,
1060 Montgomery, Boogert, and Mares served as managing editors. All authors approved the final
1061 version for submission. We certify that what we have written represents original scholarship.

1062 **BOX 1. What can comparative cognitive tests tell us?**

1063 Performance on cognitive tests arguably offers the most direct measure of ‘*intelligence*’ and can
1064 therefore be seen as an improvement over other behavioral measures – such as innovation rate –
1065 which have been used as proxies for intelligence. However, cognitive abilities can still only be
1066 *inferred* from test performance, rather than directly measured. Performance on any given test will
1067 depend on a suite of abilities and behaviors (such as motor abilities, perception, attention,
1068 motivation, fear), beyond the cognitive ability in question, which mean that successes or failures
1069 can occur for a range of different reasons across subjects (Rowe & Healy, 2014; A. Seed,
1070 Seddon, Greene, & Call, 2012).

1071 Let’s take the example of MacLean and colleagues’ (2014) collaborative study,
1072 comparing brain size with measures of self-control across species. Here, 36 species of mammals
1073 and birds received two cognitive tasks: the A-not-B task and the cylinder task. In the A-not-B
1074 task a human demonstrator places food in cup A multiple times. Once the subject has retrieved a
1075 reward from cup A three times in a row, they are given a test trial where food is first placed in
1076 cup A, and then visibly moved to cup B. Subjects have to inhibit choosing previously rewarded
1077 cup A to succeed. This is a commonly used developmental task, on which babies under 10
1078 months typically display perseverative reaching errors (Smith, Thelen, Titzer, & McLin, 1999).
1079 However, in addition to ‘self-control’, to pass this task, subjects also need to be capable of
1080 accurately tracking the movement of food by human hands. This is trivial for humans, and may
1081 also be relatively easy for great apes or other primates, given that they are closely related to
1082 humans and also possess hands. However, there is growing evidence that many species struggle
1083 to use this type of information from a human demonstrator (e.g., Erdőhegyi, Topál, Virányi, &
1084 Miklósi, 2007; Shaw, Plotnik, & Clayton, 2013); thus, this task may have systematically

1085 disadvantaged certain species for reasons unrelated to their self-control (Jelbert et al., 2016).
1086 Crucially, in MacLean and colleagues' study, although some animals (lemurs, dogs and pigeons)
1087 were explicitly trained to select experimenter-baited containers during pre-training, most other
1088 species were not. All non-primates that were given both tests, and had *not* been trained to attend
1089 to the demonstrator, performed substantially worse on the A-not-B task than on the cylinder task.
1090 Elephants, given the A-not-B task only, failed every trial. Without knowing whether subjects
1091 possessed one of the key requirements for the A-not-B test, we cannot know whether their poor
1092 test scores actually reflect poor self-control. In a study that directly tested for this effect, New
1093 Caledonian crows that had been explicitly trained to attend to a human demonstrator went on to
1094 pass 67% of A-not-B test trials, while a control group, trained on an unrelated inhibitory control
1095 task, passed only 7% (Jelbert et al., 2016).

1096 The particular use of the A-not-B task is a clear example of a situation in which
1097 comparing cognitive test scores from different species, and relating them to brain size, is unlikely
1098 to provide us with a meaningful comparison of the cognitive ability in question. Given the range
1099 of factors that can influence test success, the majority of cognitive tests will suffer from
1100 limitations like this to some degree. For example, in MacLean's cylinder task, one key variable
1101 that would influence performance was the amount of experience different subjects had with
1102 transparent materials. In the cylinder task, subjects first learn to retrieve a reward from the open
1103 end of an opaque cylinder, and are then presented with a transparent cylinder, where the reward
1104 is now visible through the tube. Subjects are considered to have failed the task if they touch or
1105 peck the front of the transparent tube, rather than detouring to the open side. Kabadayi et al.
1106 (2016) highlighted that a number of bird species showed learning effects over the course of 10

1107 trials in the cylinder task, suggesting that relative unfamiliarity with transparent Perspex might
1108 have contributed towards this behavior.

1109 Other factors that might have influenced performance can also be hypothesized. These
1110 could include morphological differences (i.e., performance in the cylinder task may vary
1111 depending on whether a subject must insert their arm or their head into the tube to obtain the
1112 reward), perceptual differences (i.e., how well can different subjects perceive the transparent
1113 material?), behavioral differences (i.e., does this species typically explore or avoid new objects?)
1114 or motivational differences (i.e., to what extent is the animal motivated to directly obtain the
1115 reward?) or any other task specific variables. Thus, while here we have highlighted some specific
1116 examples, performance on any cognitive test will be influenced by numerous sources of
1117 variation, both across individuals and across species, in addition to the cognitive ability in
1118 question.

1119 To address this, minimizing any unnecessary task demands (such as the use of human
1120 demonstrators) is the first step in designing comparative tests. It is also crucial to level the
1121 playing field from the bottom up, including baseline criteria training to ensure that all subjects
1122 meet specific key requirements, before the test of the desired cognitive ability begins. Focusing
1123 on groups of more closely related species will also help to limit the number of ways in which
1124 subjects' performance could vary beyond the ability in question.

1125 **BOX 2. Spider behavior varies according to environmental differences**

1126 Traditionally, vertebrates have been used as subjects to help answer probing questions relating to
1127 animal cognition, including studies of working memory, search images, expectancy violation and
1128 insight (e.g., Kamil & Bond, 2006; Köhler, 1924, p. 192; Pepperberg & Kozak, 1986;
1129 Shettleworth, 2010). However, recent research demonstrates that invertebrates, despite having
1130 much smaller brains, perform similarly on cognitive tasks (Jakob, Skow, & Long, 2011; Perry,
1131 Barron, & Chittka, 2017), thereby challenging notions of brain size translating into cognitive
1132 ability.

1133 Jumping spiders (family Salticidae), for example, have been of considerable interest in
1134 the context of cognition, partly because they have large forward-facing principal eyes that play a
1135 major role in high-precision visual discrimination (Harland, Li, & Jackson, 2012), supporting
1136 tasks such as selective attention and planning (Jackson & Cross, 2011). There is still much to be
1137 learned about spider neurobiology (but see Menda, Shamble, Nitzany, Golden, & Hoy, 2014);
1138 however, excellent vision may be part of the solution for how an animal with a minute brain can
1139 perform cognitive tasks in its environment (e.g., Pfeifer & Iida, 2005).

1140 Moreover, spiders are an excellent system for investigating how variation in cognitive
1141 abilities may be associated with ecology. The salticid genus *Portia* has provided us with many
1142 insights because the species from this genus eat other spiders (Jackson & Wilcox, 1998) and
1143 deploy a variety of strategies to avoid being eaten by their prey (Jackson & Cross, 2011). For
1144 example, when at the edge of another spider's web, *Portia* is known to deploy a specialized
1145 strategy of moving its eight legs and two pedipalps across the silk in ways that may mimic the
1146 movements made by a trapped insect (Jackson & Cross, 2013). These web signals are generated
1147 using trial and error (Jackson & Wilcox, 1993); *Portia* repeats a signal when the resident spider

1148 starts moving toward *Portia* across the web, and changes to using a different signal when it does
1149 not succeed at eliciting an approach by the resident spider.

1150 However, differences in the use of this tactic have been observed in different populations
1151 of the same species. In the Philippines, two populations (Los Baños and Sagada) of *Portia*
1152 *orientalis* (formerly *P. labiata*) encounter different spider species as prey. The site at Los Baños
1153 is a low-elevation tropical rainforest where *P. orientalis* encounters a wider variety of prey
1154 species, including species that are particularly dangerous, such as *Scytodes pallidus*, a spitting
1155 spider that specializes on salticids as prey (Li, Jackson, & Barrion, 1999). In contrast, the site at
1156 Sagada is a high-elevation pine-forest where *P. orientalis* encounters a smaller number of prey
1157 species, and does not encounter the particularly dangerous prey species that are present in Los
1158 Baños (Jackson & Carter, 2001; Jackson, Pollard, Li, & Fijn, 2002). For the Los Baños *P.*
1159 *orientalis*, a greater reliance on flexible predatory strategies, such as using trial and error when
1160 generating signals in other spiders' webs, is likely to be of greater importance than for the Sagada
1161 population. Indeed, when at the edge of another spider's web, the *P. orientalis* from Los Baños
1162 repeated a signal that elicited movement by the resident spider significantly more often than the
1163 Sagada population, and were also more likely to change a signal when it did not elicit movement
1164 by the resident spider (Jackson & Carter, 2001).

1165 Individuals from the Los Baños population also learned faster when faced with a novel
1166 situation of escaping from an island in a water-filled tray (Jackson, Cross, & Carter, 2006). After
1167 leaving the island, the spider first needed to reach an atoll before it could then reach the edge of
1168 the tray, but the distance was too far for the spider to clear by leaping alone. Instead, the spider
1169 could first reach the atoll by swimming across or, alternatively, it could leap before swimming
1170 the rest of the way. However, before a trial began, the researchers decided at random which of

1171 these two tactics (leaping or swimming) was the ‘successful’ tactic for reaching the atoll. If the
1172 spider used the successful tactic (e.g., when the successful tactic was leaping and the spider
1173 leaped), the researchers used a plastic scoop to make small waves in the water to help the spider
1174 across to the atoll. The spider then made its next move from the atoll. If, however, the spider
1175 used the ‘unsuccessful’ tactic (e.g., when the successful tactic was swimming and the spider
1176 leapt), the researchers used the plastic scoop to force it back to the island. To successfully reach
1177 the atoll, the spider then needed to switch to using the other tactic on its next attempt from the
1178 island. Similar to when making signals in webs, individuals of the Los Baños population
1179 repeated tactics when successful, and switched tactics when unsuccessful, significantly more
1180 often than the Sagada individuals (Jackson et al., 2006).

1181 It is currently unknown whether such observed differences in arachnid behavior are
1182 causally related to neural architecture. However, applying the bottom-up approach to the study of
1183 cognition in this taxon will likely contribute substantially to our understanding of how cognition
1184 relates to ecology and neurobiology. As well as showing a wide variety of foraging strategies in
1185 various environments, spiders are relatively easy to study in the field and laboratory, especially
1186 when compared with large vertebrates.

