

Four errors and a fallacy: pitfalls for the unwary in comparative brain analyses

Robin I. M. Dunbar^{1,*}  and Susanne Shultz²

¹*Department of Experimental Psychology, Anna Watts Building, University of Oxford, Oxford, OX2 6GG, UK*

²*Department of Earth and Environmental Sciences, Michael Smith Building, University of Manchester, Manchester, M13 9PT, UK*

ABSTRACT

Comparative analyses are the backbone of evolutionary analysis. However, their record in producing a consensus has not always been good. This is especially true of attempts to understand the factors responsible for the evolution of large brains, which have been embroiled in an increasingly polarised debate over the past three decades. We argue that most of these disputes arise from a number of conceptual errors and associated logical fallacies that are the result of a failure to adopt a biological systems-based approach to hypothesis-testing. We identify four principal classes of error: a failure to heed Tinbergen's Four Questions when testing biological hypotheses, misapplying Dobzhansky's Dictum when testing hypotheses of evolutionary adaptation, poorly chosen behavioural proxies for underlying hypotheses, and the use of inappropriate statistical methods. In the interests of progress, we urge a more careful and considered approach to comparative analyses, and the adoption of a broader, rather than a narrower, taxonomic perspective.

Key words: Tinbergen's Four Questions, Dobzhansky's Dictum, social brain hypothesis, regression analysis, sloppy definitions.

CONTENTS

I. Introduction	1279
II. Conceptual confounds	1279
(1) Conflating Tinbergen's Questions	1280
(2) The fallacy of the missing middle	1283
III. The mismeasures of fitness	1284
(1) Dobzhansky's Dictum	1284
(2) Fitness top-down and bottom-up	1285
IV. Critical tests and sloppy proxies	1287
(1) Critical tests	1287
(2) Sloppy proxies	1288
(3) The problem of relativity	1291
(4) The fallacy of 'secundum quid' (over-generalisation)	1294
V. Statistical pitfalls	1295
(1) How to test the wrong hypothesis	1295
(2) Simpson's Paradox	1297
(3) When a regression is not the right regression	1299
(4) The art of prediction	1299
VI. Discussion	1301
VII. Conclusions	1303
VIII. References	1303

* Author for correspondence (Tel.: +44(0)7909 785269; E-mail: robin.dunbar@psy.ox.ac.uk).

I. INTRODUCTION

Comparative analyses have been the workhorse of evolutionary analysis ever since Darwin first made seminal use of the method as a means of understanding the evolution of adaptations. Over the decades, ever more sophisticated statistical methods have been developed to enable more nuanced analyses to be undertaken, culminating in the widely used phylogenetic methods of the last quarter century. Although most applications of comparative analyses have been relatively uncontroversial, some topics have become so mired in controversy that even modern statistical methods seem powerless to resolve them. Attempts to explain why large brains evolved in primates, in particular, seem to have been unusually prone to vicariously polarised disputes of this kind.

Ever since Jerison (1977) first pointed it out, the fact that primates have much larger brains (absolutely and relative to body size) than any other group of animals has continued to attract interest and debate, with the debate spilling out into other vertebrate (and even invertebrate) orders over time. The issue can be stated quite simply: given that vertebrate brains are unusually expensive to evolve, grow and maintain (the Expensive Tissue Hypothesis; Aiello & Wheeler, 1995), why would any species want to invest so heavily in them? Or to put it more prosaically: how big a brain do you really need to eat a fruit? Since natural selection would not normally be so profligate as to waste resources on traits as energetically expensive as brains without good reason, the implication is that large-brained species must be doing something unusual if the evolutionary costs and benefits are going to balance. Four decades after Jerison (1977), however, there is still no consensus, with opposing views dominated by two main camps: those who argue for the pre-eminence of food-finding as the main driver of fitness (with a focus on foraging decisions and individual-level selection) and those who argue for the significance of group-living and the cognitive demands of sociality (with an emphasis on multilevel selection). The literature has become so littered with seemingly contradictory claims (Dunbar & Shultz, 2017) that some have even been led to suggest that it is not possible to draw any meaningful conclusions (Powell, Isler & Barton, 2017; Logan *et al.*, 2018; Wartel, Lindenfors & Lind, 2019; Hooper, Brett & Thornton, 2022).

We suggest that this impasse is mainly a consequence of the fact that many analyses fall foul of a series of conceptual and statistical traps, some of which are well-known logical fallacies. Most of these sources of error seem to arise because of a failure to appreciate that biology is a systems-based discipline. This can result in misinterpretations of statistical results, usually because the hypothesis actually being tested is often not the one we think we are testing. We identify four common classes of error: conceptual issues, hypothesis-testing issues, errors created by the choice of proxies used to test hypotheses, and problematic statistical analyses. We argue that all of these errors are easily resolved, and that

the resulting evolutionary picture is richer and, from a biological perspective, reassuringly more complex (given that the biological world is complex). Although we focus on mammalian (and explicitly primate) brain evolution, we suggest that these issues are a cautionary tale that applies right across the broad spectrum of comparative biology.

II. CONCEPTUAL CONFOUNDS

In a seminal paper, Tinbergen (1963) pointed out that explanations of biological phenomena naturally partition into a set of four mutually exclusive conceptual categories, or explanatory levels. These are usually identified as function (the way a trait maximises fitness), mechanisms (the complex of anatomical, behavioural, cognitive and physiological processes that allow the trait to maximise fitness – in other words, an adaptation), ontogeny (the developmental processes involving genetic, environmental and learning effects that give rise to the trait in the adult organism) and phylogeny (the sequence whereby a trait in a living species evolved from an ancestor that lacked it). Although Tinbergen originally referred to them all as *Why?* questions (four different ways in which a biologist might answer the question ‘Why is X the case?’), we might think of them as answering four different types of question: why, how, what and when, respectively.

This way of viewing biology has a long history. It was first enunciated by Aristotle nearly two and a half millennia ago, although, not knowing anything about evolution, he only identified the first three. [Besides being a philosopher, Aristotle was an exceptional hands-on biologist who pre-empted many later findings of modern evolutionary ecology and life-history theory (Dunbar, 1993b).] His insight was reinforced in the mid-twentieth century by evolutionary biologists of the stature of Julian Huxley (1942), one of the founding fathers of the New Darwinian Synthesis (the synthesis of Darwin’s theory of natural selection with Mendel’s mechanism of inheritance). However, it was Tinbergen in his classic 1963 paper that established the case for the four questions that are now recognised.

Tinbergen’s central point (following Aristotle) was that the four questions are logically independent of each other: the answer to one does not constrain, and is not constrained by, the answer to any of the others. Even though a full explanation will necessarily provide answers to all four questions, the fact that these questions are logically independent of each other means that we can ask and answer them piecemeal: we are not obliged to address all of them at the same time. Nonetheless, being able to show both that a trait has a function and that there is a convincing mechanism to support that function strengthens any claim we might make – much as fitting all the pieces of a jigsaw together creates a coherent, integrated picture.

Unfortunately, some commentators have been confused by the fact that these questions can sometimes intersect: a mechanism, for example, can have its own internal function, as when the provision of energy to the brain (a mechanism) in order to make some fitness-maximising

benefit possible (a function) may in turn create its own built-in functional explanation in terms of how energy flow around the body is optimised by a venous branching system. This is not a conceptual inadequacy as has sometimes mistakenly been claimed, but simply reflects the natural hierarchical structure of biological explanations (Dunbar, 1983). The distinction is about *types* of explanation, not about biological categories. The four types of explanation apply at each level in the hierarchy of biological explanations from genetics up through anatomy and physiology to behaviour and ecology.

Although these four questions have long provided a central framework for biology, far too many analyses conflate Tinbergen Questions by pitting against each other explanations that, in fact, answer different questions (see also Hooper *et al.*, 2022). We identify two sources of error under this heading. The first involves testing between explanations that properly belong to different Tinbergen Questions, a logical fallacy that philosophers of science refer to as a category error. The second is a derivative, but distinct, type of error that we refer to as the fallacy of the missing middle.

(1) Conflating Tinbergen's Questions

Many comparative analyses seek to test between alternative hypotheses for the selection pressures that have acted on a trait. Unfortunately, a worrying number have done this

by comparing a functional hypothesis with either a cognitive hypothesis or a constraint (both mechanisms issues). Many analyses, for example, have sought to test whether sociality (conventionally indexed by group size) or foraging skills (usually indexed by some aspect of diet, but occasionally by a cognitive index) have selected for the evolution of large brain size.

Consider the simple example involving brain size, group size and diet. Three variables can be causally related to each other in any of 18 different ways (six permutations with three different causal relationships in each case: positive, negative and no correlation). Figure 1 illustrates six of the more obvious ones, with a narrative explanation for the causal relationships in each case. Diet, for example, might be (i) a selection pressure for a large brain (larger brains are needed to allow animals to forage more efficiently, with a direct fitness consequence in terms of fertility), (ii) a cost of (i.e. constraint on) brain size (a species can evolve a large brain under selection only if it can solve the nutrient throughput problem so as to have sufficient spare energy to fuel the additional brain growth) or (iii) the lucky by-product of having a large brain for some other reason (otherwise known as an exaptation, or window of evolutionary opportunity: once you have a big brain, it can be used for many other purposes such as smart foraging that may provide additional subsidiary fitness benefits). Each of these possibilities identifies a very different

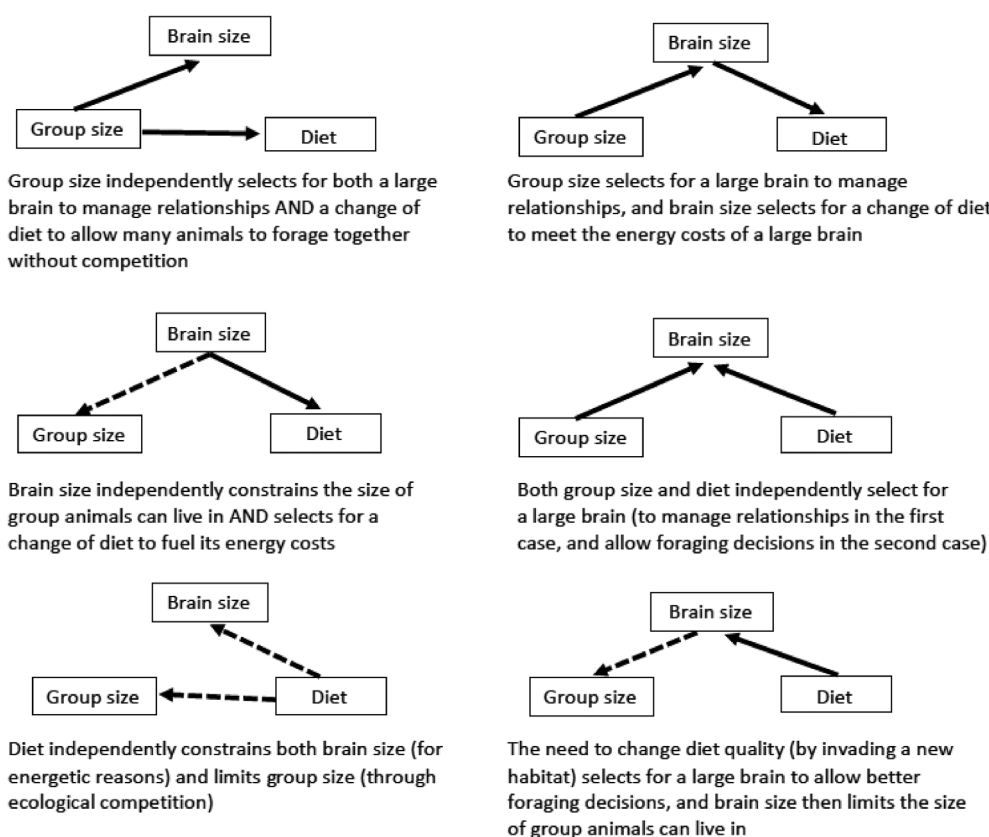


Fig. 1. Alternative causal models for three variables of interest. Solid lines indicate direction of a selection pressure; dashed lines indicate direction of a constraint.

causal pathway between the variables. In addition, a window of evolutionary opportunity can sometimes completely take over a trait, redesigning it for new purposes (as sexual selection often does).