1187 **References**

- 1188 Ahn, Y.-Y., Jeong, H., & Kim, B. J. (2006). Wiring cost in the organization of a biological
1189 neuronal network. *Physica A: Statistical Mechanics and Its Applications*, 367, 531–537.
1190 <https://doi.org/10.1016/j.physa.2005.12.013>
- 1191 Aiello, L. C., & Wheeler, P. (1995). The expensive-tissue hypothesis: The brain and the digestive
1192 system in human and primate evolution. *Current Anthropology*, 36(2), 199–221.
1193 <https://doi.org/10.1086/204350>
- 1194 Allman, J., McLaughlin, T., & Hakeem, A. (1993). Brain weight and life-span in primate species.
1195 *Proceedings of the National Academy of Sciences*, 90(1), 118–122.
- 1196 Allritz, M., Tennie, C., & Call, J. (2013). Food washing and placer mining in captive great apes.
1197 *Primates*, 54(4), 361–370. <https://doi.org/10.1007/s10329-013-0355-5>
- 1198 Andics, A., Gábor, A., Gácsi, M., Faragó, T., Szabó, D., & Miklósi, Á. (2016). Neural
1199 mechanisms for lexical processing in dogs. *Science*, 353(6303), 1030–1032.
1200 <https://doi.org/10.1126/science.aaf3777>
- 1201 Armstrong, E. (1983). Relative brain size and metabolism in mammals. *Science*, 220(4603),
1202 1302–1304. <https://doi.org/10.1126/science.6407108>
- 1203 Atchley, W. R. (1984). Ontogeny, timing of development, and genetic variance-covariances
1204 structure. *The American Naturalist*, 123(4), 519–540. <https://doi.org/10.1086/284220>
- 1205 Auersperg, A. M. I., Bayern, A. M. P. von, Gajdon, G. K., Huber, L., & Kacelnik, A. (2011).
1206 Flexibility in problem solving and tool use of kea and New Caledonian crows in a multi
1207 access box paradigm. *PLOS ONE*, 6(6), e20231.
1208 <https://doi.org/10.1371/journal.pone.0020231>

- 1209 Balanoff, A. M., Bever, G. S., Rowe, T. B., & Norell, M. A. (2013). Evolutionary origins of the
1210 avian brain. *Nature*, *501*(7465), 93–96. <https://doi.org/10.1038/nature12424>
- 1211 Balanoff, A. M., Smaers, J. B., & Turner, A. H. (2016). Brain modularity across the theropod–
1212 bird transition: Testing the influence of flight on neuroanatomical variation. *Journal of*
1213 *Anatomy*, *229*(2), 204–214. <https://doi.org/10.1111/joa.12403>
- 1214 Barton, R. A. (1998). Visual specialization and brain evolution in primates. *Proceedings of the*
1215 *Royal Society of London B: Biological Sciences*, *265*(1409), 1933–1937.
1216 <https://doi.org/10.1098/rspb.1998.0523>
- 1217 Barton, R. A. (2012). Embodied cognitive evolution and the cerebellum. *Phil. Trans. R. Soc. B*,
1218 *367*(1599), 2097–2107. <https://doi.org/10.1098/rstb.2012.0112>
- 1219 Barton, R. A., & Capellini, I. (2011). Maternal investment, life histories, and the costs of brain
1220 growth in mammals. *Proceedings of the National Academy of Sciences*, *108*(15), 6169–
1221 6174. <https://doi.org/10.1073/pnas.1019140108>
- 1222 Barton, R. A., & Harvey, P. H. (2000). Mosaic evolution of brain structure in mammals. *Nature*,
1223 *405*(6790), 1055–1058. <https://doi.org/10.1038/35016580>
- 1224 Barton, R. A., Purvis, A., & Harvey, P. H. (1995). Evolutionary radiation of visual and olfactory
1225 brain systems in primates, bats and insectivores. *Philosophical Transactions of the Royal*
1226 *Society of London B: Biological Sciences*, *348*(1326), 381–392.
1227 <https://doi.org/10.1098/rstb.1995.0076>
- 1228 Barton, R. A., & Venditti, C. (2014). Rapid evolution of the cerebellum in humans and other
1229 great apes. *Current Biology*, *24*(20), 2440–2444. <https://doi.org/10.1016/j.cub.2014.08.056>
- 1230 Beatty, J. (1994). The proximate/ultimate distinction in the multiple careers of Ernst Mayr.
1231 *Biology and Philosophy*, *9*(3), 333–356. <https://doi.org/10.1007/BF00857940>

- 1232 Beaulieu, J. M., Jhwueng, D.-C., Boettiger, C., & O'Meara, B. C. (2012). Modeling stabilizing
1233 selection: Expanding the Ornstein–Uhlenbeck model of adaptive evolution. *Evolution*,
1234 *66*(8), 2369–2383. <https://doi.org/10.1111/j.1558-5646.2012.01619.x>
- 1235 Bendesky, A., Kwon, Y.-M., Lassance, J.-M., Lewarch, C. L., Yao, S., Peterson, B. K., ...
1236 Hoekstra, H. E. (2017). The genetic basis of parental care evolution in monogamous mice.
1237 *Nature*, *544*(7651), 434–439. <https://doi.org/10.1038/nature22074>
- 1238 Bengio, Y., Simard, P., & Frasconi, P. (1994). Learning long-term dependencies with gradient
1239 descent is difficult. *IEEE Transactions on Neural Networks*, *5*(2), 157–166.
1240 <https://doi.org/10.1109/72.279181>
- 1241 Benson-Amram, S., Dantzer, B., Stricker, G., Swanson, E. M., & Holekamp, K. E. (2016). Brain
1242 size predicts problem-solving ability in mammalian carnivores. *Proceedings of the*
1243 *National Academy of Sciences*, *113*(9), 2532–2537.
1244 <https://doi.org/10.1073/pnas.1505913113>
- 1245 Benson-Amram, S., & Holekamp, K. E. (2012). Innovative problem solving by wild spotted
1246 hyenas. *Proc. R. Soc. B*, *279*(1744), 4087–4095. <https://doi.org/10.1098/rspb.2012.1450>
- 1247 Bickart, K. C., Wright, C. I., Dautoff, R. J., Dickerson, B. C., & Barrett, L. F. (2011). Amygdala
1248 volume and social network size in humans. *Nature Neuroscience*, *14*(2), 163–164.
1249 <https://doi.org/10.1038/nn.2724>
- 1250 Bogart, S. L., Bennett, A. J., Schapiro, S. J., Reamer, L. A., & Hopkins, W. D. (2014). Different
1251 early rearing experiences have long-term effects on cortical organization in captive
1252 chimpanzees (*Pan troglodytes*). *Developmental Science*, *17*(2), 161–174.
1253 <https://doi.org/10.1111/desc.12106>

- 1254 Bond, A. B., Kamil, A. C., & Balda, R. P. (2007). Serial reversal learning and the evolution of
1255 behavioral flexibility in three species of North American corvids (*Gymnorhinus*
1256 *cynocephalus*, *Nucifraga columbiana*, *Aphelocoma californica*). *Journal of Comparative*
1257 *Psychology*, *121*(4), 372–379. <https://doi.org/10.1037/0735-7036.121.4.372>
- 1258 Boogert, N. J., Anderson, R. C., Peters, S., Searcy, W. A., & Nowicki, S. (2011). Song repertoire
1259 size in male song sparrows correlates with detour reaching, but not with other cognitive
1260 measures. *Animal Behaviour*, *81*(6), 1209–1216.
1261 <https://doi.org/10.1016/j.anbehav.2011.03.004>
- 1262 Boogert, N. J., Monceau, K., & Lefebvre, L. (2010). A field test of behavioural flexibility in
1263 Zenaida doves (*Zenaida aurita*). *Behavioural Processes*, *85*(2), 135–141.
1264 <https://doi.org/10.1016/j.beproc.2010.06.020>
- 1265 Boogert, N. J., Reader, S. M., Hoppitt, W., & Laland, K. N. (2008). The origin and spread of
1266 innovations in starlings. *Animal Behaviour*, *75*(4), 1509–1518.
1267 <https://doi.org/10.1016/j.anbehav.2007.09.033>
- 1268 Brown, M. F., & Sayde, J. M. (2013). Same/different discrimination by bumblebee colonies.
1269 *Animal Cognition*, *16*(1), 117–125. <https://doi.org/10.1007/s10071-012-0557-z>
- 1270 Bullmore, E., & Sporns, O. (2012). The economy of brain network organization. *Nature Reviews*
1271 *Neuroscience*, *13*(5), 336–349. <https://doi.org/10.1038/nrn3214>
- 1272 Burkart, J. M., Schubiger, M. N., & van Schaik, C. P. (2016). The evolution of general
1273 intelligence. *Behavioral and Brain Sciences*, 1–65.
1274 <https://doi.org/10.1017/S0140525X16000959>

- 1275 Burns, J. G., Foucaud, J., & Mery, F. (2010). Costs of memory: Lessons from ‘mini’ brains.
1276 *Proceedings of the Royal Society of London B: Biological Sciences*, rspb20102488.
1277 <https://doi.org/10.1098/rspb.2010.2488>
- 1278 Buzsáki, G., Geisler, C., Henze, D. A., & Wang, X.-J. (2004). Interneuron Diversity series:
1279 Circuit complexity and axon wiring economy of cortical interneurons. *Trends in*
1280 *Neurosciences*, 27(4), 186–193. <https://doi.org/10.1016/j.tins.2004.02.007>
- 1281 Calcott, B. (2013). Why the Proximate-Ultimate Distinction Is Misleading, and Why It Matters
1282 for Understanding the Evolution of Cooperation. In K. Sterelny, R. Joyce, B. Calcott, & B.
1283 Fraser (Eds.), *Cooperation and Its Evolution* (pp. 249–264). Cambridge, USA: MIT Press.
- 1284 Cartmill, M. (1972). Arboreal adaptations and the origin of the Order Primates. *The Function*
1285 *and Evolutionary Biology of Primates*, 97–122.
- 1286 Cauchoix, M., & Chaine, A. S. (2016). How can we study the evolution of animal minds?
1287 *Frontiers in Psychology*, 7. <https://doi.org/10.3389/fpsyg.2016.00358>
- 1288 Chancellor, L. V., Roth, T. C., LaDage, L. D., & Pravosudov, V. V. (2011). The effect of
1289 environmental harshness on neurogenesis: A large-scale comparison. *Developmental*
1290 *Neurobiology*, 71(3), 246–252. <https://doi.org/10.1002/dneu.20847>
- 1291 Chen, B. L., Hall, D. H., & Chklovskii, D. B. (2006). Wiring optimization can relate neuronal
1292 structure and function. *Proceedings of the National Academy of Sciences of the United*
1293 *States of America*, 103(12), 4723–4728. <https://doi.org/10.1073/pnas.0506806103>
- 1294 Chittka, L., & Niven, J. (2009). Are bigger brains better? *Current Biology*, 19(21), R995–R1008.
1295 <https://doi.org/10.1016/j.cub.2009.08.023>

- 1296 Chittka, L., Rossiter, S. J., Skorupski, P., & Fernando, C. (2012). What is comparable in
1297 comparative cognition? *Phil. Trans. R. Soc. B*, 367(1603), 2677–2685.
1298 <https://doi.org/10.1098/rstb.2012.0215>
- 1299 Clancy, B., Darlington, R. B., & Finlay, B. L. (2001). Translating developmental time across
1300 mammalian species. *Neuroscience*, 105(1), 7–17. <https://doi.org/10.1016/S0306->
1301 [4522\(01\)00171-3](https://doi.org/10.1016/S0306-4522(01)00171-3)
- 1302 Clayton, N. S., & Dickinson, A. (1998). Episodic-like memory during cache recovery by scrub
1303 jays. *Nature*, 395(6699), 272–274. <https://doi.org/10.1038/26216>
- 1304 Clutton-Brock, T. H., & Harvey, P. H. (1980). Primates, brains and ecology. *Journal of Zoology*,
1305 190(3), 309–323. <https://doi.org/10.1111/j.1469-7998.1980.tb01430.x>
- 1306 Cnotka, J., Güntürkün, O., Rehkämper, G., Gray, R. D., & Hunt, G. R. (2008). Extraordinary
1307 large brains in tool-using New Caledonian crows (*Corvus moneduloides*). *Neuroscience*
1308 *Letters*, 433(3), 241–245. <https://doi.org/10.1016/j.neulet.2008.01.026>
- 1309 Cross, F. R., & Jackson, R. R. (2014). Specialised use of working memory by *Portia africana*, a
1310 spider-eating salticid. *Animal Cognition*, 17(2), 435–444. <https://doi.org/10.1007/s10071->
1311 [013-0675-2](https://doi.org/10.1007/s10071-013-0675-2)
- 1312 Cross, F. R., & Jackson, R. R. (2016). The execution of planned detours by spider-eating
1313 predators. *Journal of the Experimental Analysis of Behavior*, 105(1), 194–210.
1314 <https://doi.org/10.1002/jeab.189>
- 1315 Cross, F. R., & Jackson, R. R. (2017). Representation of different exact numbers of prey by a
1316 spider-eating predator. *Interface Focus*, 7(3), 20160035.
1317 <https://doi.org/10.1098/rsfs.2016.0035>

- 1318 Currie, A. (2012). Convergence, contingency & morphospace. *Biology & Philosophy*, 27(4),
1319 583–593. <https://doi.org/10.1007/s10539-012-9319-1>
- 1320 Currie, A. (2013). Convergence as evidence. *The British Journal for the Philosophy of Science*,
1321 64(4), 763–786. <https://doi.org/10.1093/bjps/axs027>
- 1322 Currie, A. (2014). Venomous dinosaurs and rear-fanged snakes: Homology and homoplasy
1323 characterized. *Erkenntnis*, 79(3), 701–727. <https://doi.org/10.1007/s10670-013-9533-5>
- 1324 Currie, A., & Walsh, P. D. (n.d.). *Ontic-driven explanation of scientific method*.
- 1325 Dacke, M., & Srinivasan, M. V. (2008). Evidence for counting in insects. *Animal Cognition*,
1326 11(4), 683–689. <https://doi.org/10.1007/s10071-008-0159-y>
- 1327 De Groof, G., Jonckers, E., Güntürkün, O., Denolf, P., Van Audekerke, J., & Van der Linden, A.
1328 (2013). Functional MRI and functional connectivity of the visual system of awake pigeons.
1329 *Behavioural Brain Research*, 239, 43–50. <https://doi.org/10.1016/j.bbr.2012.10.044>
- 1330 Deaner, R. O., Barton, R. A., & van Schaik, C. P. (2003). Primate Brains and Life Histories:
1331 Renewing the Connection. In P. Kappeler & M. Pereira (Eds.), *Primate Life Histories and*
1332 *Socioecology* (pp. 233–265). Chicago, USA: University of Chicago Press.
- 1333 Deaner, R. O., Isler, K., Burkart, J., & Schaik, C. van. (2007). Overall brain size, and not
1334 encephalization quotient, best predicts cognitive ability across non-human primates. *Brain,*
1335 *Behavior and Evolution*, 70(2), 115–124. <https://doi.org/10.1159/000102973>
- 1336 Deaner, R. O., van Schaik, C. P., & Johnson, V. (2006). Do some taxa have better domain-general
1337 cognition than others? A meta-analysis of nonhuman primate studies. *Evolutionary*
1338 *Psychology*, 4(1), 147470490600400130. <https://doi.org/10.1177/147470490600400114>

- 1339 DeCasien, A. R., Williams, S. A., & Higham, J. P. (2017). Primate brain size is predicted by diet
1340 but not sociality. *Nature Ecology & Evolution*, *1*, 0112. [https://doi.org/10.1038/s41559-](https://doi.org/10.1038/s41559-017-0112)
1341 017-0112
- 1342 Ducatez, S., Clavel, J., & Lefebvre, L. (2015). Ecological generalism and behavioural innovation
1343 in birds: Technical intelligence or the simple incorporation of new foods? *Journal of*
1344 *Animal Ecology*, *84*(1), 79–89. <https://doi.org/10.1111/1365-2656.12255>
- 1345 Dunbar, R. I. M., & Shultz, S. (2007a). Evolution in the social brain. *Science*, *317*(5843), 1344–
1346 1347. <https://doi.org/10.1126/science.1145463>
- 1347 Dunbar, R. I. M., & Shultz, S. (2007b). Understanding primate brain evolution. *Philosophical*
1348 *Transactions of the Royal Society of London B: Biological Sciences*, *362*(1480), 649–658.
1349 <https://doi.org/10.1098/rstb.2006.2001>
- 1350 Eberhard, W. G., & Weislo, W. T. (2011). Grade Changes in Brain-Body Allometry:
1351 Morphological and Behavioural Correlates of Brain Size in Miniature Spiders, Insects and
1352 Other Invertebrates. In J. Casas (Ed.), *Advances in Insect Physiology* (Vol. 40, pp. 155–
1353 214). London, UK: Academic Press.
- 1354 Emery, N. J., & Clayton, N. S. (2004). Comparing the Complex Cognition of Birds and Primates.
1355 In L. J. Rogers & G. Kaplan (Eds.), *Comparative Vertebrate Cognition* (pp. 3–55). USA:
1356 Springer. https://doi.org/10.1007/978-1-4419-8913-0_1
- 1357 Enard, W., Gehre, S., Hammerschmidt, K., Hölter, S. M., Blass, T., Somel, M., ... Pääbo, S.
1358 (2009). A humanized version of Foxp2 affects cortico-basal ganglia circuits in mice. *Cell*,
1359 *137*(5), 961–971. <https://doi.org/10.1016/j.cell.2009.03.041>

- 1360 Enard, W., Przeworski, M., Fisher, S. E., Lai, C. S. L., Wiebe, V., Kitano, T., ... Pääbo, S. (2002).
1361 Molecular evolution of *FOXP2*, a gene involved in speech and language. *Nature*,
1362 *418*(6900), 869–872. <https://doi.org/10.1038/nature01025>
- 1363 Erdőhegyi, Á., Topál, J., Virányi, Z., & Miklósi, Á. (2007). Dog-logic: Inferential reasoning in a
1364 two-way choice task and its restricted use. *Animal Behaviour*, *74*(4), 725–737.
1365 <https://doi.org/10.1016/j.anbehav.2007.03.004>
- 1366 Eriksen, N., & Pakkenberg, B. (2007). Total neocortical cell number in the mysticete brain. *The*
1367 *Anatomical Record: Advances in Integrative Anatomy and Evolutionary Biology*, *290*(1),
1368 83–95. <https://doi.org/10.1002/ar.20404>
- 1369 Farris, S. M., & Schulmeister, S. (2011). Parasitoidism, not sociality, is associated with the
1370 evolution of elaborate mushroom bodies in the brains of hymenopteran insects.
1371 *Proceedings of the Royal Society of London B: Biological Sciences*, *278*(1707), 940–951.
1372 <https://doi.org/10.1098/rspb.2010.2161>
- 1373 Finlay, B. L., Darlington, R. B., & Nicastro, N. (2001). Developmental structure in brain
1374 evolution. *Behavioral and Brain Sciences*, *24*(2), 263–278.
1375 <https://doi.org/10.1017/S0140525X01003958>
- 1376 Fitzpatrick, J. L., Almbro, M., Gonzalez-Voyer, A., Hamada, S., Pennington, C., Scanlan, J., &
1377 Kolm, N. (2012). Sexual selection uncouples the evolution of brain and body size in
1378 pinnipeds. *Journal of Evolutionary Biology*, *25*(7), 1321–1330.
1379 <https://doi.org/10.1111/j.1420-9101.2012.02520.x>
- 1380 Freas, C. A., LaDage, L. D., Roth II, T. C., & Pravosudov, V. V. (2012). Elevation-related
1381 differences in memory and the hippocampus in mountain chickadees, *Poecile gambeli*.
1382 *Animal Behaviour*, *84*(1), 121–127. <https://doi.org/10.1016/j.anbehav.2012.04.018>