Failure to keep Tinbergen's Questions properly segregated results in a tendency to adopt a psychological (or mechanisms) mindset rather than a biological (or systems-based) one. This causes us to ask simple mechanisms questions: given that A and B are potential causes of C, which one is the more important? But if B is a constraint (mechanism), and not an evolutionary cause (function) (e.g. energy intake imposes a constraint on growing a large brain, even when group size selects for a larger brain), this framing will be very misleading. In this case, the correct formulation should have the form: A causes C, which in turn causes (= requires) B to change so as to make this possible. The contrast is obvious when we specify the structure of the explanation in symbolic logic form. In formal notation, the mechanisms (psychological) version has the form:

$$\{(A \Rightarrow C) \vee (B \Rightarrow C)\} \quad (1)$$

where the operator \vee indicates disjunction ('either...or.... but not both') and the arrow specifies causality: either causal path (A determines C) is true or causal path (B determines C) is true, but not both. By contrast, the biological version might have the form:

$$\{(A \Rightarrow B) \Rightarrow (B \Leftrightarrow C)\} \quad (2)$$

A selects for a change in B, which causes B to select for a change in C, where C is a cost that has to be adjusted at the same time in order for A to be able to move B up the selection gradient (with the doubleheaded arrow signifying that B can only change if, and only if, C changes with it). Or, if C is a window of evolutionary opportunity:

$$\{(A \Rightarrow B) \Rightarrow C\} \quad (3)$$

(once B is in place, it provides an opportunity for C to appear at a later time). In other words, pitching a social explanation against an ecological one risks misconstruing the underlying biological causality.

This confusion seems to arise because many analyses assume that the social outcome is an evolutionary end in itself, comparable to food-finding. It then seems obvious to compare the role of group size and diet directly as determinants of brain size, as a great many recent studies have done (DeCasien, Williams & Higham, 2017; Powell *et al.*, 2017; Hardie & Cooney, 2023). The problem should be obvious: group-living is not an end in itself, but a stepping stone to an (ecological) end (see Dunbar, 1998b). The actual contrast here is not between a social *versus* an ecological explanation, but between two alternative ecological explanations, namely individual *versus* social (i.e. group-based) ways of solving the same ecological problem (be that food-finding, avoiding predators or combatting ecological rivals). The first implies that

animals deal with the challenges of survival and reproduction largely on the basis of individual trial-and-error, while the second implies that animals solve these problems through a group-level mechanism (i.e. group-level efficiencies, the emergent properties of groups, or cultural transmission). The group-level mechanism implies that there is an intervening *behavioural* step between the individual's cognition (brain) and the fitness outcome – namely, the demands of maintaining the stability and coherence of large social groups, which in turn selects for the cognitive skills needed to achieve this. It is important, by the way, to be clear that this is *not* a case of group selection, but rather one of group-level, or group *augmentation* (Kokko, Johnstone & Clutton-Brock, 2001; Kingma *et al.*, 2014) or multilevel, selection, a standard form of Darwinian selection. Analogous conceptual mistakes were made by van der Bijl & Kolm (2016) who wanted to test between group size and predation risk as drivers of brain evolution, and by Ashton, Kennedy & Radford (2020) who wanted to test between food-finding and inter-group conflict as drivers of cognitive (i.e. brain) evolution.

In fact, the question we should be asking is: which component of fitness is the most limiting for the animals? Biologists tend to assume, largely as a matter of convention, that the limiting factor is always energy throughput: surplus energy over and above that needed to sustain life is what determines fertility. This naturally predisposes us to assume that species whose brains allow them to forage more efficiently will be more likely to achieve higher fitness. This probably is broadly true for small-bodied species. It is not, however, necessarily the case for large-bodied species. For large-bodied slowly reproducing species, predation is often a more serious problem than food-finding: it can prevent species from occupying habitats where they would otherwise be under no nutrient constraint (Shultz *et al.*, 2004; Dunbar, Korstens & Lehmann, 2009; Lehmann & Dunbar, 2009; Shultz & Finlayson, 2010; Bettridge, Lehmann & Dunbar, 2010). There is, for example, no ecological (i.e. foraging) reason why chimpanzees (genus *Pan*) could not live in the forested region south of the Congo River, but they do not. That they do not seems to be because, unusually, both lion and leopard occur there; apes can apparently cope with either one of these predators, but not with both at the same time (Lehmann & Dunbar, 2009).

The point may be clearer if we think of this in terms of the classic life-history equation:

$$LRS = \sum_{x=0} (l_x b_x) \quad (4)$$

where LRS (lifetime reproductive success) is a proxy for fitness, l_x is age-specific survivorship (the probability of surviving from birth to age x) and b_x is the age-specific fertility (same-sex birth rate per annum), with their annual product summed over a lifetime. Since investment in survival and growth is necessarily inversely related to fertility, animals can maximise lifetime output by emphasising either survival or fertility. This is where the classic r - K selection trade-off comes from: species that give greater weight to the first term

(*K*-selected) emphasise survival at the expense of fertility (humans and apes with their slow life histories), whereas those that give greater weight to the second (*r*-selected) emphasise fertility at the expense of survival (many rodents with 'fast' life histories). Predation often plays the key role in tipping the balance between the two (Reznick & Endler, 1982; Charnov, 1993; Charnov & Berrigan, 1993). In effect, those who champion ecological explanations implicitly prioritise fertility as the fitness-limiting factor in animals' lives, while those who champion social explanations prioritise survival. Both obviously affect fitness, but it is an empirical, not a theoretical, question as to which is actually the more limiting in any given case.

Although the balance between survival and fertility as the two main components of fitness may well vary among taxonomic groups, evidence collated by Clutton-Brock (1988) suggests that, at least for larger-bodied species like primates, individual differences in longevity (i.e. survival) have a consistently bigger effect on fitness than individual differences in fertility. This conclusion is bolstered by findings from the empirically derived time budget models that have been built for a dozen primate and ungulate genera. These models [which are more accurate than conventional climate envelope models in predicting the biogeographical distributions of individual genera (Dunbar *et al.*, 2009, 2019; Willems & Hill, 2009; Korstjens, Lehmann & Dunbar, 2018)] indicate that, other than at the edges of their ranges where populations will always be under significant ecological stress, the majority of populations could live in much larger groups than they actually do. More importantly, perhaps, evidence from population studies of taxa as diverse as humans (Stein & Susser, 1975; Arends *et al.*, 2012), primates (Gesquiere *et al.*, 2018), mongoose (Creel *et al.*, 2013) and ungulates (Clutton-Brock, Guinness & Albon, 1983; Albon, Mitchell & Staines, 1983) provide clear evidence that food shortage only starts to impact on fertility once loss of body mass exceeds ~15% – in effect, starvation conditions (Dunbar & Shultz, 2021b). In fact, one reason for evolving large body size is precisely to capitalise on the metabolic savings of scale offered by Kleiber's Law (Kleiber, 1961). This buffers large-bodied animals against starvation and allows them to survive unpredictable periods of food shortage in a way that small-bodied species cannot. Many small-bodied mammals and birds, by contrast, starve to death overnight if they do not eat the equivalent of a significant proportion of their own body mass in food each day (Peters, 1986; Hatchwell *et al.*, 2009).

The substantive issue is that large brains do not come for free: brain tissue is unusually expensive compared to all other somatic tissue [the Expensive Tissue Hypothesis (Mink, Blumenshine & Adams, 1981; Tsuboi *et al.*, 2015; Liao *et al.*, 2016)]. Whatever else may be the case, species that need to evolve large brains will need to solve the mechanisms problem of how to make sufficient spare energy available to fuel a larger brain. Even if it might sometimes act as a selection factor, diet will always act as a brake, or constraint, on brain evolution. Indeed, we know this to be the case from

population-level developmental studies (Isaacs *et al.*, 2010; Staff *et al.*, 2012; Prado & Dewey, 2014). Given this, a species will always resist increasing its brain size beyond what is immediately necessary because doing so incurs energetic costs. Remove the benefit of having a large brain and there will be selection pressure (commensurate with the energetic cost of neural tissue) to reduce brain size. This seems to have happened several times during ungulate and carnivore evolution, but is extremely rare in primate evolution where sociality and brain size seem to be in a very tight co-evolutionary ratchet (Pérez-Barbería, Shultz & Dunbar, 2007). Montgomery *et al.* (2010), for example, found that although the callitrichids and *Microcebus* have both undergone marked reduction in body mass (dwarfism), brain size has been conserved.

Before we proceed to test between our hypotheses, then, we first need to be clear about the logical status of the variables we propose to include because their position in the evolutionary equation will be very different depending on their biological role.

Cognition is a more worrying case, because not only is it quite uncontroversially a mechanisms issue (it is about *how* animals make decisions, not *why* they make them), its inclusion as a variable in a comparative analysis is in danger of committing one of the more insidious of all the logical fallacies: a circular argument. The brain *is* cognition, so, in effect, the hypothesis being tested is whether brain size predicts brain size. Not surprisingly, it does so rather well. A second issue is that it implicitly assumes that the cognitive mechanisms involved are dedicated modular or 'closed-loop' (i.e. 'domain-specific') processes that function *only* in a specific context.

Of course, some cognitive processes genuinely are domain-specific in this sense: the visual system or the subcortical mechanisms that manage physiological homeostasis are obvious examples. But the kinds of high-level cognition that underpin decision-making typically involve distributed systems (mainly, but not entirely, in the neocortex) and are often 'domain general' (i.e. are involved in finding solutions for many different kinds of task). The claim that anthropoid primates developed a unique cognitive capacity for a generalised form of rule-learning that allows rapid one-trial learning (the capacity to infer a rule from a single observation, in contrast to the long haul of associative trial-and-error learning) has been cogently made by Passingham & Wise (2012) and Passingham (2021). This does not, however, tell us which kinds of tasks initially selected for this ability, nor which other tasks are emergent properties. Choosing between two behavioural outcomes involves comparisons that are independent of the specific task involved, and necessarily so because no two social or foraging situations are ever identical. A social decision (choosing between two grooming partners, or whether or not to threaten someone) involves exactly the same reasoning processes as choosing which of two food items to eat, or whether this branch or that one would make a better base for building a night nest. The misperception that cognition relates mainly to food is largely due to the fact that most of the indices of cognition developed for use in experiments use food as a reward purely for practical

convenience; at the same time, the inevitable constraints imposed by laboratory environments mean that the tasks involved are not always especially ecologically relevant.

By pitting a functional explanation against a mechanisms explanation, we are, in effect, asking whether a mechanisms explanation is more important than a functional one – a question that is, as Tinbergen reminded us, meaningless. Every biological phenomenon needs *both* a function *and* a mechanism to underpin that function: you cannot have one without the other. The mistake lies in a failure to parse correctly the causal relationships between the variables being tested.

(2) The fallacy of the missing middle

In the previous subsection, we pointed out that incorporating cognition into an analysis can result in a circular argument. The reason has to do with a derivative problem, the fallacy of the missing middle. A number of studies (e.g. MacLean *et al.*, 2014; Stevens, 2014; Benson-Amram *et al.*, 2016) have asked whether inhibition (or temporal discounting) as a putative index of foraging skills is a better predictor of brain size than an index of sociality such as group size. On finding that it is (and that the influence of group size in a multivariate regression is not significant), the obvious temptation has been to conclude that this is evidence that foraging skills have selected for large brains, and hence that group size is irrelevant.

What these analyses overlook is that, while inhibition may well play a role in the context of food choice decisions, it is also essential (and perhaps more so) for the existence of stable, bonded social groups – for two very good reasons. First, the stability of these groups depends on individuals being able to resist acting in ways that might destabilise relationships (e.g. by unnecessarily escalating agonistic encounters to the point where an opponent decides to leave the group, thereby causing the aggressor to lose the size-dependent benefits of the group). Second, and perhaps more importantly, animals need to be able to resist continuing to feed when others want to rest (or to rest when others want to continue foraging) – otherwise groups will very quickly break up and disperse (King & Cowlishaw, 2009), as happens in herding ungulates that do not have bonded groups (Ruckstuhl & Kokko, 2002; Ruckstuhl & Neuhaus, 2002; Calhim, Shi & Dunbar, 2006; Dunbar & Shi, 2008). When animals differ in the rate of gut-fill, some will inevitably need to go to rest in order to clear the gut while those that have only half-filled their stomachs will want to carry on feeding. Since the latter will drift away as they continue to feed, the group will inevitably fragment. Being able to resist this temptation requires the capacity to inhibit prepotent actions (self-control). More importantly, the opportunity costs incurred in these social contexts (predation) are much higher than that incurred by deciding not to pick one fruit in order to pick another that is some distance away.

In the light of Tinbergen's Questions, the obvious question we should be asking is whether inhibition (a form of cognition) is the intervening (mechanisms) variable between the

independent (brain size) and dependent (group size or diet) variables rather than itself being a selection factor. Mediation analysis allows us to test for this. To illustrate this, we combined the data from two widely used inhibition indices [a Go/No-go task from Stevens (2014) and an A-not-B task from MacLean *et al.* (2014); the tasks correlate well ($r = 0.681$, $N = 10$, $P = 0.03$), and appear to index the same underlying cognitive ability]. To do this, we converted the scores in each data set to standard deviates from their respective mean values, averaging the standard deviate scores where a species was sampled on both tasks. We then ran separate mediation analyses with brain size (indexed as endocranial volume, ECV) as the predictor variable, inhibition as the mediator and either group size (Fig. 2A) or diet (indexed as the percentage of fruit in the diet) (Fig. 2B) as the dependent variable.

If we follow what most analyses have done (a simple multiple regression with brain size as the outcome variable), we find, like them, that inhibitory control is a much stronger predictor of ECV ($\beta = 0.669$, $P = 0.005$) than group size ($\beta = 0.174$, $P = 0.341$). However, a mediation analysis with the correct

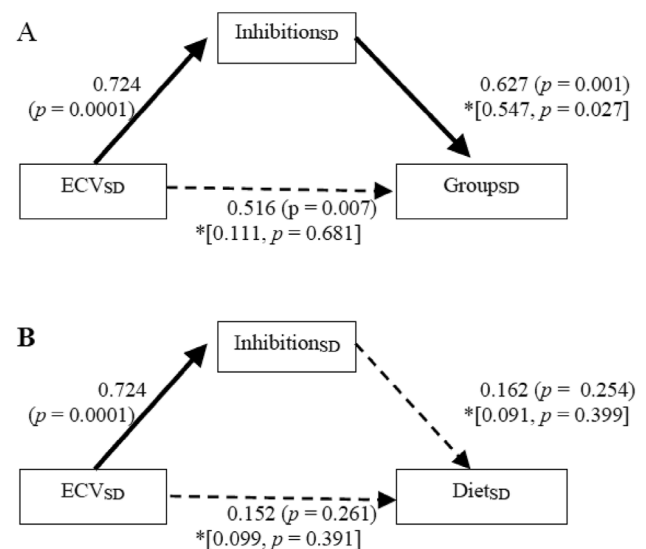


Fig. 2. Mediation analysis of the influence of brain size (indexed as endocranial volume, ECV) and inhibitory capacity (averaging two different inhibition/self-control tasks: Go/No-go and A-not-B) on (A) mean species group size and (B) diet (indexed as percentage fruit in the diet). All values are standardised deviates (calculated separately for each inhibition task before averaging). Values against arrows are standardised β , and their associated p values. Solid arrows: significant relationships; dashed arrows: non-significant relationships. Asterisks indicate values of β (and associated p value) for ECV and Inhibition as predictors in a multivariate regression. The marked difference between the bivariate regression and the multivariate regression indicates that the relationship between brain size and group size is indirect *via* the mediator Inhibition. Data for ECV, inhibitory capacity and diet are from MacLean *et al.* (2014) and Stevens (2014), and for group size are from Dunbar *et al.* (2018b).

outcome variable (i.e. group size or diet) yields a significant indirect relationship between brain size and group size *via* inhibition (Fig. 2A: Sobel test, $z = 2.12$, $P = 0.034$), whereas that for diet is not in fact significant (Fig. 2B: Sobel test, $z = 0.457$, $P = 0.648$). The log-likelihood ratio favouring the first over the second is $\lambda = 15.59$ ($P \ll 0.0001$). In sum, brain size determines inhibitory capacity, and inhibitory capacity then determines group size, but neither of them influences diet.

That inhibition is more intimately related to grouping variables than to foraging variables has also been shown using factor analysis for these same indices, using two variables that influence group cohesion during foraging (group size and day journey length: groups are demonstrably more likely to fragment when they are larger and have to travel further) and two variables associated directly with food-finding (diet and territory size) (Dunbar & Shultz, 2022). (In this context, territory size influences both the number of food patches available to the animals and their capacity to exclude rival groups, but not which food patches to visit.) Separate factor analyses for the MacLean *et al.* (2014) and Stevens (2014) indices yield identical results: in each case, the inhibition index clusters with the two social cohesion variables and not with the foraging variables (Fig. 3).

III. THE MISMEASURES OF FITNESS

Ultimately, evolutionary explanations are about fitness (the success with which individual genes are propagated down the generations) and how this is maximised by the adaptations to which they give rise (Dunbar, 1982, 2019). Two related issues arise in this context. One concerns how evolutionary biologists test for adaptations (Dobzhansky's Dictum); the other is how we measure the fitness associated with these processes.

(1) Dobzhansky's Dictum

Dobzhansky (1973) famously distinguished between two equally valid methods that biologists use to test hypotheses about adaptation: by testing for *being adapted* (fitness of design

as a consequence of selection that acted in the past) and by testing for *becoming adapted* (selection observed in action in the present). When we test a 'becoming adapted' hypothesis, we test a direct causal relationship in the here-and-now (e.g. males with longer tails or bigger antlers mate with more females). By contrast, most comparative analyses use species' mean values on traits, and therefore test a 'being adapted' hypothesis: they test the *outcome* of adaptation (the product of a selection pressure that acted in the past), not the *process* of becoming adapted. These are not the same. In the historical past, the need for more efficient foraging (perhaps as habitat quality deteriorated) might have selected for a larger brain that enabled smarter foraging skills, but in the here-and-now this is expressed as a species' foraging skills being constrained by the size of the brain it currently has, not by the effect that its foraging skills have on the size of its brain in the here-and-now. In effect, the causal logic, and hence the hypothesis we test, is reversed in the two cases (Fig. 4).

Consider the relationship between brain size, diet and group size that we discussed in the previous section. Some recent studies have regressed brain size on both group size and diet, taking body mass into account, with a view to determining which was the more important determinant of brain size. Finding that diet predicts brain size, whereas group size does not, they have concluded that it was the cognition underpinning diet choice (i.e. foraging decisions) that selected for large brains. But if we re-run the analysis the other way around with group size as the dependent variable, we get a *very* different answer (Table 1). With brain size (ECV) as the dependent variable, we find that body size, diet (indexed as percentage fruit in the diet) and group size are all significant predictors. But, with group size as the dependent variable, *only* brain size is a significant predictor; neither body size nor diet play a role.

The problem is that, when we choose brain size as the outcome measure – as, for example, was done by DeCasien *et al.* (2017), Powell *et al.* (2017) and Hardie & Cooney (2023) – we inadvertently interpret a 'being adapted' analysis as though it was a 'becoming adapted' one. But it should be obvious that it is biologically (and psychologically) implausible that group size could ever constrain brain size, if only because

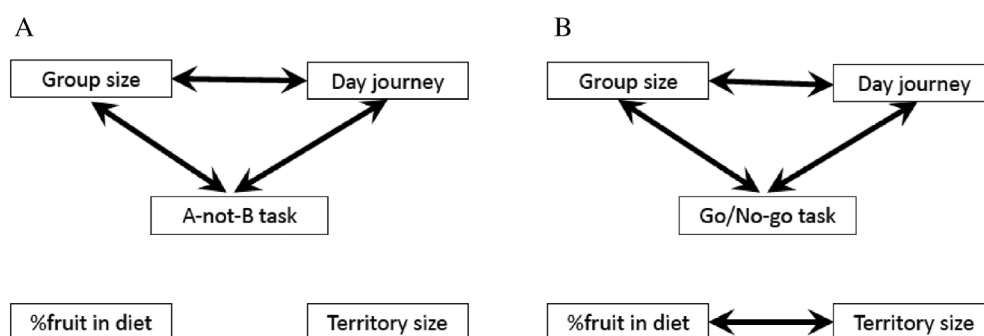


Fig. 3. Factor analysis of relationships between two different inhibition indices [A: A-not-B task of MacLean *et al.* (2014); B: Go/No-go task of Stevens (2014)], two variables known to influence group cohesion during foraging (group size and day journey length) and two variables that reflect foraging efficiency (percentage fruit in the diet and territory size). $N = 18$ species (A) and $N = 13$ species (B). Arrows link variables that cluster on the same factor.

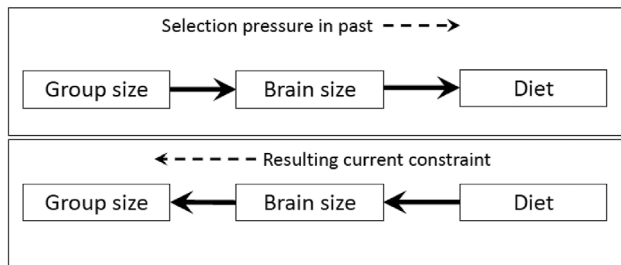


Fig. 4. When we see selection in action ('becoming adapted'), we observe a cascade of direct cause–effect relationships like that shown in the upper diagram. However, when the selection occurred in the past, in the here-and-now we only see the outcome of the selection process (the state of 'being adapted'). In this case, what we observe is a reversed causality in which the effect acts as a constraint on the cause (the counter-selection effect that the object of selection imposes as the cost of selection).

Table 1. Statistical analysis of the DeCasien *et al.* (2017) data on primate brain size evolution. Note that DeCasien *et al.* (2017) claimed that \log_{10} group size is not a significant predictor of \log_{10} brain size ($P = 0.74$). In three separate analyses of their data, with and without phylogenetic controls, we have not been able to replicate this result: however we run the analysis, group size is always a *very* significant ($P = 0.001$) predictor of brain size, and is actually slightly more significant than diet.

Predictor variables	Dependent variable			
	\log_{10} ECV*		\log_{10} group size	
	β^{\dagger}	P	β^{\dagger}	P
\log_{10} ECV			1.081	0.001
\log_{10} body mass	0.869	<0.001	−0.507	0.097
% fruit in diet	0.096	0.001	−0.001	0.992
\log_{10} group size	0.113	0.001		

*Endocranial volume (mm^3).

\dagger Standardised slope.

an individual's brain size is determined soon after birth whereas the size of group it lives in is determined by the environment it experiences as an adult. To treat this as a 'becoming adapted' process comes perilously close to assuming that causes can act backwards in time. In fact, the only conclusion that can legitimately be drawn from these results is the none-too-surprising one that group size does not, in the here-and-now, *constrain* brain size, implying that in the evolutionary past brain size did not select for group size. But what we really want to know is whether or not the need for a larger group size (to solve some ecological problem) imposed a selection pressure on brain size.

Unfortunately, it is all too easy to make inferential mistakes of this kind. To reinforce the point, we illustrate it with an example from another context that does not involve brains. Lukas & Huchard's (2014) wanted to test

whether monogamy evolves in order to minimise the risk of infanticide. To do this, they asked whether the rate of infanticide correlates with monogamy across species. When they found that infanticide was lower in monogamous species than in polygamous ones, they concluded that infanticide could not have selected for monogamy. But, in reality, their result provides strong direct support for the claim that infanticide *does* select for monogamy, at least in the special case of primates. Unfortunately, it seems they assumed they were testing a 'becoming adapted' hypothesis (individuals, or species, that exhibit a trait will gain higher fitness – paradoxically in this case, more infanticide) when actually they were testing a 'being adapted' one (if the trait is successful in its objectives, species that exhibit it more will incur *less* of the cost it is meant to counteract).

This highlights a confusion that seems to be disturbingly common in comparative analyses – a confusion between *risk* and *rate*. Infanticide rate (the actual observed frequency of infanticide events) is the level of risk (exposure to infanticide that a species faces, in the absence of any counterstrategies, in the environment of selection) that a species' behavioural adaptations have failed to control. It is *not* the level of selection it is under. Analysis of infanticide *risk* using van Schaik's (2000) risk index indicates that monogamous species typically experience a high *risk* of infanticide (because they have long interbirth intervals) but a low *rate* of infanticide as a result of adopting an effective counterstrategy (Opie *et al.*, 2013, 2014). The same problem arises in discussions of whether animals live in groups to reduce the risk of predation. In this case, predation *risk* (the likelihood of encountering a predator) should be higher in the kinds of habitats occupied by species with larger groups [as, indeed, it is (Hill & Lee, 1998; Hill & Dunbar, 1998; Dunbar, MacCarron & Robertson, 2018a; Dunbar & Shultz, 2021b)]. But, if living in large groups genuinely does protect individuals from predators, predation *rate* should be lower in large groups than in small ones in these habitats [as indeed it is (Shultz *et al.*, 2004; Shultz & Finlayson, 2010; Bettridge *et al.*, 2010)].

(2) Fitness top-down and bottom-up

The importance of investigating 'becoming adapted' hypotheses at the level of the individual has rightly been highlighted (Logan *et al.*, 2018; Hooper *et al.*, 2022). That is how evolution works. Being able to show both that putative benefits accrue at the level of the individual and how these are enabled by the appropriately designed behavioural or cognitive mechanisms provides important evidence that the traits of interest really are adaptations whose evolution has been driven by selection.