- 1383 Garland, T., Bennett, A. F., & Rezende, E. L. (2005). Phylogenetic approaches in comparative
1384 physiology. *Journal of Experimental Biology*, 208(16), 3015–3035.
1385 <https://doi.org/10.1242/jeb.01745>
- 1386 Glorot, X., & Bengio, Y. (2010). Understanding the difficulty of training deep feedforward neural
1387 networks. In Y. W. Teh & M. Titterton (Eds.), *Proceedings of the Thirteenth*
1388 *International Conference on Artificial Intelligence and Statistics* (Vol. 9, pp. 249–256).
1389 Sardinia, Italy: PMLR. Retrieved from <http://proceedings.mlr.press/v9/glorot10a.html>
- 1390 Gonda, A., Herczeg, G., & Merilä, J. (2011). Population variation in brain size of nine-spined
1391 sticklebacks (*Pungitius pungitius*) - local adaptation or environmentally induced variation?
1392 *BMC Evolutionary Biology*, 11, 75. <https://doi.org/10.1186/1471-2148-11-75>
- 1393 González-Lagos, C., Sol, D., & Reader, S. M. (2010). Large-brained mammals live longer.
1394 *Journal of Evolutionary Biology*, 23(5), 1064–1074. [https://doi.org/10.1111/j.1420-](https://doi.org/10.1111/j.1420-9101.2010.01976.x)
1395 [9101.2010.01976.x](https://doi.org/10.1111/j.1420-9101.2010.01976.x)
- 1396 Gonzalez-Voyer, A., & Hardenberg, A. von. (2014). An Introduction to Phylogenetic Path
1397 Analysis. In L. Z. Garamszegi (Ed.), *Modern Phylogenetic Comparative Methods and*
1398 *Their Application in Evolutionary Biology* (pp. 201–229). Berlin, Germany: Springer.
1399 https://doi.org/10.1007/978-3-662-43550-2_8
- 1400 Gonzalez-Voyer, A., Winberg, S., & Kolm, N. (2009). Distinct evolutionary patterns of brain and
1401 body size during adaptive radiation. *Evolution*, 63(9), 2266–2274.
1402 <https://doi.org/10.1111/j.1558-5646.2009.00705.x>
- 1403 Goodrum, M. R. (2016). The beginnings of human palaeontology: Prehistory, craniometry and
1404 the “fossil human races.” *The British Journal for the History of Science*, 49(3), 387–409.
1405 <https://doi.org/10.1017/S0007087416000674>

- 1406 Grasso, F. W., & Basil, J. A. (2009). The evolution of flexible behavioral repertoires in
1407 cephalopod molluscs. *Brain, Behavior and Evolution*, 74(3), 231–245.
1408 <https://doi.org/10.1159/000258669>
- 1409 Greenwood, A. K., Wark, A. R., Yoshida, K., & Peichel, C. L. (2013). Genetic and neural
1410 modularity underlie the evolution of schooling behavior in threespine sticklebacks. *Current*
1411 *Biology*, 23(19), 1884–1888. <https://doi.org/10.1016/j.cub.2013.07.058>
- 1412 Griffiths, P. E. (1994). Cladistic classification and functional explanation. *Philosophy of Science*,
1413 61(2), 206–227. <https://doi.org/10.1086/289796>
- 1414 Grinberg, L. T., Amaro, E., Teipel, S., Santos, D. D. dos, Pasqualucci, C. A., Leite, R. E. P., ...
1415 Group, B. A. B. S. (2008). Assessment of factors that confound MRI and neuropathological
1416 correlation of human postmortem brain tissue. *Cell and Tissue Banking*, 9(3), 195.
1417 <https://doi.org/10.1007/s10561-008-9080-5>
- 1418 Gronenberg, W., & Couvillon, M. J. (2010). Brain composition and olfactory learning in honey
1419 bees. *Neurobiology of Learning and Memory*, 93(3), 435–443.
1420 <https://doi.org/10.1016/j.nlm.2010.01.001>
- 1421 Gupta, A. S., van der Meer, M. A. A., Touretzky, D. S., & Redish, A. D. (2010). Hippocampal
1422 replay is not a simple function of experience. *Neuron*, 65(5), 695–705.
1423 <https://doi.org/10.1016/j.neuron.2010.01.034>
- 1424 Hager, R., Lu, L., Rosen, G. D., & Williams, R. W. (2012). Genetic architecture supports mosaic
1425 brain evolution and independent brain–body size regulation. *Nature Communications*, 3,
1426 1079. <https://doi.org/10.1038/ncomms2086>

- 1427 Hall, B. K. (2003). Descent with modification: The unity underlying homology and homoplasy
1428 as seen through an analysis of development and evolution. *Biological Reviews*, 78(3), 409–
1429 433. <https://doi.org/10.1017/S1464793102006097>
- 1430 Hall, Z. J., Street, S. E., & Healy, S. D. (2013). The evolution of cerebellum structure correlates
1431 with nest complexity. *Biology Letters*, 9(6), 20130687.
1432 <https://doi.org/10.1098/rsbl.2013.0687>
- 1433 Hampton, R. R., & Shettleworth, S. J. (1996). Hippocampal lesions impair memory for location
1434 but not color in passerine birds. *Behavioral Neuroscience*, 110(4), 831–835.
1435 <https://doi.org/10.1037/0735-7044.110.4.831>
- 1436 Harland, D., Li, D., & Jackson, R. (2012). How Jumping Spiders See the World. In O. Lazareva,
1437 T. Shimizu, & E. Wasserman (Eds.), *How Animals See the World: Comparative Behavior,*
1438 *Biology, and Evolution of Vision* (pp. 133–164). New York, USA: Oxford University Press.
- 1439 Harvey, P. H., & Bennett, P. M. (1983). Evolutionary biology: Brain size, energetics, ecology and
1440 life history patterns. *Nature*, 306(5941), 314–315.
- 1441 Harvey, P. H., & Pagel, M. D. (1991). *The Comparative Method in Evolutionary Biology* (Vol. 1).
1442 Oxford, UK: Oxford University Press.
- 1443 Healy, S. D., & Rowe, C. (2007). A critique of comparative studies of brain size. *Proceedings of*
1444 *the Royal Society of London B: Biological Sciences*, 274(1609), 453–464.
1445 <https://doi.org/10.1098/rspb.2006.3748>
- 1446 Herculano-Houzel, S. (2017). Numbers of neurons as biological correlates of cognitive
1447 capability. *Current Opinion in Behavioral Sciences*, 16, 1–7.
1448 <https://doi.org/10.1016/j.cobeha.2017.02.004>

- 1449 Herculano-Houzel, S., Catania, K., Manger, P. R., & Kaas, J. H. (2015). Mammalian brains are
1450 made of these: A dataset of the numbers and densities of neuronal and nonneuronal cells in
1451 the brain of glires, primates, scandentia, eulipotyphlans, afrotherians and artiodactyls, and
1452 their relationship with body mass. *Brain, Behavior and Evolution*, *86*(3–4), 145–163.
1453 <https://doi.org/10.1159/000437413>
- 1454 Herculano-Houzel, S., Collins, C. E., Wong, P., & Kaas, J. H. (2007). Cellular scaling rules for
1455 primate brains. *Proceedings of the National Academy of Sciences*, *104*(9), 3562–3567.
1456 <https://doi.org/10.1073/pnas.0611396104>
- 1457 Hochner, B. (2010). Functional and comparative assessments of the octopus learning and
1458 memory system. *Frontiers in Bioscience (Scholar Edition)*, *2*, 764–771.
1459 <https://doi.org/10.2741/s99>
- 1460 Hochner, B., Shomrat, T., & Fiorito, G. (2006). The octopus: A model for a comparative analysis
1461 of the evolution of learning and memory mechanisms. *The Biological Bulletin*, *210*(3),
1462 308–317. <https://doi.org/10.2307/4134567>
- 1463 Hopkins, W. D., Cantalupo, C., & Tagliabata, J. (2007). Handedness is associated with
1464 asymmetries in gyrification of the cerebral cortex of chimpanzees. *Cerebral Cortex*, *17*(8),
1465 1750–1756. <https://doi.org/10.1093/cercor/bhl085>
- 1466 Hunt, G. R., & Gray, R. D. (2002). Species-wide manufacture of stick-type tools by New
1467 Caledonian crows. *Emu - Austral Ornithology*, *102*(4), 349–353.
1468 <https://doi.org/10.1071/mu01056>
- 1469 Isler, K. (2011). Energetic trade-offs between brain size and offspring production: Marsupials
1470 confirm a general mammalian pattern. *BioEssays*, *33*(3), 173–179.
1471 <https://doi.org/10.1002/bies.201000123>

- 1472 Isler, K., Christopher Kirk, E., Miller, J. M. A., Albrecht, G. A., Gelvin, B. R., & Martin, R. D.
1473 (2008). Endocranial volumes of primate species: Scaling analyses using a comprehensive
1474 and reliable data set. *Journal of Human Evolution*, 55(6), 967–978.
1475 <https://doi.org/10.1016/j.jhevol.2008.08.004>
- 1476 Isler, K., & van Schaik, C. P. (2009). The Expensive Brain: A framework for explaining
1477 evolutionary changes in brain size. *Journal of Human Evolution*, 57(4), 392–400.
1478 <https://doi.org/10.1016/j.jhevol.2009.04.009>
- 1479 Iwaniuk, A. N., & Nelson, J. E. (2002). Can endocranial volume be used as an estimate of brain
1480 size in birds? *Canadian Journal of Zoology*, 80(1), 16–23. <https://doi.org/10.1139/z01-204>
- 1481 Iwaniuk, A. N., & Nelson, J. E. (2003). Developmental differences are correlated with relative
1482 brain size in birds: A comparative analysis. *Canadian Journal of Zoology*, 81(12), 1913–
1483 1928. <https://doi.org/10.1139/z03-190>
- 1484 Jackson, R. R., & Carter, C. M. (2001). Geographic variation in reliance on trial-and-error signal
1485 derivation by *Portia labiata*, an araneophagic jumping spider from the Philippines. *Journal*
1486 *of Insect Behavior*, 14(6), 799–827. <https://doi.org/10.1023/A:1013041501584>
- 1487 Jackson, R. R., Clark, R. J., & P. Harland, D. (2002). Behavioural and cognitive influences of
1488 kairomones on an araneophagic jumping spider. *Behaviour*, 139(6), 749–775.
1489 <https://doi.org/10.1163/156853902320262808>
- 1490 Jackson, R. R., & Cross, F. R. (2011). Spider Cognition. In J. Casas (Ed.), *Spider Physiology and*
1491 *Behaviour: Behaviour* (Vol. 41, pp. 115–174). London, UK: Academic Press.
- 1492 Jackson, R. R., & Cross, F. R. (2013). A cognitive perspective on aggressive mimicry. *Journal of*
1493 *Zoology*, 290(3), 161–171. <https://doi.org/10.1111/jzo.12036>