As it happens, the Social Brain Hypothesis is particularly well supported in terms of evidence for direct fitness outcomes. There is considerable evidence, for example from longitudinal studies of known individuals, that socially well-embedded females recover faster from injuries, have lower physiological stress responses to disruptive events, live longer, have higher fertility and have more offspring that survive to

adulthood than socially less successful females [chimpanzees *Pan troglodytes* (Wittig *et al.*, 2016); baboons (genus *Papio*) (Silk, Alberts & Altmann, 2003; Silk *et al.*, 2009, 2010; Wittig *et al.*, 2008; Archie, Altmann & Alberts, 2012; Cheney, Silk & Seyfarth, 2016); macaques (genus *Macaca*) (McFarland & Majolo, 2013; Young *et al.*, 2014; Lehmann, Majolo & McFarland, 2016; Brent, Ruiz-Lambides & Platt, 2017); wild horses *Equus caballus* (Linklater *et al.*, 1999; Cameron, Setsaas & Linklater, 2009; Nuñez, Adelman & Rubenstein, 2015); dolphins *Tursiops australis* (Frère *et al.*, 2010)]. Although they have seldom been considered in this respect, there is even some evidence for a similar effect in males: socially well-embedded male orcas (*Orcinus orca*) experience significantly lower mortality than less social males (Ellis *et al.*, 2017).

Comparable effects have been reported for humans. The last decade or so has witnessed a veritable deluge of very large scale correlational as well as prospective epidemiological studies showing that the number and quality of close friendships is the single best predictor of mental health and well-being, physical health and wellbeing, and even future longevity (among many other examples, see Holt-Lunstad, Smith & Layton, 2010; Rosenquist, Fowler & Christakis, 2011; Cruwys *et al.*, 2013; van Harmelen *et al.*, 2016; Yang *et al.*, 2016; Kim *et al.*, 2016; Cundiff & Matthews, 2018; Santini *et al.*, 2021; for biochemical-level reasons why this might be so, see Dunbar, 2018).

The link between these outcomes and the Social Brain Hypothesis is direct, with mentalising (the ability to understand others' perspectives and intentions) as the key intervening variable. Upwards of two dozen neuroimaging studies (some with very large samples indeed) for both humans (Bickart *et al.*, 2011, 2012; Lewis *et al.*, 2011; Horváth *et al.*, 2011; Powell *et al.*, 2012; Kanai *et al.*, 2012; Von Der Heide, Vyas & Olson, 2014; Dziura & Thompson, 2014; Hampton *et al.*, 2016; Pillemer, Holtzer & Blumen, 2017; Kwak *et al.*, 2018; Noonan *et al.*, 2018; Spagna *et al.*, 2018; Kiesow *et al.*, 2020; Peer *et al.*, 2021; Hyon *et al.*, 2022) and Old World monkeys (Sallet *et al.*, 2011; Meguerditchian *et al.*, 2020; Testard *et al.*, 2022) provide evidence that individual differences in the volume of specific brain regions [notably those known to be associated with mentalising skills in both macaques (Roumazeilles *et al.*, 2021) and humans (Carrington & Bailey, 2009; van Overwalle, 2009; Andrews-Hanna *et al.*, 2010)] correlate with individual-level differences in social network (or living-group) size (see also Bzdok & Dunbar, 2020, 2022). Powell *et al.* (2012) used path analysis with individual-level neuroimaging data to show that the causal sequence is: prefrontal cortex volume determines mentalising skills which, in turn, determine social network size.

Mentalising (or mindreading) is not the only cognitive ability that is important in a social context. There are now several large-scale prospective studies showing that individual differences in the capacity for self-control (behavioural inhibition) in childhood [a trait that is largely dependent on the brain's frontal pole (Brodman areas BA9/10); Passingham & Wise, 2012] strongly predict differences in adult social skills, relationship

stability and (negatively) the likelihood of being in trouble with the law (Robins & Ratcliff, 1978; Tremblay *et al.*, 1994; Moffitt *et al.*, 2001; Molero Samuelson *et al.*, 2010). These effects are specifically associated with structural differences in brain organisation (Carlisi *et al.*, 2020) as well as differences in specific genetic alleles (Moffitt *et al.*, 2001). In addition, across primate species, differences in brain (or brain region) volume correlate with a range of socially relevant cognitive skills, including the capacity for self-control (Amici, Aureli & Call, 2008; MacLean *et al.*, 2014; Stevens, 2014; Dunbar & Shultz, 2021a, 2022), mentalising (Hermann *et al.*, 2007; Krupenye *et al.*, 2016; Devaine *et al.*, 2017), tactical deception (Byrne & Corp, 2004), the ability to exploit coalitions to gain a fitness advantage (Pawlowski, Lowen & Dunbar, 1998) and the ability to reason inferentially (Dunbar, McAdam & O'Connell, 2005; Deaner *et al.*, 2007; Shultz & Dunbar, 2010c). Most of these cognitive processes are computationally very demanding (Dàvid-Barrett & Dunbar, 2013; Lewis *et al.*, 2017), offering a direct causal explanation for why socially sophisticated species need large brains.

By contrast, the evidence for direct fitness consequences of foraging skills is, at best, meagre. Many studies certainly provide evidence that primates engage in sophisticated ecological decision-making (Janson, 1990; Berghänel, Schülke & Ostner, 2015; Rosati, 2017). However, none of these provide evidence that species differences in foraging ability correlate with differences in brain size, or that individual differences in foraging skills have direct fitness consequences in terms of longevity or lifetime fecundity. We know of only one study that provides such evidence: Altmann's (1991, 1998) study of foraging skills in yearling baboons (*Papio cynocephalus*) and their consequences for longevity and lifetime fecundity. Although the data in this case are truly impressive and the correlation near-linear, the sample size is very small (just $N = 6$ females), there was no control for their social embeddedness (a factor whose impact on fitness only became apparent decades later: see Dunbar & Shultz, 2021a) and the population in question is on the ecological margin of the species' biogeographical distribution, and was undergoing demographic contraction due to deteriorating environmental conditions – precisely where one might expect foraging skills to have most influence. That the primary selection factors for group-living can switch from predation risk to food-finding as habitat quality deteriorates should not be a surprise: it has been shown in other primates (Dunbar, 1989). However, we cannot consider populations on the edge of a species' biogeographical range as being representative of the norm.

There is one final caveat we need to add. The fitness consequences of foraging and sociality arise at different levels in the system, and this can make them difficult to compare directly. Foraging skills can be measured directly at the level of the individual in terms of nutrient intake per unit time. The fitness consequences of some social skills can also be measured directly (e.g. how many matings a male with particular skills achieves), but the skills that influence group cohesion can only be measured as the sum of the social

competences of all the animals in the group over their lifetimes (in effect, neighbourhood-modulated fitness *sensu* Hamilton, 1964). In all obligately social species, the outcome measure is not whether an individual achieves an outcome, but whether the group is sufficiently well coordinated to maximise a collective benefit. This component of fitness is a function of the *average* fitness of the individuals concerned, not that of any one individual. There is a similar issue with cooperative breeding and pairbonded monogamy: ultimately, the success of the breeding pair lies not in their individual performances (even though those may be additively contributory) but in how well they cooperate in the complex business of reproduction.

IV. CRITICAL TESTS AND SLOPPY PROXIES

The physicist Isaac Newton famously defined a *critical test* as one whose outcome unequivocally discriminates between the hypotheses under test. In other words, the behavioural index we use to test between two hypotheses needs to predict an outcome in one direction if hypothesis A is true and in the *diametrically opposite* direction if the competing hypothesis B is true. This remains a benchmark of good experimental design, but it applies equally to statistical testing of hypotheses based on observational data. Far too many analyses fail on this account. We identify four potential traps under this heading, all of which result in a different hypothesis being tested to the one we think we are testing. These concern the design of critical tests, problems that arise when the variables used to test hypotheses are poorly defined, the common practice of relativising traits against body size (or anything else), and the tendency to over-generalise hypotheses (i.e. test an hypothesis on taxonomic groups or in contexts to which it does not, and was not intended to, apply).

(1) Critical tests

In a conceptually important paper, van Schaik (1983) identified lack of critical tests as a common problem in comparative analyses. Far too often, we use an outcome variable that does not discriminate between alternative predictor variables. In testing between predation risk and defence of food sources as the explanation for group-living in primates, for example, it does not make sense to use group size or fecundity (lifetime reproductive output) as the outcome measure, since both hypotheses predict that successful groups will be larger with more fecund females. What differentiates well-formed hypotheses, van Schaik argued, is the *mechanism* that makes the outcome possible in each case. In this example, one hypothesis (predation risk) identifies survival as the issue, the other (resource defence) identifies fertility (the more surplus energy acquired, the higher will be an individual's fertility). As a result, they make contrasting predictions about how within-group competition impacts females' fertility (their birth rates per year). The foraging hypothesis predicts

that fertility will increase with group size because a large group's ability to monopolise rich food sources offsets the fertility costs of competition (at least up to the point where within-group competition starts to overwhelm this benefit). By contrast, the predation risk hypothesis predicts that fertility will decline linearly with group size because the hypothesis offers no antidote to the insidious effects of within-group competition (Fig. 5). The result is two patterns that differ from each other, and which, at least within the range of the boxed region on the left side of Fig. 5, form a critical test (they make diametrically opposite predictions). Alternatively, we might test whether fertility has a negative or a quadratic form. van Schaik's (1983) original analyses, and subsequent tests by Dunbar (1988) and Dunbar & Shultz (2021b), confirm that the driver of group-living in primates is indeed predation risk, not resource defence. In short, effective tests between alternative hypotheses need to identify the right level of analysis.

The Social Brain Hypothesis builds directly on van Schaik's (1983) predation risk hypothesis as the principal driver for group-living (see Dunbar, 1998b; Dunbar & Shultz, 2007, 2017, 2021b), but identifies *social* stress created by living in close spatial proximity as the cause of declining fertility in larger groups (the 'infertility trap': Dunbar & Shultz, 2021b) rather than food access. van Schaik's view assumes that the fertility cost of large groups is simply a cost animals have to accept. However, the whole point of having

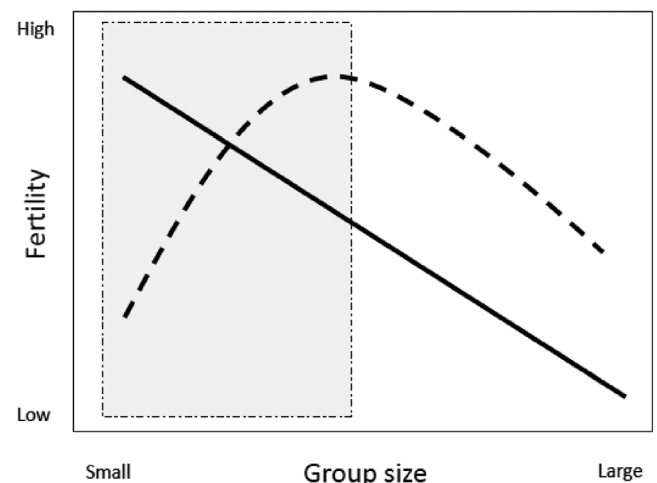


Fig. 5. Relationship between fertility and group size predicted by the two alternatives for group-living in primates: predation risk (solid line) and (between-group) resource defence (dashed line). In both cases, within-group competition for resources impacts negatively on fertility, but the benefits of between-group resource defence defer the effect until larger group sizes. As a result, the two hypotheses predict different relationships between the two variables. On the left side of the graph (boxed area), these predictions are in diametrically opposite directions, forming a classic critical test that unequivocally differentiates between the two hypotheses. Redrawn after van Schaik (1983).

a large brain according to the Social Brain Hypothesis is to be able to devise strategies to defuse these costs, thereby buying demographic space that will make it possible to occupy high-risk habitats (Dunbar, 1998b; Dunbar & Shultz, 2021b). In other words, the social brain is necessary not to create large groups *per se* but rather to buffer the females against the stresses of living in large groups by deploying cognitively expensive social strategies such as coalition formation and relationship management with their associated skills of diplomacy, understanding third-party relationships, and self-control (Dunbar, 1998b). The resource defence hypothesis (originally proposed by Wrangham, 1980) is, of course, also a social hypothesis, but it makes no assumptions about the cognitive demands of group-living. The cognitive demands of food-finding might provide an answer, but it leaves unanswered the question of why primates should be willing to incur such significant fertility costs by living in large groups – or, given that they in fact clearly do, how they avoid being overwhelmed by these costs. The merit of the Social Brain Hypothesis is that it provides a single unified explanation for all these elements.

We might note in conclusion that a comparison between biologically plausible alternatives lends itself to Bayesian statistical analysis. A Bayesian approach is always a more powerful form of hypothesis-testing than the conventional frequentist approach because it allows us not merely to reject the null hypothesis but also to assert that the evidence uncontroversially favours one hypothesis over the other. This is inherent in the conceptual design of Bayesian statistics: not only must the posterior probability for one hypothesis evolve across successive tests towards $p_{\text{posterior}[A]} > 0.95$ (close to certainty) but those for the alternative hypotheses must correspondingly tend towards $p_{\text{posterior}[B]} \approx 0.00$. More importantly, a Bayesian approach allows us to test between multiple hypotheses simultaneously: animals live in groups EITHER to manage predation risk OR to defend their territory against competitors OR to rear offspring cooperatively OR to forage more efficiently. We do not even need a null hypothesis, given that this is not likely to be either helpful or interesting. Most people's experience of Bayesian statistics is probably limited to their use as a more sophisticated form of parameter estimation in statistical packages. In fact, they are much more useful as a way of testing for goodness-of-fit to a theoretical prediction in reverse engineering designs [e.g. Hill & Dunbar (2003); Dunbar & Shultz (2021a)] or testing between alternative hypotheses (e.g. Dunbar, 1989).

(2) Sloppy proxies

When we test hypotheses, we rarely test the hypothesis as it is framed, unless it is an exceptionally low-level mechanisms hypothesis. Instead, we test a proxy for it based implicitly on the mechanism that underlies the proposed hypothesis (Altmann, 1974; Dunbar, 1976). This is because such hypotheses are usually high-level theoretical claims derived from overarching theory, and these cannot usually be tested directly. For example, evolutionary theory predicts that males who are

more successful in mating contests will have higher fitness. It is difficult to measure fitness directly, especially for long-lived species, because, formally, it requires data from a minimum of three successive generations to do so reliably (Dunbar, 1982). This being so, we usually test a derivative proxy that we believe is a correlate of fitness (e.g. males who win more fights will mate with more females or sire more progeny). These proxies are, however, often sloppy in the sense that they incorporate varying degrees of error variance, not just as measurement error but also in how uniquely the proxy correlates with the underlying theoretical concept.

Number of matings is an imperfect proxy for fitness because many other factors intervene between the two. As Lack (1954) reminded us, there is an important distinction between the number of offspring born and the number that actually survive and, in their turn, breed successfully: natural selection acts on the second, not on the first (Lack's Principle). Normally, we just hope that the error variance is not so great as to overwhelm what, due to all the intervening steps, can often become a weak causal relationship [a version of Grafen's (1991) *phenotypic gambit*]. Sometimes, however, the data are so sloppy that it becomes impossible to get anything but a non-significant result (a Type II error). Philosophers of science (Popper, 1962; Lakatos, 1980) remind us that, in such cases, we should not reject the hypothesis under test out of hand, but we should first ask whether our test has been a fair test – have we omitted some key variable? We will meet another example of this problem in Section V.2.

This problem frequently arises in contexts where one might least expect it. One of these is group size. Intuitively, we all think we understand what we mean by group size, but when we apply that definition to actual populations it can be subject to considerable slippage as we try to force what we see on the ground to fit our definition. Patterson *et al.* (2014), for example, noted that estimates of mean species group size in primates vary considerably across compilations, and questioned whether analyses that use these data had any real meaning. Others have noted comparable problems with how we classify species' social and mating systems, mainly because we fail to note the variability in what species actually do (Kappeler & Pozzi, 2019). However, none of these concerns are quite what they seem. To see why, we focus on Patterson *et al.*'s (2014) analysis of group size.

The first point to note is that, despite their concerns, the five largest samples in their data set (those with $N > 10$ species sampled) actually correlate significantly with each other (pairwise comparisons: mean $r = 0.756$, range $0.674 \leq r \leq 0.907$, $P \leq 0.030$), and all five correlate significantly with the most recent independent sample provided by Dunbar, MacCarron & Shultz (2018b) (Fig. 6: mean $r = 0.820$, all $P \leq 0.001$). There is certainly some variability, and this will undoubtedly introduce some error variance into any statistical analysis. But, since the estimates all broadly agree with each other, the effect will actually be modest: increased variance can only reduce statistical power and hence increase the risk of Type II errors (failing to reject the null hypothesis when it is in fact false).

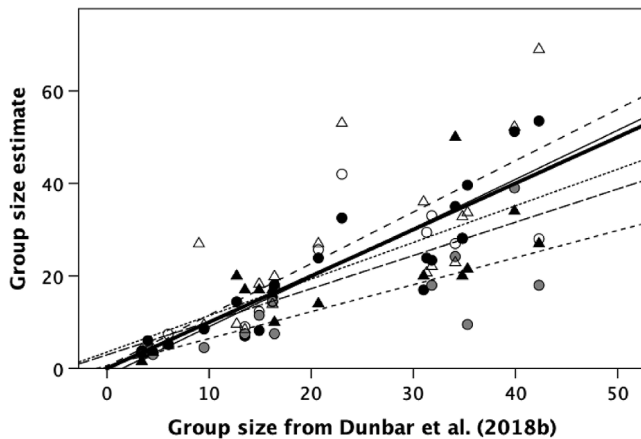


Fig. 6. Estimates of mean group size for 23 primate species from five different compilations, as given by Patterson *et al.* (2014), plotted against the equivalent values given by Dunbar *et al.* (2018b). The thick line is the line of equivalence where samples should match. Filled circles and thin solid line: Dunbar (1992); note that the regression line for these data is almost indistinguishable from the line of equality. Filled triangles and long dashed line: Clutton-Brock & Harvey (1977); open triangles and medium dashed line: Patterson *et al.* (2014); open circles and short dashed line: Chapman & Chapman (2000); grey circles and dotted line: Majolo *et al.* (2008). Samples with $N < 10$ species given by Patterson *et al.* (2014) are not included.

More importantly, however, their analysis confuses four separate issues.

One is that estimates of group size will always vary because of small sample bias effects. However, the statistical Law of Large Numbers guarantees that estimates will converge on the true mean as sample size increases with time. We could easily deal with this by setting a minimum research effort criterion for including a taxon in our sample – although that is bound to reduce sample size. The real issue here is the trade-off between the quantity and quality of data. Data quality is more important when you only have small samples. As more populations are sampled, however, the problem becomes less and less serious.

Second, they seem to take the view that group size is a species characteristic in the way that fur colour or the presence/absence of horns are. But group size is the outcome of the momentary decisions that animals make about the costs and benefits of living in groups of different size under particular environmental conditions, and the average describes exactly what it says: the time-weighted mean outcome of these decisions. This is why there is little or no phylogenetic signal in species group sizes in primates (Kamilar & Cooper, 2013). Group size is not the outcome of a simple genetic effect; it is the consequence of the interaction of many different environmental and psychological factors, as is true of much mammal (and probably bird) behaviour. We are not dealing with a simple one-cause/one-effect phenomenon here, but that does not make statistical analysis invalid. Error variance in data is precisely what statistics was designed to deal with.

The third issue is a consequence of the processes that underpin the demography of all species that, like primates, have bonded social groups. These groups ('congregations') cannot lose members by individual trickle emigration the moment their size exceeds some ecologically ideal value in the way that optimal foraging theory predicts for casual flocks and herds ('aggregations'). Bonded groups can only lose significant numbers of members by group fission, and fission is only possible when current group size is at least double the minimum size required for predator defence in that specific habitat (Dunbar *et al.*, 2009, 2018a; Dunbar & MacCarron, 2019). As a result, primate group dynamics takes the form of a non-linear oscillator: there is a target value set by the environment, and the group oscillates around this over a period of years as the group increases naturally in size (slowly at first, but increasingly fast as it gets larger and accumulates more breeding females), and then undergoes a precipitous crash when the group is finally able to fission (Fig. 7). Since fertility declines as groups get larger (Dunbar & Shultz, 2021b), the process of fission can take many months, sometimes years, because groups can become locked in a form of demographic stasis where births only just offset deaths (see Strier, Lee & Ives, 2014), unable either to increase in size or to undergo fission. Small-scale human groups have similar dynamics (Dunbar & Sosis, 2018). Group fission has been widely documented in primates, but is rare in the lifetime of any one group (probably occurring at intervals of 10 years or more). It should be no surprise that mean group size estimates vary, at least within a range, because they will depend on where in this cycle groups are sampled, and whether environmental conditions (especially predator density) favour the lower or higher end of the oscillator (see Dunbar *et al.*, 2018a; Dunbar & MacCarron, 2019).

The fourth issue is the most troubling, and this is the fact that many compilations suffer from definitional slippage, mixing foraging groups for some species with social groups for other species – another form of category error (see Section II.1). In many primate species, these are one and the same, but in a significant number of cases they are not. These latter species come in two varieties, those like chimpanzees, orang utans (*Pongo* spp.), spider monkeys (*Ateles* spp.) and many nocturnal prosimians that have stable communities but spend most of their time in small dispersed foraging parties (atomistic fission–fusion societies) and those like gelada (*Theropithecus gelada*) and hamadryas (*Papio hamadryas*) baboons or snubnosed (*Rhinopithecus* spp.) and proboscis (*Nasalis larvatus*) monkeys that live small stable harems (one-male groups) that cluster into larger groupings of variable stability during foraging (modular fission–fusion societies).

Orang utans, to take one example, are invariably listed in databases as being solitary because the animals are usually seen alone in most populations. However, most fieldworkers have noted that there appear to be distinct communities who know (and tolerate) each other (MacKinnon, 1974; Singleton & van Schaik, 2002). Indeed, where conditions allow, as in northern Sumatra, orangs may even forage in small groups (Sugardjito, Te Boekhorst & van Hooff, 1987). The

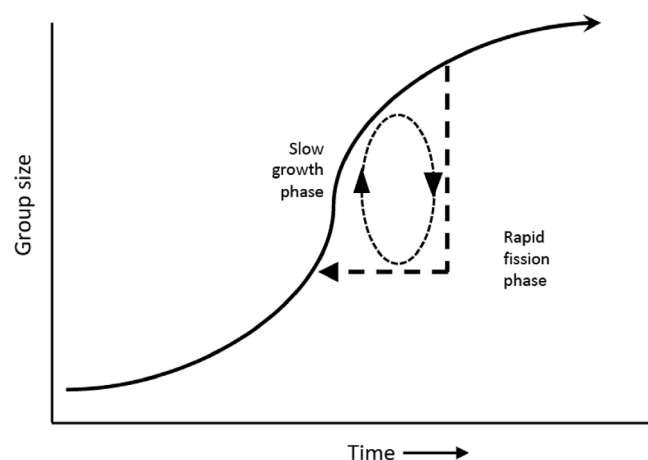


Fig. 7. The non-linear oscillator that describes the dynamic size trajectory of a typical primate group. The oscillator consists of two phases: (1) a long slow growth phase that follows a sigmoid trajectory of increasingly rapid early growth followed by a slow phase as the stresses due to increasing group size reduce female fertility, and hence growth rates, and (2) a catastrophic reduction in group size following group fission. A group will cycle continuously round the oscillator so long as there are no changes in environmental conditions.

average size of these communities is ~ 14 individuals. In fact, in captivity, oranges are at least as social as gorillas, and most zoos house them in groups for precisely this reason (Lardeux-Gilloux, 1997). The species is solitary now only because, thanks to climate warming, it lives in a marginal habitat at the limits of its ecological tolerances (Carne, Semple & Lehmann, 2015). Using a (social) group size of $N = 14$ places the species exactly where its neocortex size predicts, but using a (foraging) group size of $N = 1$ leaves it far adrift of all other species. That surely tells us something.

Much the same is true of the aye aye (*Daubentonia madagascariensis*) of Madagascar which, likewise, usually forages alone (and so is always listed as being solitary), but seems to live in local communities ('neighbourhoods') of up to eight individuals that share a home range and may even on occasion forage and nest together (Iwano, 1991; Ancrenaz, Lackman-Ancrenaz & Mundy, 1994; Sterling & McCreless, 2006). Using a group size of $N = 1$ makes the species a puzzling outlier on the Social Brain graph; using a group size of $N = 8$ places it where it might be expected to lie given its brain size. In this, aye aye resemble other 'semi-solitary' nocturnal lemurines and galagines who are now considered to live in social groups ('nest groups') (Bearder, 2008; Nekaris & Bearder, 2007).

A comparable problem arises in the case of species that have modular fission–fusion social systems. The multilevel social systems of gelada and hamadryas baboons have several layers that have consistently stable memberships (harems of 5–15 individuals, clans of 30–50, bands of 100–150). Since these can differ by an order of magnitude in size, choosing the wrong level will have a dramatic effect on any

hypothesis that is being tested. The problem, once again, lies not with the theory or the data, but with researchers' preconceptions about the animals' natural history.