- 1494 Jackson, R. R., Cross, F. R., & Carter, C. M. (2006). Geographic variation in a spider's ability to
1495 solve a confinement problem by trial and error. *International Journal of Comparative*
1496 *Psychology*, 19(3).
- 1497 Jackson, R. R., & Li, D. (2004). One-encounter search-image formation by araneophagic spiders.
1498 *Animal Cognition*, 7(4), 247–254. <https://doi.org/10.1007/s10071-004-0219-x>
- 1499 Jackson, R. R., & Wilcox, R. S. (1993). Observations in nature of detouring behaviour by *Portia*
1500 *fimbriata*, a web-invading aggressive mimic jumping spider from Queensland. *Journal of*
1501 *Zoology*, 230(1), 135–139. <https://doi.org/10.1111/j.1469-7998.1993.tb02677.x>
- 1502 Jackson, R. R., & Wilcox, R. S. (1998). Spider-Eating Spiders: Despite the small size of their
1503 brain, jumping spiders in the genus *Portia* outwit other spiders with hunting techniques that
1504 include trial and error. *American Scientist*, 86(4), 350–357.
- 1505 Jacobs, I. F., & Osvath, M. (2015). The string-pulling paradigm in comparative psychology.
1506 *Journal of Comparative Psychology*, 129(2), 89–120. <https://doi.org/10.1037/a0038746>
- 1507 Jakob, E., Skow, C. D., & Long, S. M. (2011). Plasticity, Learning and Cognition. In M. E.
1508 Herberstein (Ed.), *Spider Behavior: Flexibility and Versatility* (pp. 307–347). Cambridge,
1509 UK: Cambridge University Press.
- 1510 Jarvis, E. D., Güntürkün, O., Bruce, L., Csillag, A., Karten, H., Kuenzel, W., ... Butler, A. B.
1511 (2005). Avian brains and a new understanding of vertebrate brain evolution. *Nature*
1512 *Reviews Neuroscience*, 6(2), 151–159. <https://doi.org/10.1038/nrn1606>
- 1513 Jelbert, S. A., Taylor, A. H., Cheke, L. G., Clayton, N. S., & Gray, R. D. (2014). Using the
1514 Aesop's Fable paradigm to investigate causal understanding of water displacement by New
1515 Caledonian crows. *PLOS ONE*, 9(3), e92895.
1516 <https://doi.org/10.1371/journal.pone.0092895>

- 1517 Jelbert, S. A., Taylor, A. H., & Gray, R. D. (2015). Investigating animal cognition with the
1518 Aesop's Fable paradigm: Current understanding and future directions. *Communicative &*
1519 *Integrative Biology*, 8(4), e1035846. <https://doi.org/10.1080/19420889.2015.1035846>
- 1520 Jelbert, S. A., Taylor, A. H., & Gray, R. D. (2016). Does absolute brain size really predict self-
1521 control? Hand-tracking training improves performance on the A-not-B task. *Biology*
1522 *Letters*, 12(2), 20150871. <https://doi.org/10.1098/rsbl.2015.0871>
- 1523 Jerison, H. J. (1969). Brain evolution and dinosaur brains. *The American Naturalist*, 103(934),
1524 575–588. <https://doi.org/10.1086/282627>
- 1525 Jerison, H. J. (1973). *Evolution of The Brain and Intelligence*. London, UK: Academic Press.
- 1526 Jerison, H. J. (1985). Animal intelligence as encephalization [and discussion]. *Philosophical*
1527 *Transactions of the Royal Society of London B: Biological Sciences*, 308(1135), 21–35.
1528 <https://doi.org/10.1098/rstb.1985.0007>
- 1529 Kaas, J. H., & Collins, C. E. (2001). Variability in the sizes of brain parts. *Behavioral and Brain*
1530 *Sciences*, 24(2), 288–290. <https://doi.org/10.1017/S0140525X01333952>
- 1531 Kabadayi, C., Taylor, L. A., Bayern, A. M. P. von, & Osvath, M. (2016). Ravens, New
1532 Caledonian crows and jackdaws parallel great apes in motor self-regulation despite smaller
1533 brains. *Royal Society Open Science*, 3(4), 160104. <https://doi.org/10.1098/rsos.160104>
- 1534 Kamil, A. C., & Bond, A. B. (2006). Selective Attention, Priming, and Foraging Behavior. In E.
1535 A. Wasserman & T. R. Zentall (Eds.), *Comparative Cognition: Experimental Explorations*
1536 *of Animal Intelligence* (pp. 106–126). New York, USA: Oxford University Press.
- 1537 Kanai, R., Bahrami, B., Roylance, R., & Rees, G. (2011). Online social network size is reflected
1538 in human brain structure. *Proc. R. Soc. B*, rspb20111959.
1539 <https://doi.org/10.1098/rspb.2011.1959>

- 1540 Kawai, M. (1965). Newly-acquired pre-cultural behavior of the natural troop of Japanese
1541 monkeys on Koshima islet. *Primates*, 6(1), 1–30. <https://doi.org/10.1007/BF01794457>
- 1542 Köhler, W. (1924). *The Mentality of Apes*. Oxford, UK: Harcourt, Brace.
- 1543 Kotrschal, A., Buechel, S. D., Zala, S. M., Corral-Lopez, A., Penn, D. J., & Kolm, N. (2015).
1544 Brain size affects female but not male survival under predation threat. *Ecology Letters*,
1545 18(7), 646–652. <https://doi.org/10.1111/ele.12441>
- 1546 Kotrschal, A., Deacon, A. E., Magurran, A. E., & Kolm, N. (2017). Predation pressure shapes
1547 brain anatomy in the wild. *Evolutionary Ecology*, 1–15. [https://doi.org/10.1007/s10682-](https://doi.org/10.1007/s10682-017-9901-8)
1548 017-9901-8
- 1549 Kotrschal, A., Lievens, E. J., Dahlbom, J., Bundsen, A., Semenova, S., Sundvik, M., ... Kolm, N.
1550 (2014). Artificial selection on relative brain size reveals a positive genetic correlation
1551 between brain size and proactive personality in the guppy. *Evolution*, 68(4), 1139–1149.
1552 <https://doi.org/10.1111/evo.12341>
- 1553 Kotrschal, A., Rogell, B., Bundsen, A., Svensson, B., Zajitschek, S., Brännström, I., ... Kolm, N.
1554 (2013). Artificial selection on relative brain size in the guppy reveals costs and benefits of
1555 evolving a larger brain. *Current Biology*, 23(2), 168–171.
1556 <https://doi.org/10.1016/j.cub.2012.11.058>
- 1557 Kowalko, J. E., Rohner, N., Rompani, S. B., Peterson, B. K., Linden, T. A., Yoshizawa, M., ...
1558 Tabin, C. J. (2013). Loss of schooling behavior in cavefish through sight-dependent and
1559 sight-independent mechanisms. *Current Biology*, 23(19), 1874–1883.
1560 <https://doi.org/10.1016/j.cub.2013.07.056>

- 1561 Krebs, J. R., Clayton, N. S., Healy, S. D., Cristol, D. A., Patel, S. N., & Jolliffe, A. R. (1996).
1562 The ecology of the avian brain: Food-storing memory and the hippocampus. *Ibis*, *138*(1),
1563 34–46. <https://doi.org/10.1111/j.1474-919X.1996.tb04311.x>
- 1564 LaBarbera, M. (1986). The Evolution and Ecology of Body Size. In D. Raup & D. Jablonski
1565 (Eds.), *Patterns and Processes in the History of Life* (Vol. 36, pp. 69–98). Berlin, Germany:
1566 Springer. https://doi.org/10.1007/978-3-642-70831-2_5
- 1567 Lai, C. S. L., Fisher, S. E., Hurst, J. A., Vargha-Khadem, F., & Monaco, A. P. (2001). A forkhead-
1568 domain gene is mutated in a severe speech and language disorder. *Nature*, *413*(6855), 519–
1569 523. <https://doi.org/10.1038/35097076>
- 1570 Laland, K. N., Odling-Smee, J., Hoppitt, W., & Uller, T. (2013). More on how and why: Cause
1571 and effect in biology revisited. *Biology & Philosophy*, *28*(5), 719–745.
1572 <https://doi.org/10.1007/s10539-012-9335-1>
- 1573 Laland, K. N., Sterelny, K., Odling-Smee, J., Hoppitt, W., & Uller, T. (2011). Cause and effect in
1574 biology revisited: Is Mayr’s proximate-ultimate dichotomy still useful? *Science*,
1575 *334*(6062), 1512–1516. <https://doi.org/10.1126/science.1210879>
- 1576 Larson, T. A., Thatra, N. M., Lee, B. H., & Brenowitz, E. A. (2014). Reactive neurogenesis in
1577 response to naturally occurring apoptosis in an adult brain. *Journal of Neuroscience*,
1578 *34*(39), 13066–13076. <https://doi.org/10.1523/JNEUROSCI.3316-13.2014>
- 1579 Larson, T. A., Wang, T.-W., Gale, S. D., Miller, K. E., Thatra, N. M., Caras, M. L., ... Brenowitz,
1580 E. A. (2013). Postsynaptic neural activity regulates neuronal addition in the adult avian
1581 song control system. *Proceedings of the National Academy of Sciences*, *110*(41), 16640–
1582 16644. <https://doi.org/10.1073/pnas.1310237110>