This raises an important issue concerning the nature of sociality in primates (and the handful of other mammalian orders that have bonded relationships). The groups of these taxa are characterised by relationships that have considerable stability over time. The members of the group know each other well at a cognitive level, have stable long-term relationships, and are tolerant of each others' close physical proximity. These traits are all lacking in the more transient groupings of herd-forming species, where most relationships are of-the-moment and lack the personalised depth of bonded relationships; in effect, every interaction is with a stranger. Probably the closest we get to bonded relationships of the primate intensity in other mammals and birds are the pairbonded societies of canids (Macdonald *et al.*, 2019) and some miniature antelope (e.g. klipspringer, *Oreotragus oreotragus*; Dunbar & Dunbar, 1980), and the lifelong pairbonders among the birds.

In primates, these relationships are created and expressed through social grooming. This does not mean that everyone grooms, or has a bonded relationship, with everyone else in the group, especially in very large groups. In bonded social groups, individuals devote almost all their grooming to a very limited number of group members (Kudo & Dunbar, 2001; Dunbar, 2003, 2023). In humans, for example, 60% of total social effort (whether measured as time invested, frequency of contact or emotional closeness) is devoted to just 15 people (Sutcliffe *et al.*, 2012). What seems to hold the group together is a 'friends-of-friends' effect that links these grooming subgroups together into a grooming chain, creating a form of 'gravitational field' (Fig. 8). The result is a fractal structure to social groups, which, when seen from the individual's viewpoint, has a hierarchically inclusive layered structure with layers of very similar size across a wide variety of mammalian species including dolphins, elephants, cercopithecine monkeys, apes and humans (Hill & Dunbar, 2003; Wittemyer, Douglas-Hamilton & Getz, 2005; Hamilton *et al.*, 2007; Hill, Bentley & Dunbar, 2008; Zhou *et al.*, 2005; Waller, 2011; Moss, Croze & Lee, 2011; Wakefield, 2013; MacCarron & Dunbar, 2016; Escribano *et al.*, 2022). These size regularities derive from the mathematical properties of networks and the way animals choose to allocate their limited social time (Tamarit *et al.*, 2018; Tamarit, Sánchez & Cuesta, 2022; West *et al.*, 2020, 2023). All that animals need do is maintain visual (or even auditory) contact with their one or two closest grooming partners, and the more casual (weak) links between subnetworks are sufficient to maintain group cohesion (Castles *et al.*, 2014; Dunbar, 2023) – unless, of course, groups get very large and/or day journeys very long, in which case groups may fission down the fracture line created by the weak links between sub-networks (Dunbar & Shultz, 2022).

A final issue to consider is that, although analyses invariably focus on species mean group sizes, the Social Brain Hypothesis has always been conceptualised in terms of an upper limit on the size of group that can be maintained as a coherent, stable

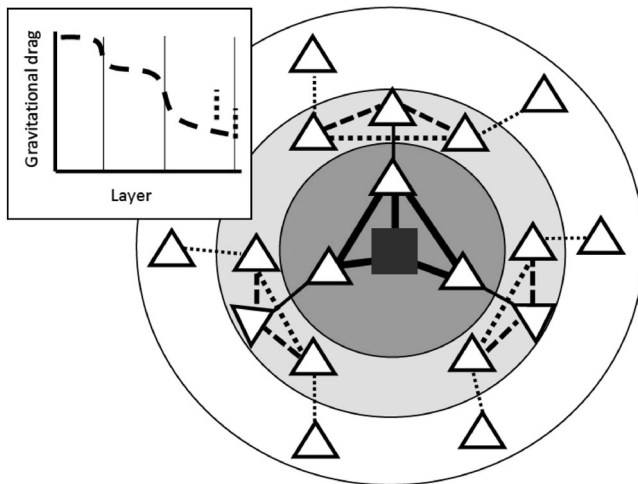


Fig. 8. The layered fractal structure characteristic of anthropoid primate (and human) egocentric social networks. Seen from the perspective of any given individual (the central square), grooming relationships with other individuals (triangles) spread out through a series of increasingly distant layers. The processes of bonding created by social grooming and friends-of-friends effects create a form of gravitational drag (inset) that holds a group together. Human egocentric social networks have exactly this form (Dunbar, 2020).

entity (Dunbar, 1998b). A species does not *have* to live in the largest group size its brain will allow; this simply sets the limit it can manage. Because of the non-linear oscillator (Fig. 7), the limiting size is not the maximum group size ever observed but the size at which groups start to become unstable. This value, however, is difficult to determine. Fortunately, primate group sizes are almost always Poisson-distributed (Dunbar *et al.*, 2018b), and Poisson distributions have the convenient property that the mean and variance are identical. This means that it should not matter too much which statistical moment (mean, variance, limiting size, maximum size) we use in an analysis, as these are all closely correlated, if not identical. Whichever index we use to test the Social Brain Hypothesis we get the same answer, as Sandel *et al.* (2016) showed.

The fact that group sizes are Poisson-distributed offers us a way to estimate where this limit might be. If we plot a species' group sizes as a cumulative distribution, the limiting group size is given by the upper inflection point where the cumulative distribution changes slope, since this demarcates the point of diminishing returns. Figure 9 plots the cumulative distribution for 376 woodland baboon (genus *Papio*) groups as an example. The inflection point can be estimated in either of two ways. On a sigmoid cumulative distribution, the theoretical inflection point is the value on the x -axis that corresponds to the point that is $1/e^{\text{th}}$ down from the asymptote on the y -axis (Slatkin & Hausfater, 1976). If we take the asymptote to be defined by the 360th group (since there are only a very small number of groups larger than this), the inflection point corresponds to the $360 * (1 - e^{-1}) = 227^{\text{th}}$ ranked group (the horizontal dotted line), and this has a

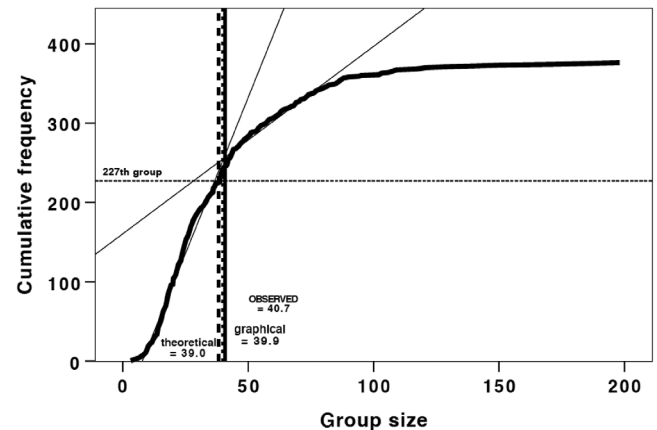


Fig. 9. Cumulative plot of group size for 376 groups from 40 populations of woodland baboons (*Papio anubis*, *P. cynocephalus* and *P. ursinus*). The observed mean group size for the three species is 40.7 (solid vertical line). Dashed vertical lines give estimates of the upper limit on group size defined by the inflection point in the cumulative distribution. The theoretical estimate (long-dashed vertical line) is obtained by finding the value on the x -axis that corresponds to the value on the y -axis that is $1/e^{\text{th}}$ down from the asymptotic value (taken to be at the 360th group), identified as the 227th group (horizontal dotted line). The graphical estimate (dotted vertical line) is the 'broken stick' estimate, defined by the intersection of the two regressions (thin lines) that yield the best fit when the x -axis is partitioned into two parts.

group size of 39.0 (long-dashed vertical line). Alternatively, we can determine the inflection point graphically using the classic broken stick method widely used in ecology (Magurran, 1988). We partition the x -axis serially into two parts and set regressions to each half, searching for the partition that maximises overall fit. The point where the two best-fit regression lines intersect defines the inflection point. The relevant regressions are shown as the thin lines fitted to each half of the distribution in Fig. 9. They cross over at a group size of 39.9 (thin dashed vertical line). The observed mean group size for the three species of woodland baboons is 40.7 (solid vertical line). These results suggest that the average group size, for this genus at least, is identical to its limiting group size. Once a group's size exceeds this value, it is straying into the region of its demographic state space where both social cohesion and fertility are rapidly declining. This value will, of course, vary across genera (Dunbar & Shultz, 2021b) as a consequence of the genus' ability to evolve behavioural and cognitive mechanisms for counteracting the stresses involved (Dunbar & Shultz, 2021a; Dunbar, 2023).

(3) The problem of relativity

Many comparative analyses automatically relativise traits of interest against body mass, or alternatively include body mass as a covariate in their statistical analysis (which, statistically speaking, comes to the same thing). There are good reasons for including body mass as a covariate in an analysis, such

as when we want to know whether a trait increases in size over evolutionary time merely because it scales with body size. Here, we ask whether trait size is a non-functional by-product of whatever is driving the change in body size, or whether body size causally determines, or is determined by, the trait in question. Heart size, for example, is highly correlated with body size because a bigger body needs a proportionately larger heart to pump blood around it. However, it seems that many comparative analyses feel compelled to include body mass in their analyses simply because Jerison (1977) did so. Aside from the fact that adding unnecessary extra variables inevitably reduces statistical power and is bad statistical practice, there are four separate issues here.

First, there seems to be a widespread misunderstanding (for an example, see Logan *et al.*, 2018) as to why Jerison (1977) originally calculated his Encephalisation Quotient (EQ, the residual of an individual species' brain from the overall regression line relating brain to body size across species). Jerison (1977) was not seeking to determine whether species had smaller or larger brains than we would expect for their body size; rather, as the title of his book indicated, he was trying to remove that part of the brain which is solely concerned with managing somatic tissue and other physiological processes (which is therefore likely to be isometrically scaled with body mass) in order to isolate out that part of the brain (in effect, the neocortex) that is available for higher cognitive functions (smart foraging, clever behaviour, etc.) when he only had total brain size (actually, cranial volume) available (Jerison, 1977; H.J. Jerison, personal communication). In practice, both EQ and ECV are poor estimates of the socially functional brain and hence yield only modest correlations with either 'smart' cognition (i.e. decision-making competences) (Dunbar *et al.*, 2005; Deaner *et al.*, 2007; Shultz & Dunbar, 2010c) or social group size (Dunbar, 1992; Dunbar & Shultz, 2017, 2021a; Shultz & Dunbar, 2022) compared to indices based on neocortex size itself.

Second, using body mass as a covariate (or as the base for a residual) unavoidably changes the question we are asking. When we take residuals for wing area or brain size regressed against body size (or include body size as a covariate), we are asking whether a species has a wing, or a brain, that is larger (or smaller) than we would expect for an animal of its body size. This is, of course, a perfectly legitimate question to ask: in the case of wing size, for example, we might be interested in whether the lift properties of a wing are proportional to the mass it has to lift, or whether (as in the case of basal metabolic rate, BMR) there are savings of scale that could be invested in other organs (Kleiber, 1961). In this case, we are *not* asking whether one species has an absolutely larger wing than other species, or what external (environmental) factors might have determined why it has a large wing. Asking whether a species has a brain bigger than expected for an animal of its size is not a functional (or *why?*) question, but a question about developmental constraints (a *what?* question). It asks not about what the brain does for you, but simply how you get a relatively bigger brain *if* you happen to want one.

If we are interested in cognitive *performance*, then absolute neural volume is the only variable that matters. More importantly, this relationship is likely to be order-specific because the brains of different taxonomic orders are organised in different ways and have different neural densities (Collins *et al.*, 2010; Herculano-Houzel *et al.*, 2007). For example, dolphin neocortices have only five cellular layers instead of the six present in primates, and in addition accommodate a very substantial specialised neural system for echolocation that primates, obviously, lack (Hof, Chanis & Marino, 2005; Marino *et al.*, 2007; Oelschläger, 2008). Similarly, most mammals have a very large olfactory cortex (and a well-developed sense of smell), but this is greatly reduced in anthropoid primates whose sense of smell is relatively poor by comparison. It is surely significant that the cognitive neuroscience community never uses relative volumes and would be astonished by any suggestion that they should.

It is important to appreciate that while much low-level cognition is dealt with by specialised (often subcortical) units, high level (i.e. 'smart') cognition is associated mainly with the neocortex, and is often unspecialised and highly distributed. Duncan (Duncan, 2001; Duncan *et al.*, 2000; Duncan, Assem & Sashidhara, 2020) has argued that one reason why the brain's prefrontal cortex is relatively undifferentiated is that it provides a generalised source for neural recruitment when tasks demand more processing capacity, a function that may even extend to recruiting the default mode network (Crittenden, Mitchell & Duncan, 2015). The neocortex makes up a variable proportion of total brain size across mammal species, ranging from 10% in insectivores to 20–40% in artiodactyls and carnivores, and, within the primates, from 50% in prosimians up to 80% in humans (Finlay & Darlington, 1995; Finlay, Darlington & Nicastro, 2001).