- 1583 Laughlin, S. B., de Ruyter van Steveninck, R. R., & Anderson, J. C. (1998). The metabolic cost
1584 of neural information. *Nature Neuroscience*, *1*(1), 36–41. <https://doi.org/10.1038/236>
- 1585 Lefebvre, L., Whittle, P., Lascaris, E., & Finkelstein, A. (1997). Feeding innovations and
1586 forebrain size in birds. *Animal Behaviour*, *53*(3), 549–560.
1587 <https://doi.org/10.1006/anbe.1996.0330>
- 1588 Legg, S., & Hutter, M. (2007). Universal intelligence: A definition of machine intelligence.
1589 *Minds and Machines*, *17*(4), 391–444. <https://doi.org/10.1007/s11023-007-9079-x>
- 1590 Li, D., Jackson, R. R., & Barrion, A. T. (1999). Parental and predatory behaviour of *Scytodes* sp.,
1591 an araneophagic spitting spider (Araneae: Scytodidae) from the Philippines. *Journal of*
1592 *Zoology*, *247*(3), 293–310. <https://doi.org/10.1111/j.1469-7998.1999.tb00993.x>
- 1593 Logan, C. J. (2016a). Behavioral flexibility and problem solving in an invasive bird. *PeerJ*, *4*,
1594 e1975. <https://doi.org/10.7717/peerj.1975>
- 1595 Logan, C. J. (2016b). How far will a behaviourally flexible invasive bird go to innovate? *Royal*
1596 *Society Open Science*, *3*(6), 160247. <https://doi.org/10.1098/rsos.160247>
- 1597 Logan, C. J., Harvey, B. D., Schlinger, B. A., & Rensel, M. (2016). Western scrub-jays do not
1598 appear to attend to functionality in Aesop’s Fable experiments. *PeerJ*, *4*, e1707.
1599 <https://doi.org/10.7717/peerj.1707>
- 1600 Logan, C. J., Jelbert, S. A., Breen, A. J., Gray, R. D., & Taylor, A. H. (2014). Modifications to the
1601 Aesop’s Fable paradigm change New Caledonian crow performances. *PLOS ONE*, *9*(7),
1602 e103049. <https://doi.org/10.1371/journal.pone.0103049>
- 1603 Logan, C. J., Kruuk, L. E. B., Stanley, R., Thompson, A. M., & Clutton-Brock, T. H. (2016).
1604 Endocranial volume is heritable and is associated with longevity and fitness in a wild
1605 mammal. *Royal Society Open Science*, *3*(12), 160622. <https://doi.org/10.1098/rsos.160622>

- 1606 MacLean, E. L., Hare, B., Nunn, C. L., Addessi, E., Amici, F., Anderson, R. C., ... Zhao, Y.
1607 (2014). The evolution of self-control. *Proceedings of the National Academy of Sciences*,
1608 *111*(20), E2140–E2148. <https://doi.org/10.1073/pnas.1323533111>
- 1609 Macphail, E. M. (1982). *Brain and Intelligence in Vertebrates*. Oxford, UK: Clarendon Press.
- 1610 Madison, P. (2016). The most brutal of human skulls: Measuring and knowing the first
1611 Neanderthal. *British Journal for the History of Science*, *49*(3), 411–432.
1612 <https://doi.org/10.1017/S0007087416000650>
- 1613 Manrique, H. M., Völter, C. J., & Call, J. (2013). Repeated innovation in great apes. *Animal*
1614 *Behaviour*, *85*(1), 195–202. <https://doi.org/10.1016/j.anbehav.2012.10.026>
- 1615 Mars, R. B., Neubert, F.-X., Verhagen, L., Sallet, J., Miller, K. L., Dunbar, R. I. M., & Barton, R.
1616 A. (2014). Primate comparative neuroscience using magnetic resonance imaging: Promises
1617 and challenges. *Frontiers in Neuroscience*, *8*. <https://doi.org/10.3389/fnins.2014.00298>
- 1618 Maseko, B. C., Spocter, M. A., Haagensen, M., & Manger, P. R. (2012). Elephants have
1619 relatively the largest cerebellum size of mammals. *The Anatomical Record: Advances in*
1620 *Integrative Anatomy and Evolutionary Biology*, *295*(4), 661–672.
1621 <https://doi.org/10.1002/ar.22425>
- 1622 Matthewson, J. (2011). Trade-offs in model-building: A more target-oriented approach. *Studies*
1623 *in History and Philosophy of Science Part A*, *42*(2), 324–333.
1624 <https://doi.org/10.1016/j.shpsa.2010.11.040>
- 1625 McGhee, G. R. (2011). *Convergent Evolution: Limited Forms Most Beautiful*. MIT Press.
- 1626 Meitzen, J., Thompson, C. K., Choi, H., Perkel, D. J., & Brenowitz, E. A. (2009). Time course of
1627 changes in Gambel's white-crowned sparrow song behavior following transitions in

- 1628 breeding condition. *Hormones and Behavior*, 55(1), 217–227.
1629 <https://doi.org/10.1016/j.yhbeh.2008.10.006>
- 1630 Menda, G., Shamble, P. S., Nitzany, E. I., Golden, J. R., & Hoy, R. R. (2014). Visual perception
1631 in the brain of a jumping spider. *Current Biology*, 24(21), 2580–2585.
1632 <https://doi.org/10.1016/j.cub.2014.09.029>
- 1633 Menzel, R., & Giurfa, M. (2001). Cognitive architecture of a mini-brain: The honeybee. *Trends*
1634 *in Cognitive Sciences*, 5(2), 62–71. [https://doi.org/10.1016/S1364-6613\(00\)01601-6](https://doi.org/10.1016/S1364-6613(00)01601-6)
- 1635 Meyer, W., & Idel, K. (1977). The distribution of acetylcholinesterase in the central nervous
1636 system of jumping spiders and wolf spiders (*Arachnida, Araneida: Salticidae et*
1637 *Lycosidae*). *The Journal of Comparative Neurology*, 173(4), 717–743.
1638 <https://doi.org/10.1002/cne.901730406>
- 1639 Mikhalevich, I., Powell, R., & Logan, C. (2017). Is behavioural flexibility evidence of cognitive
1640 complexity? How evolution can inform comparative cognition. *Interface Focus*, 7(3),
1641 20160121. <https://doi.org/10.1098/rsfs.2016.0121>
- 1642 Miller, K. L., Stagg, C. J., Douaud, G., Jbabdi, S., Smith, S. M., Behrens, T. E. J., ... McNab, J.
1643 A. (2011). Diffusion imaging of whole, post-mortem human brains on a clinical MRI
1644 scanner. *NeuroImage*, 57(1), 167–181. <https://doi.org/10.1016/j.neuroimage.2011.03.070>
- 1645 Miller, R., Jelbert, S. A., Taylor, A. H., Cheke, L. G., Gray, R. D., Loissel, E., & Clayton, N. S.
1646 (2016). Performance in object-choice Aesop's Fable tasks are influenced by object biases
1647 in New Caledonian crows but not in human children. *PLOS ONE*, 11(12), e0168056.
1648 <https://doi.org/10.1371/journal.pone.0168056>

- 1649 Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., ... Hassabis, D.
1650 (2015). Human-level control through deep reinforcement learning. *Nature*, *518*(7540),
1651 529–533. <https://doi.org/10.1038/nature14236>
- 1652 Montgomery, S. H. (2017). Evolution of Large Brain and Body Size in Mammals. In J. H. Kaas
1653 (Ed.), *Evolution of Nervous Systems* (2nd ed., pp. 103–136). Oxford, UK: Academic Press.
1654 <https://doi.org/10.1016/B978-0-12-804042-3.00034-8>
- 1655 Montgomery, S. H., Capellini, I., Barton, R. A., & Mundy, N. I. (2010). Reconstructing the ups
1656 and downs of primate brain evolution: Implications for adaptive hypotheses and *Homo*
1657 *floresiensis*. *BMC Biology*, *8*, 9. <https://doi.org/10.1186/1741-7007-8-9>
- 1658 Montgomery, S. H., Geisler, J. H., McGowen, M. R., Fox, C., Marino, L., & Gatesy, J. (2013).
1659 The evolutionary history of cetacean brain and body size. *Evolution*, *67*(11), 3339–3353.
1660 <https://doi.org/10.1111/evo.12197>
- 1661 Montgomery, S. H., & Merrill, R. M. (2017). Divergence in brain composition during the early
1662 stages of ecological specialization in *Heliconius* butterflies. *Journal of Evolutionary*
1663 *Biology*, *30*(3), 571–582. <https://doi.org/10.1111/jeb.13027>
- 1664 Montgomery, S. H., & Mundy, N. I. (2013). Microcephaly genes and the evolution of sexual
1665 dimorphism in primate brain size. *Journal of Evolutionary Biology*, *26*(4), 906–911.
1666 <https://doi.org/10.1111/jeb.12091>
- 1667 Mulcahy, N. J., & Call, J. (2006). How great apes perform on a modified trap-tube task. *Animal*
1668 *Cognition*, *9*(3), 193–199. <https://doi.org/10.1007/s10071-006-0019-6>
- 1669 Nair, A., Srinivasan, P., Blackwell, S., Alcicek, C., Fearon, R., De Maria, A., ... Silver, D.
1670 (2015). Massively parallel methods for deep reinforcement learning. Presented at the Deep

- 1671 Learning Workshop, International Conference on Machine Learning, Lille, France.
1672 Retrieved from <http://arxiv.org/abs/1507.04296>
- 1673 Nakagawa, S., & Santos, E. S. A. (2012). Methodological issues and advances in biological
1674 meta-analysis. *Evolutionary Ecology*, 26(5), 1253–1274. [https://doi.org/10.1007/s10682-](https://doi.org/10.1007/s10682-012-9555-5)
1675 012-9555-5
- 1676 Navarrete, A. F., Reader, S. M., Street, S. E., Whalen, A., & Laland, K. N. (2016). The
1677 coevolution of innovation and technical intelligence in primates. *Phil. Trans. R. Soc. B*,
1678 371(1690), 20150186. <https://doi.org/10.1098/rstb.2015.0186>
- 1679 Noreikiene, K., Herczeg, G., Gonda, A., Balázs, G., Husby, A., & Merilä, J. (2015). Quantitative
1680 genetic analysis of brain size variation in sticklebacks: Support for the mosaic model of
1681 brain evolution. *Proc. R. Soc. B*, 282(1810), 20151008.
1682 <https://doi.org/10.1098/rspb.2015.1008>
- 1683 Nyberg, D. (1971). A hypothesis concerning the larger brains of homoiotherms. *The American*
1684 *Naturalist*, 105(942), 183–185. <https://doi.org/10.1086/282713>
- 1685 Olkowicz, S., Kocourek, M., Lučan, R. K., Porteš, M., Fitch, W. T., Herculano-Houzel, S., &
1686 Němec, P. (2016). Birds have primate-like numbers of neurons in the forebrain.
1687 *Proceedings of the National Academy of Sciences*, 113(26), 7255–7260.
1688 <https://doi.org/10.1073/pnas.1517131113>
- 1689 O’regan, H. J., & Kitchener, A. C. (2005). The effects of captivity on the morphology of captive,
1690 domesticated and feral mammals. *Mammal Review*, 35(3–4), 215–230.
1691 <https://doi.org/10.1111/j.1365-2907.2005.00070.x>