Within the neocortex, a very substantial neural network, or connectome, known as the default mode neural network that connects processing units in different parts of the cortex, is heavily involved in managing social relationships in both primates and humans (Mars *et al.*, 2012, 2016; Rushworth, Mars & Sallet, 2013; Li, Mai & Liu, 2014; Roumazeilles *et al.*, 2020; Yokoyama *et al.*, 2021). In anthropoid primates, the default mode network (and its ancillary connections into the limbic system and the cerebellum) forms a very substantial proportion of the neocortex. This goes some way to explaining both why the social brain relationship holds (with varying degrees of precision) irrespective of what measure of brain size is used, and at the same time why the fit gets better the closer the index focusses on the socially functional components of the neocortex (Dunbar, 1992; Joffe & Dunbar, 1997; Dunbar & Shultz, 2021a) – and perhaps why the Social Brain Hypothesis does not seem to hold for most non-primate mammals (see Section IV.4). Indeed, even within the primates, there are important quantitative differences in the size and structure of major neural tracts. Despite being the most social of the prosimians, *Lemur catta*, for example, has a disproportionately small dedicated social cognition neural tract compared to the anthropoid primates (Roumazeilles *et al.*, 2022), while semi-solitary nocturnal

prosimians like *Galago* lack the diversified temporal lobe connections characteristic of the more intensely social Old World monkeys and apes (Braunsdorf *et al.*, 2021).

The distinction between total brain size and neocortex size, and the confusion this can cause, is particularly well illustrated by the two largest brained primates, the gorilla (*Gorilla* spp.) and the orang utan. Both have very large brains (mainly because they have a large cerebellum, usually thought necessary to manage coordination of a very large body in trees), but surprisingly small neocortices. In both cases, their neocortex size is very close to what we would predict for their respective social group sizes (not foraging group size in the case of the orang!), helping to create a tightly linear relationship within the apes; total and relative brain size, on the other hand, are way off-line and yield no meaningful correlations with anything in particular (see Dunbar & Shultz, 2021a). Although it is inevitable that large neocortices need large brains to house them, the brain can, and does, evolve in a mosaic fashion (Barton & Harvey, 2000), as is very conspicuously the case in respect of neocortex size in primates (Finlay & Darlington, 1995).

Third, using the residuals from a body size equation for our analyses can have the unfortunate consequence of obscuring the fact that the causality may actually be the other way around: some species might have solved the problem of how to grow a bigger brain simply by growing a bigger body so as to exploit savings of scale provided by large body size without needing to change diet (Kleiber, 1961; Martin, 1990). What constrains brain size is not, of course, relevant to the question of what brains are used for: it is an answer to a question about the *costs* against which natural selection has to work, not the *benefits* it seeks to maximise. More worryingly, Rogell, Dowling & Husby (2020) draw attention to the fact that controlling for body size in this way can cause unpredictable sign reversals in multiple regressions, and this seems to be especially problematic in brain/body size data. This arises when collinearity between a predictor variable (here, brain size) and a third variable (body size) is high but functionally irrelevant. For a more general discussion of the problems created for multiple regression by ‘suppressor variables’, see Friedman & Wall (2005) and Smele (2023).

Kronmal (1993) also cautioned against the use of ratios in regression and correlation analyses because we cannot tell whether any resulting relationship is due to a change in the numerator or a change in the denominator, or both. He recommended that regression analyses be run with both components of the ratio as separate predictor variables in the form:

$$Y = X + Z^{-1} \quad (5)$$

Dunbar & Shultz (2021b) re-analysed the social brain data with $\log_{10}(\text{Group size})$ plotted against $\log_{10}(\text{Neocortex volume})$ and $\log_{10}(\text{Rest of brain volume})$ (with the latter as both the raw value and the reciprocal value) (Table 2). Three points should be noted. First, the regression equations are all highly significant: brain size is a good predictor of group

size, irrespective of how we index it. Second, this is mainly because neocortex volume rather than the volume of the rest of the brain drives the relationship, reflecting the fact that, in primates, the neocortex makes up the bulk of the brain. Third, rest-of-brain is a better predictor than its reciprocal, but its effect is strongly negative. This no doubt explains why neocortex ratio produces much stronger results than absolute brain volume does (Dunbar, 1992; Shultz & Dunbar, 2022). In effect, it indexes relative investment in ‘smart’ cognition as opposed to somatic management.

The fourth issue concerns the fact that brain size and body mass often have different evolutionary trajectories. Deacon (1990) pointed out that the interpretation of most relativised brain indices is made difficult by the fact that there is no independent baseline against which to assess allometric trends. In particular, the use of residuals from the regression line against body mass fails to recognise that, if the selection factors acting on brain size and body size differ, the two components can evolve at very different rates, often independently of each other, as Hager *et al.* (2012) showed on a sample of 10,000 mice (see also Gonzalez-Voyer, Winberg & Kolm, 2009; Smaers *et al.*, 2012). Fitzpatrick *et al.* (2012) found that, although there appears to be an effect of sexual selection on relative brain size in pinnipeds, this is entirely due to a change in male body mass; in fact, male and female brain sizes remain in close lockstep across species. Montgomery *et al.* (2010) found that there is a directional trend in brain mass but not body mass in primates; more importantly, temporal trends in body mass over geological time are not correlated with trends in brain mass (see also Aristide *et al.*, 2016).

Lande (1979) used brain/body size allometry across mammals to examine the evolutionary coupling of these traits. He argued that the comparatively weak genetic correlation of primate brain and body size as compared to other mammalian orders suggests that evolutionary changes in primate brain size are only weakly coupled with changes in body size. Moreover, the variance in the allometric relationship

Table 2. Regression analysis of the components of neocortex ratio as predictors of $\log_{10}(\text{group size})$ in 25 species of anthropoid primates.

Variable	Parameter estimate	<i>t</i>	<i>P</i>
<i>Model 1</i> ($F_{1,23} = 36.78$, $P \ll 0.0001$)			
Intercept	0.02	0.12	0.907
$\log_{10}(\text{Neocortex ratio})$	3.34	6.07	<0.001
<i>Model 2</i> ($F_{2,22} = 9.72$, $P = 0.001$)			
Intercept	-15.76	-2.53	0.019
$\log_{10}(\text{Neocortex volume})$	1.99	3.03	0.006
$\log_{10}(\text{Rest-of-brain volume})^{-1}$	32.68	2.41	0.025
<i>Model 3</i> ($F_{2,22} = 34.73$, $P \ll 0.0001$)			
Intercept	2.22	3.63	0.001
$\log_{10}(\text{Neocortex volume})$	6.37	6.71	<0.001
$\log_{10}(\text{Rest-of-brain volume})$	-7.14	-6.27	<0.001

Brain volume data from Stephan *et al.* (1981); group size data from Dunbar *et al.* (2018b).

increases with body size, suggesting that the two become increasingly decoupled as bodies get larger (perhaps because of the energetic savings of scale that large bodies allow: Martin, 1990). This is further compounded by the fact that if body mass changes faster than brain size [as, contrary to the claims of Deaner & Nunn (1999), is in fact the case in primates: Dunbar, 2015], using body size as the baseline will result in uninterpretable estimates of predicted size for brain regions. This may be one reason why brain size rather than body size acts as the biological constant determining most life-history variables (Mace, Harvey & Clutton-Brock, 1981; Clutton-Brock & Harvey, 1980; Harvey & Clutton-Brock, 1985; Harvey & Pagel, 1988).

(4) The fallacy of ‘secundum quid’ (over-generalisation)

There is a widespread tendency to assume that any biological principle or ‘law’ must be true for all taxa if it is to qualify as a biological universal (see, e.g. Grueter *et al.*, 2013). If this is not the case, so the argument runs, it must be a case of special pleading. Unfortunately, claims of this kind fail to appreciate the difference between universal laws and how these laws are instantiated in particular cases (which, in biological systems, will depend on the influence of many contextual variables). It is worth exploring this issue in a little more detail because it has profound implications for how we interpret attempts to test any hypothesis using comparative data.

The Social Brain Hypothesis was originally proposed to explain a particular feature of primate biology – the fact that, as Jerison (1977) originally pointed out, primates have much larger brains (absolutely and relative to body size) than all other groups of vertebrates (with the arguable exception of the cetaceans in respect of absolute brain size). In essence, the claim was that this reflected the fact that primates live in more complex societies than other vertebrates, and hence need a larger computer to manage the relationships involved (Jolly, 1966; Humphrey, 1976; Byrne & Whiten, 1988). Subsequent research has revealed that its basis lies in the specialised cognition required for bonded social groups (Shultz & Dunbar 2010a; Dunbar & Shultz, 2021a; Dunbar, 2023).

The original finding spawned a very large number of applications to other taxonomic groups, including various orders of mammals, birds, fish and, somewhat implausibly, even insects. The results have, however, been something of a mixed bag. In some cases, a significant quantitative relationship between group size and brain size has been identified [cetaceans (Fox, Muthukrishna & Shultz, 2017); some carnivores (Dunbar & Bever, 1998; Sakai *et al.*, 2011)], but in many cases not [insectivores, most other carnivores, artiodactyl ungulates, bats, birds, fish, and, perhaps not too surprisingly, insects (Gittleman, 1986; Dunbar & Bever, 1998; Pérez-Barbería *et al.*, 2007; Pollen *et al.*, 2007; Shultz & Dunbar, 2007, 2010a; Shumway, 2010; Smith *et al.*, 2010; O'Donnell, Clifford & Molina, 2011; Farris, 2016; Fedorova, Evans & Byrne, 2017; Hardie & Cooney, 2023)].

A more nuanced consideration of the data reveals that, while the Social Brain Hypothesis in its quantitative form applies to only a handful of mammalian orders [anthropoid primates, delphinids, equids, tylopods and elephantids – all of whom have large brains, exhibit a steady increase in brain size across geological time and live in bonded social groups (Shultz & Dunbar, 2007, 2010a; Fox *et al.*, 2017)], it appears as a qualitative difference (or phase shift) between monogamous and polygynous species (monogamous species have bigger brains) in insectivores, prosimians, bats, carnivores and ungulates, and most (but perhaps not all) birds (Dunbar & Bever, 1998; Beauchamp & Fernández-Juricic, 2004; Shultz & Dunbar, 2006, 2007, 2010a,b; Pitnick, Jones & Wilkinson, 2006; Pérez-Barbería *et al.*, 2007; West, 2014; Fedorova *et al.*, 2017; Hardie & Cooney, 2023), although some of the more social rodents (marmots; Matějů *et al.*, 2016) and the mustelids (Dunbar & Bever, 1998) may be exceptions.

Paradoxically, this monogamy effect has been interpreted as a negative relationship between brain size and group size, leading some to conclude that this is evidence against the Social Brain Hypothesis (e.g. Fedorova *et al.*, 2017; Hardie & Cooney, 2023). But this is a rather naïve interpretation of both the data and the theory, not to mention the behaviour of the animals concerned: the Social Brain Hypothesis is, as we have emphasised, about the formation of bonded relationships (‘friendships’) as a solution to the problem of how to create stable social groups in the face of environmental threats (the alternative being temporary aggregations). It is not about group size *per se*. Testing the social brain relationship by comparing pair-bonded species with those that live in large, anonymous flocks or herds rather misses the point (see Sections II.1 and IV.2). Living in anonymous, unstable, casual herds does not require significant cognitive skills. By contrast, pair bonds in mammals and birds alike are cognitively demanding in exactly the same way that ‘friendships’ are in anthropoid primates (and humans): although the number of relationships that needs to be managed is different, they involve similar kinds of decisions, trade-offs, coordination problems and investment in social bonding (grooming, huddling). In birds, species that have lifelong pair bonds (raptors, corvids, Psittaciformes, many seabirds) have significantly larger brains than annual pair-bonders whose pair bonds only last a single breeding season (many songbirds) and these, in turn, consistently have larger brains than species that have promiscuous mating systems associated with anonymous flocks (peacocks, ostriches, most Anseriformes) (Shultz & Dunbar, 2007, 2010b).

In other words, these findings actually provide support for the Social Brain Hypothesis properly understood, and do so in a way that greatly adds to our understanding of the phenomenon. A more nuanced evaluation of the differences among taxa suggests that bonded social relationships are one solution to the scalar stresses created by living in very large groups (Dunbar & Shultz, 2021b; Dunbar, 2023). If a taxon does not occupy habitats that require it to live in large groups, it will never exhibit any form of social brain relationship [unless it evolves pair-bonded monogamy for other

reasons (van Schaik & Dunbar, 1990; van Schaik, 2000; Opie *et al.*, 2013; Dunbar, 2022a). But when it does, a species can choose between incurring the cognitive and neurobiological costs of bonded social groups or opting for the less costly (but less effective) strategy of casual herding (Dunbar & Shultz, 2021b). Both strategies solve the ecological problem of predation risk, albeit in different ways and at different costs.