- 1692 Overington, S. E., Morand-Ferron, J., Boogert, N. J., & Lefebvre, L. (2009). Technical
1693 innovations drive the relationship between innovativeness and residual brain size in birds.
1694 *Animal Behaviour*, 78(4), 1001–1010. <https://doi.org/10.1016/j.anbehav.2009.06.033>
- 1695 Packard, A. (1972). Cephalopods and fish: The limits of convergence. *Biological Reviews*, 47(2),
1696 241–307. <https://doi.org/10.1111/j.1469-185X.1972.tb00975.x>
- 1697 Pagel, M. (1999). Inferring the historical patterns of biological evolution. *Nature*, 401(6756),
1698 877–884. <https://doi.org/10.1038/44766>
- 1699 Pagel, M., & Meade, A. (2006). Bayesian analysis of correlated evolution of discrete characters
1700 by reversible–jump Markov chain Monte Carlo. *The American Naturalist*, 167(6), 808–
1701 825. <https://doi.org/10.1086/503444>
- 1702 Park, P. J., & Bell, M. A. (2010). Variation of telencephalon morphology of the threespine
1703 stickleback (*Gasterosteus aculeatus*) in relation to inferred ecology. *Journal of*
1704 *Evolutionary Biology*, 23(6), 1261–1277. [https://doi.org/10.1111/j.1420-](https://doi.org/10.1111/j.1420-9101.2010.01987.x)
1705 9101.2010.01987.x
- 1706 Pascanu, R., Mikolov, T., & Bengio, Y. (2013). On the difficulty of training recurrent neural
1707 networks. In S. Dasgupta & D. McAllester (Eds.), *Proceedings of the 30th International*
1708 *Conference on Machine Learning* (Vol. 28, pp. 1310–1318). Atlanta, USA: PMLR.
1709 Retrieved from <http://proceedings.mlr.press/v28/pascanu13.html>
- 1710 Patel, S. N., Clayton, N. S., & Krebs, J. R. (1997). Hippocampal tissue transplants reverse lesion-
1711 induced spatial memory deficits in zebra finches (*Taeniopygia guttata*). *Journal of*
1712 *Neuroscience*, 17(10), 3861–3869.

- 1713 Pearce, T. (2012). Convergence and parallelism in evolution: A neo-Gouldian account. *The*
1714 *British Journal for the Philosophy of Science*, 63(2), 429–448.
1715 <https://doi.org/10.1093/bjps/axr046>
- 1716 Pedersen, N. L., Plomin, R., Nesselroade, J. R., & McClearn, G. E. (1992). A quantitative genetic
1717 analysis of cognitive abilities during the second half of the life span. *Psychological*
1718 *Science*, 3(6), 346–353. <https://doi.org/10.1111/j.1467-9280.1992.tb00045.x>
- 1719 Pepperberg, I. M., & Kozak, F. A. (1986). Object permanence in the African grey parrot
1720 (*Psittacus erithacus*). *Animal Learning & Behavior*, 14(3), 322–330.
1721 <https://doi.org/10.3758/BF03200074>
- 1722 Perry, C. J., Barron, A. B., & Chittka, L. (2017). The frontiers of insect cognition. *Current*
1723 *Opinion in Behavioral Sciences*, 16, 111–118. <https://doi.org/10.1016/j.cobeha.2017.05.011>
- 1724 Petraitis, P. S., Dunham, A. E., & Niewiarowski, P. H. (1996). Inferring multiple causality: The
1725 limitations of path analysis. *Functional Ecology*, 10(4), 421–431.
1726 <https://doi.org/10.2307/2389934>
- 1727 Pfeifer, R., & Iida, F. (2005). Morphological computation: Connecting body, brain and
1728 environment. *Japanese Scientific Monthly*, 58(2), 48–54.
- 1729 Pfenning, A. R., Hara, E., Whitney, O., Rivas, M. V., Wang, R., Roulhac, P. L., ... Jarvis, E. D.
1730 (2014). Convergent transcriptional specializations in the brains of humans and song-
1731 learning birds. *Science*, 346(6215), 1256846. <https://doi.org/10.1126/science.1256846>
- 1732 Powell, R. (2007). Is convergence more than an analogy? Homoplasy and its implications for
1733 macroevolutionary predictability. *Biology & Philosophy*, 22(4), 565–578.
1734 <https://doi.org/10.1007/s10539-006-9057-3>

- 1735 Pravosudov, V. V., & Clayton, N. S. (2002). A test of the adaptive specialization hypothesis:
1736 Population differences in caching, memory, and the hippocampus in black-capped
1737 chickadees (*Poecile atricapilla*). *Behavioral Neuroscience*, 116(4), 515–522.
1738 <https://doi.org/10.1037/0735-7044.116.4.515>
- 1739 Pravosudov, V. V., & de Kort, S. R. (2006). Is the western scrub-jay (*Aphelocoma californica*)
1740 really an underdog among food-caching corvids when it comes to hippocampal volume and
1741 food caching propensity? *Brain, Behavior and Evolution*, 67(1), 1–9.
1742 <https://doi.org/10.1159/000088855>
- 1743 Punzo, F., & Ludwig, L. (2002). Contact with maternal parent and siblings affects hunting
1744 behavior, learning, and central nervous system development in spiderlings of *Hogna*
1745 *carolinensis* (Araeneae: Lycosidae). *Animal Cognition*, 5(2), 63–70.
1746 <https://doi.org/10.1007/s10071-002-0128-9>
- 1747 Ramnani, N. (2006). The primate cortico-cerebellar system: Anatomy and function. *Nature*
1748 *Reviews Neuroscience*, 7(7), 511–522. <https://doi.org/10.1038/nrn1953>
- 1749 Reader, S. M., Hager, Y., & Laland, K. N. (2011). The evolution of primate general and cultural
1750 intelligence. *Philosophical Transactions of the Royal Society of London B: Biological*
1751 *Sciences*, 366(1567), 1017–1027. <https://doi.org/10.1098/rstb.2010.0342>
- 1752 Reader, S. M., & Laland, K. N. (2002). Social intelligence, innovation, and enhanced brain size
1753 in primates. *Proceedings of the National Academy of Sciences*, 99(7), 4436–4441.
1754 <https://doi.org/10.1073/pnas.062041299>
- 1755 Ridgway, S. H., Carlin, K. P., Alstyne, K. R. V., Hanson, A. C., & Tarpley, R. J. (2016).
1756 Comparison of dolphins' body and brain measurements with four other groups of cetaceans

- 1757 reveals great diversity. *Brain, Behavior and Evolution*, 88(3–4), 235–257.
1758 <https://doi.org/10.1159/000454797>
- 1759 Ringo, J. L. (1991). Neuronal interconnection as a function of brain size. *Brain, Behavior and*
1760 *Evolution*, 38(1), 1–6. <https://doi.org/10.1159/000114375>
- 1761 Rodríguez, R. L., Briceño, R. D., Briceño-Aguilar, E., & Höbel, G. (2015). *Nephila clavipes*
1762 spiders (Araneae: Nephilidae) keep track of captured prey counts: testing for a sense of
1763 numerosity in an orb-weaver. *Animal Cognition*, 18(1), 307–314.
1764 <https://doi.org/10.1007/s10071-014-0801-9>
- 1765 Roth, G. (2013). *The Long Evolution of Brains and Minds*. Heidelberg, Germany: Springer.
- 1766 Roth, T. C., & Pravosudov, V. V. (2009). Hippocampal volumes and neuron numbers increase
1767 along a gradient of environmental harshness: A large-scale comparison. *Proceedings of the*
1768 *Royal Society of London B: Biological Sciences*, 276(1656), 401–405.
1769 <https://doi.org/10.1098/rspb.2008.1184>
- 1770 Rowe, C., & Healy, S. D. (2014). Measuring variation in cognition. *Behavioral Ecology*, 25(6),
1771 1287–1292. <https://doi.org/10.1093/beheco/aru090>
- 1772 Rubinstein, H. S. (1936). The effect of the growth hormone upon the brain and brain weight-
1773 body weight relations. *The Journal of Comparative Neurology*, 64(3), 469–496.
1774 <https://doi.org/10.1002/cne.900640308>
- 1775 Sallet, J., Mars, R. B., Noonan, M. P., Andersson, J. L., O'Reilly, J. X., Jbabdi, S., ... Rushworth,
1776 M. F. S. (2011). Social network size affects neural circuits in macaques. *Science*,
1777 334(6056), 697–700. <https://doi.org/10.1126/science.1210027>
- 1778 Schaaffhausen, H. (1858). Zur Kenntnis der ältesten Rasseschädel. *Archiv Für Anatomie,*
1779 *Physiologie Und Wissenschaftliche Medicin*, 453–478.

- 1780 Schönbrodt, F. D., & Perugini, M. (2013). At what sample size do correlations stabilize? *Journal*
1781 *of Research in Personality*, 47(5), 609–612. <https://doi.org/10.1016/j.jrp.2013.05.009>
- 1782 Schreiweis, C., Bornschein, U., Burguière, E., Kerimoglu, C., Schreiter, S., Dannemann, M., ...
1783 Graybiel, A. M. (2014). Humanized Foxp2 accelerates learning by enhancing transitions
1784 from declarative to procedural performance. *Proceedings of the National Academy of*
1785 *Sciences*, 111(39), 14253–14258. <https://doi.org/10.1073/pnas.1414542111>
- 1786 Scott-Phillips, T. C., Dickins, T. E., & West, S. A. (2011). Evolutionary theory and the ultimate–
1787 proximate distinction in the human behavioral sciences. *Perspectives on Psychological*
1788 *Science*, 6(1), 38–47. <https://doi.org/10.1177/1745691610393528>
- 1789 Seed, A. M., Tebbich, S., Emery, N. J., & Clayton, N. S. (2006). Investigating physical cognition
1790 i n r o o k s , *Corvus frugilegus*. *Current Biology*, 16(7), 697–701.
1791 <https://doi.org/10.1016/j.cub.2006.02.066>
- 1792 Seed, A., Seddon, E., Greene, B., & Call, J. (2012). Chimpanzee “folk physics”: Bringing
1793 failures into focus. *Philosophical Transactions of the Royal Society of London B:*
1794 *Biological Sciences*, 367(1603), 2743–2752. <https://doi.org/10.1098/rstb.2012.0222>
- 1795 Shaw, R. C., Plotnik, J. M., & Clayton, N. S. (2013). Exclusion in corvids: The performance of
1796 food-caching Eurasian jays (*Garrulus glandarius*). *Journal of Comparative Psychology*,
1797 127(4), 428–435. <https://doi.org/10.1037/a0032010>
- 1798 Shettleworth, S. J. (2010). *Cognition, Evolution, and Behavior* (2nd ed., Vol. xiii). New York,
1799 USA: Oxford University Press.
- 1800 Shigeno, S. (2017). Brain Evolution as an Information Flow Designer: The Ground Architecture
1801 for Biological and Artificial General Intelligence. In S. Shigeno, Y. Murakami, & T.

- 1802 Nomura (Eds.), *Brain Evolution by Design* (pp. 415–438). Japan: Springer.
1803 https://doi.org/10.1007/978-4-431-56469-0_19
- 1804 Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., van den Driessche, G., ... Hassabis,
1805 D. (2016). Mastering the game of Go with deep neural networks and tree search. *Nature*,
1806 *529*(7587), 484–489. <https://doi.org/10.1038/nature16961>
- 1807 Smith, L. B., Thelen, E., Titzer, R., & McLin, D. (1999). Knowing in the context of acting: The
1808 task dynamics of the A-not-B error. *Psychological Review*, *106*(2), 235–260.
1809 <https://doi.org/10.1037/0033-295X.106.2.235>
- 1810 Snell, O. (1892). Die Abhängigkeit des Hirngewichtes von dem Körpergewicht und den geistigen
1811 Fähigkeiten. *European Archives of Psychiatry and Clinical Neuroscience*, *23*(2), 436–446.
1812 <https://doi.org/10.1007/BF01843462>
- 1813 Snell-Rood, E. C., Papaj, D. R., & Gronenberg, W. (2009). Brain size: A global or induced cost
1814 of learning? *Brain, Behavior and Evolution*, *73*(2), 111–128.
1815 <https://doi.org/10.1159/000213647>
- 1816 Sol, D., & Lefebvre, L. (2000). Behavioural flexibility predicts invasion success in birds
1817 introduced to New Zealand. *Oikos*, *90*(3), 599–605. <https://doi.org/10.1034/j.1600-0706.2000.900317.x>
- 1818
- 1819 Sol, D., Székely, T., Liker, A., & Lefebvre, L. (2007). Big-brained birds survive better in nature.
1820 *Proceedings of the Royal Society of London B: Biological Sciences*, *274*(1611), 763–769.
1821 <https://doi.org/10.1098/rspb.2006.3765>
- 1822 Sol, D., Timmermans, S., & Lefebvre, L. (2002). Behavioural flexibility and invasion success in
1823 birds. *Animal Behaviour*, *63*(3), 495–502. <https://doi.org/10.1006/anbe.2001.1953>

- 1824 Strausfeld, N. J., & Hirth, F. (2013). Deep homology of arthropod central complex and vertebrate
1825 basal ganglia. *Science*, 340(6129), 157–161. <https://doi.org/10.1126/science.1231828>
- 1826 Swaddle, J. P. (2016). Evolution and Conservation Behavior. In O. Berger-Tal & D. Saltz (Eds.),
1827 *Conservation Behavior: Applying Behavioral Ecology to Wildlife Conservation and*
1828 *Management* (Vol. 21, pp. 36–65). Cambridge, UK: Cambridge University Press.
- 1829 Tarsitano, M. S., & Jackson, R. R. (1997). Araneophagic jumping spiders discriminate between
1830 detour routes that do and do not lead to prey. *Animal Behaviour*, 53(2), 257–266.
1831 <https://doi.org/10.1006/anbe.1996.0372>
- 1832 Tebbich, S., Sterelny, K., & Teschke, I. (2010). The tale of the finch: Adaptive radiation and
1833 behavioural flexibility. *Philosophical Transactions of the Royal Society of London B:*
1834 *Biological Sciences*, 365(1543), 1099–1109. <https://doi.org/10.1098/rstb.2009.0291>
- 1835 Thornton, A., & Lukas, D. (2012). Individual variation in cognitive performance: Developmental
1836 and evolutionary perspectives. *Philosophical Transactions of the Royal Society of London*
1837 *B: Biological Sciences*, 367(1603), 2773–2783. <https://doi.org/10.1098/rstb.2012.0214>
- 1838 Thornton, A., & Samson, J. (2012). Innovative problem solving in wild meerkats. *Animal*
1839 *Behaviour*, 83(6), 1459–1468. <https://doi.org/10.1016/j.anbehav.2012.03.018>
- 1840 Tomasello, M. (1999). *The Cultural Origins of Human Cognition*. Cambridge, USA: Harvard
1841 University Press.
- 1842 Tomer, R., Denes, A. S., Tessmar-Raible, K., & Arendt, D. (2010). Profiling by image
1843 registration reveals common origin of annelid mushroom bodies and vertebrate pallium.
1844 *Cell*, 142(5), 800–809. <https://doi.org/10.1016/j.cell.2010.07.043>
- 1845 Tramontin, A. D., & Brenowitz, E. A. (2000). Seasonal plasticity in the adult brain. *Trends in*
1846 *Neurosciences*, 23(6), 251–258. [https://doi.org/10.1016/S0166-2236\(00\)01558-7](https://doi.org/10.1016/S0166-2236(00)01558-7)

- 1847 Tucker, A. (1998). Unique events: The underdetermination of explanation. *Erkenntnis*, 48(1), 59–
1848 80.
- 1849 van der Bijl, W., Thyselius, M., Kotschal, A., & Kolm, N. (2015). Brain size affects the
1850 behavioural response to predators in female guppies (*Poecilia reticulata*). *Proc. R. Soc. B*,
1851 282(1812), 20151132. <https://doi.org/10.1098/rspb.2015.1132>
- 1852 van Schaik, C. P., & Burkart, J. M. (2011). Social learning and evolution: The cultural
1853 intelligence hypothesis. *Philosophical Transactions of the Royal Society of London B:*
1854 *Biological Sciences*, 366(1567), 1008–1016. <https://doi.org/10.1098/rstb.2010.0304>
- 1855 van Schaik, C. P., Isler, K., & Burkart, J. M. (2012). Explaining brain size variation: From social
1856 to cultural brain. *Trends in Cognitive Sciences*, 16(5), 277–284.
1857 <https://doi.org/10.1016/j.tics.2012.04.004>
- 1858 Ventura-Antunes, L., Mota, B., & Herculano-Houzel, S. (2013). Different scaling of white matter
1859 volume, cortical connectivity, and gyrification across rodent and primate brains. *Frontiers*
1860 *in Neuroanatomy*, 7. <https://doi.org/10.3389/fnana.2013.00003>
- 1861 Walsh, M. R., Broyles, W., Beston, S. M., & Munch, S. B. (2016). Predator-driven brain size
1862 evolution in natural populations of Trinidadian killifish (*Rivulus hartii*). *Proc. R. Soc. B*,
1863 283(1834), 20161075. <https://doi.org/10.1098/rspb.2016.1075>
- 1864 Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of “small-world” networks. *Nature*,
1865 393(6684), 440–442. <https://doi.org/10.1038/30918>
- 1866 Weber, J. N., Peterson, B. K., & Hoekstra, H. E. (2013). Discrete genetic modules are
1867 responsible for complex burrow evolution in *Peromyscus* mice. *Nature*, 493(7432), 402–
1868 405. <https://doi.org/10.1038/nature11816>

- 1869 Whitehead, A. N. (1925). The Origins of Modern Science. In A. I. Tauber (Ed.), *Science and the*
1870 *Quest for Reality* (pp. 53–69). New York, USA: Palgrave Macmillan.
1871 https://doi.org/10.1007/978-1-349-25249-7_2
- 1872 Whiting, B. A., & Barton, R. A. (2003). The evolution of the cortico-cerebellar complex in
1873 primates: Anatomical connections predict patterns of correlated evolution. *Journal of*
1874 *Human Evolution*, 44(1), 3–10. [https://doi.org/10.1016/S0047-2484\(02\)00162-8](https://doi.org/10.1016/S0047-2484(02)00162-8)
- 1875 Workman, A. D., Charvet, C. J., Clancy, B., Darlington, R. B., & Finlay, B. L. (2013). Modeling
1876 transformations of neurodevelopmental sequences across mammalian species. *Journal of*
1877 *Neuroscience*, 33(17), 7368–7383. <https://doi.org/10.1523/JNEUROSCI.5746-12.2013>
- 1878 Wyles, J. S., Kunkel, J. G., & Wilson, A. C. (1983). Birds, behavior, and anatomical evolution.
1879 *Proceedings of the National Academy of Sciences*, 80(14), 4394–4397.
- 1880 Zatorre, R. J., Chen, J. L., & Penhune, V. B. (2007). When the brain plays music: Auditory–
1881 motor interactions in music perception and production. *Nature Reviews Neuroscience*, 8(7),
1882 547–558. <https://doi.org/10.1038/nrn2152>
- 1883 Zhang, S., Bock, F., Si, A., Tautz, J., & Srinivasan, M. V. (2005). Visual working memory in
1884 decision making by honey bees. *Proceedings of the National Academy of Sciences of the*
1885 *United States of America*, 102(14), 5250–5255. <https://doi.org/10.1073/pnas.0501440102>
- 1886 Zilles, K., Amunts, K., & Smaers, J. B. (2011). Three brain collections for comparative
1887 neuroanatomy and neuroimaging. *Annals of the New York Academy of Sciences*, 1225(S1),
1888 E94–E104. <https://doi.org/10.1111/j.1749-6632.2011.05978.x>