Group-living is not, of course, the only way to solve the predation risk problem. Evolving a large brain might instead allow individuals to deploy detection and evasion strategies that directly minimise predation risk without necessarily living in groups. Jerison (1977), for example, noted that carnivore and ungulate brain sizes exhibit a highly synchronised ratcheted trajectory through geological time: prey brain sizes initially outstrip predator brain sizes, who then respond by increasing brain size, which in turn causes prey brain size to increase further as though each taxon is adjusting its cognitive competence in response to the other's evolution of smarter counter-strategies. Most of these ungulates would have been herd-forming species. More generally, large-brained mammals experience less predation from the same guild of predators than smaller-brained species living in the same habitat, independently of group size (Shultz *et al.*, 2004; Shultz & Dunbar, 2006; Shultz & Finlayson, 2010). Similar results have been reported for fish (Kotrschal *et al.*, 2015). Other taxa avoid the costs of evolving large brains by adopting some form of crypsis associated with solitary foraging (e.g. the nocturnal prosimians; Burnham *et al.*, 2012). Each strategy has its own costs and benefits. Anthropoid primates seem to have adopted bonded sociality despite the neural costs involved because this offers a risk-averse solution: an individual is less likely to be caught on its own by a predator than is the case for species that form casual herds.

The fact that primates, in particular, might behave in different ways to birds and other mammals has prompted some authors (e.g. Logan *et al.*, 2018; Hooper *et al.*, 2022) to lament an overly anthropocentric approach to the Social Brain Hypothesis and argue for an approach that excludes humans, if not all primates. Doing so seems ill-advised, for several good biological reasons. First, since the quantitative version of the Social Brain Hypothesis applies only to anthropoid primates and not to prosimians, should we exclude all primates or only some of them? Second, it risks falling prey to speciesism: if we exclude primates for being too social, should we also exclude hoofed mammals and cetaceans because of their peculiar forms of locomotion? The answer, obviously, is: of course not. We want to be able to explain the diversity of life on Earth, not just some of it. If very big brains or particular kinds of behaviour are rare, we want to know why. Third, by excluding the best studied of all vertebrate species (humans) we risk ignoring a major source of knowledge: some things (notably neuroimaging studies of cognition) can be studied experimentally in humans much more easily than in other animals. Rather than narrowing the taxonomic focus, we need, if anything, to broaden it.

In short, demanding that a hypothesis must be universally supported across taxonomic groups and environments risks obscuring the range of solutions animals have evolved to solve the problems they face. For example, monogamy appears to have evolved under different environmental pressures in bony fish (Stanbrook *et al.*, 2022) than in primates (Opie *et al.*, 2013; Dunbar, 2022a), while social grouping has evolved as an anti-predator strategy in primates and ungulates, but as a hunting strategy in cooperative hunters like hyaena (*Crocuta* spp.), the African wild dog (*Lycaon pictus*) and the lion (*Panthera leo*).

V. STATISTICAL PITFALLS

Perhaps the most egregious problems arise in the statistical analyses used to test comparative hypotheses. We identify three issues under this heading: why causality matters in regression analysis, the presence of grades in the data (Simpson's Paradox), and the consequences of choosing the wrong regression model for the question being asked (in respect of which we identify two separate contexts: when testing causal hypotheses and when using these relationships in reverse engineering and other kinds of predictive analyses).

(1) How to test the wrong hypothesis

We noted, in Section II, that failure to identify the correct causal structure can lead to misleading results. A related problem emerges in respect of the statistical models used to test for correlated effects in comparative data sets. The problem stems from the fact that, in conventional regression analysis, it is only possible to have one dependent variable, although we can have as many independent predictor variables as we like. This can have the unfortunate consequence of forcing us to reverse the natural causal structure of our hypothesis in order to be able to run any analysis at all. As a result, almost every recent comparative analysis that has tried to test between alternative hypotheses for the evolution of large brains has fallen foul of this problem: among recent examples, see MacLean *et al.* (2014), Benson-Amram *et al.* (2016), DeCasien *et al.* (2017), Powell *et al.* (2017), Rosati (2017), Hardie & Cooney (2023) and Grabowskia *et al.* (2023). Although a simple bivariate regression will usually yield the same result whichever way round it is run, this is not true for multiple regression: which variable we choose to use as the dependent variable can yield *very* different results.

To see this, let us return to the example we discussed in Section II. We want to know whether the evolution of large brains was driven by group size or by diet. It might seem logical to do this by regressing brain size on body size, group size and diet as independent variables. If we do this, we find that brain size is significantly determined (in the statistical sense) by both diet and group size ($P < 0.001$), with body size as a

significant covariate (Fig. 10A). However, as Fig. 10A–D show, depending on which variable we place in pole position as the dependent variable, we get four completely different answers. At this point, we might be tempted to conclude, as Wartel *et al.* (2019) and others have done, that you can get any result that suits you from analyses of the social brain data, and, since there are no consistent patterns, we ought to abandon the entire research programme as conceptually flawed. The problem, however, is not the analyses or the data (you will get exactly the same results with *any* data set, whether or not it has anything to do with brains), but that the different regression models test completely different hypotheses, all of which are biologically perfectly sensible. In complex systems, causality matters.

Path analysis is the only sensible method to use in such cases because it allows all possible causal models to be considered, as well as allowing for feedback loops to be incorporated where appropriate. Although this can be

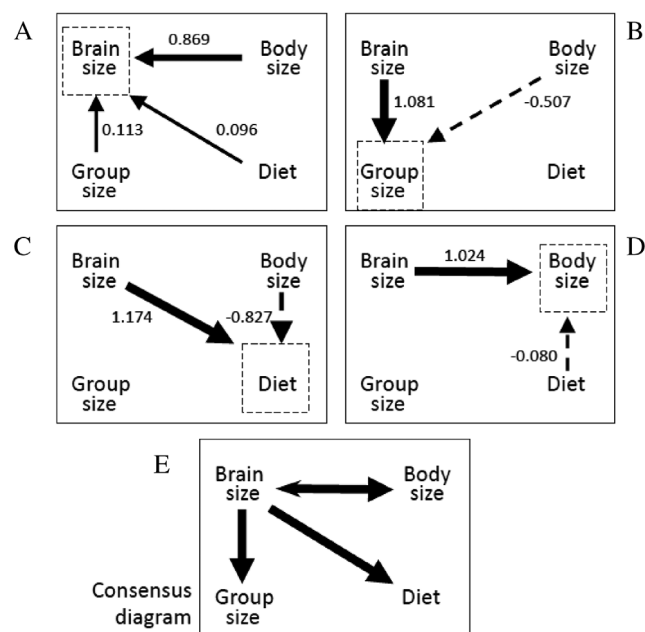


Fig. 10. Alternative versions of a multiple regression analysis between four variables testing for causal relationships, using primate data given by Powell *et al.* (2017). Diet is the percentage of diet accounted for by fruits (high-energy foods). Brain size, body size and group size are all log₁₀-transformed. In each case, the variable enclosed in the dashed square is the dependent variable in a multiple regression equation with the other three variables as predictors. Solid arrows indicate significant positive effects; dashed arrows indicate significant negative effects. Width of arrows is proportional to effect size. Only significant effects ($P < 0.05$) are shown. Numbers beside arrows are standardised β s. Group size correlates negatively, but not significantly, with diet in both directions. All four individual multiple regressions (A–D) are highly significant ($P \leq 0.0001$). The consensus diagram in (E) summarises all the significant positive effects, giving preference to the stronger effect if β values are significant in both directions.

computationally daunting when many variables are involved (although there are methods for reducing this: Watts *et al.*, 2022), it has the merit of allowing us to search through the set of possible models to find the one (or ones) that yield the best fit. For our present example, Fig. 10E gives the consensus model that takes into account all the significant results. This clearly indicates that the causal structure has a very specific form: brain size determines (i.e. influences) both group size and diet, with a tight coevolutionary loop between brain size and body size. Note, once again, that the causality of these relationships has the form of a here-and-now *constraints* model, not the form of an historical *selection* model; in selection terms, it implies that group size and a more frugivorous diet independently promoted large brain size (rather than the other way around). This implies that, historically, diet determined (i.e. constrained) brain size, not that a large brain enabled a change in diet (the assumption made by all studies that have favoured an ecological explanation for the evolution of large brains).

The advantage of path analysis is that it allows us to include a wider range of variables so as to build a more comprehensive model of the biological system. Figure 11 shows consensus, phylogenetically controlled path models, using data from two different studies [DeCasien *et al.* (2017) and Powell *et al.* (2017)] that claim to find results contrary to the Social Brain Hypothesis. In stark contrast to the results obtained by the original studies (which used simple multiple regression), both path analyses agree that brain size (and specifically neocortex size) is closely correlated with group size, while diet quality (indexed by the percentage of fruit in the diet) is better correlated with total brain size (but specifically subcortical brain volume rather than the neocortex, with which it is negatively correlated) suggesting an energetic trade-off between neocortex size and body size when species occupy habitats with high predation risk. Similar findings were reported by two other phylogenetically controlled path analyses (Dunbar & Shultz, 2007; Navarrete *et al.*, 2018) that used different data sets and took into account a much wider array of life-history, cognitive and ecological variables. As in Fig. 11, diet was related to brain size (not directly in this case, but indirectly *via* life history and body mass), while group size was directly correlated with brain size.

In other words, a more sophisticated statistical approach that takes feedback loops and biological causality into account gives us a very different, but quite consistent, picture. All four of these path models, using different data sets and different algorithms, agree that, in primates at least, brain size (and in particular neocortex size) has a tight co-evolutionary relationship with social group size, while foraging decisions are mainly a function of the demands imposed by the subcortical brain (essentially acting as a proxy for body mass). These analyses agree with the consensus view from a meta-analysis of all 44 analyses of primate comparative brain evolution that have been published to date (Shultz & Dunbar, 2022).

It is worth noting that both DeCasien *et al.* (2017) and Powell *et al.* (2017) claimed that they obtained different results to previous analyses because they had analysed data

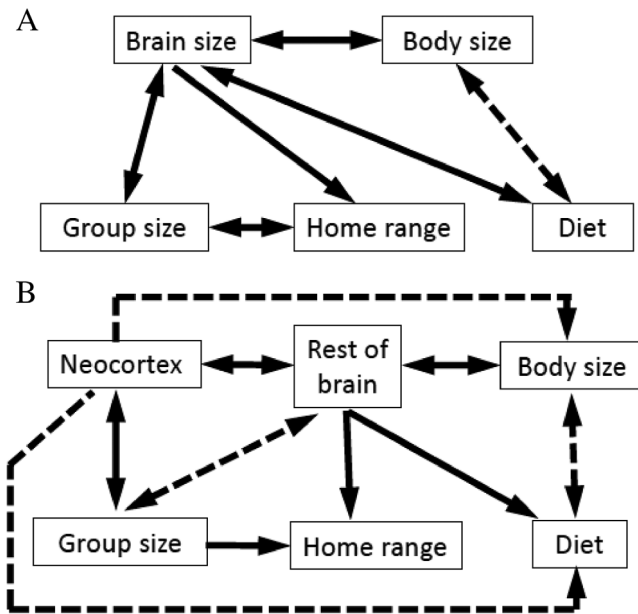


Fig. 11. Phylogenetically controlled path analyses of the causal pathways influencing primate brain size evolution for (A) total brain size [indexed as endocranial volume (ECV) using data from Powell *et al.*, 2017] or (B) histologically determined neocortex and rest of brain volumes (from Stephan *et al.*, 1981). In both cases, behavioural and demographic data are from Powell *et al.* (2017). The analysis was carried out by multi-model dredging using the *dredge* function in the MuMin R package to select the best candidate models based on Akaike Information Criterion corrected for small sample size (AIC_c) and model weight (for details, see Shultz & Dunbar, 2022). The *dredge* procedure permutes all possible phylogenetic generalised least squares models. Solid lines: significant positive causal relationships (causal direction indicated by arrows); dashed lines: statistically significant negative relationships. Redrawn from Shultz & Dunbar (2022).

from a larger sample of species and used ‘modern’ phylogenetic methods. In fact, this is not actually true: they used the same data sets and the same species, as well as the same phylogenetic methods, as most of the other studies. The reason they got different results is rather more prosaic: it is simply that they tested a different hypothesis. Unlike all other studies which all regressed group size on brain size, they opted to regress brain size on group size. In other words, they asked whether group size determines (constrains) brain size (a ‘becoming adapted’ hypothesis) rather than whether group size selected for (is constrained by) brain size (a ‘being adapted’ hypothesis) like everyone else.

The moral of this particular story is that multiple regression is not to be recommended unless we have thought through the implications of all the relationships involved and/or are very certain about the causal relationships we are testing. It should not be chosen – as, we suspect, it usually is – simply for statistical convenience. When we are dealing with complex biological systems with many feedback loops hidden in the mix, path analysis is always the safest way to

unpack the causal structure. At the very least, we should consider alternative forms of the regression model before drawing any conclusions.

We should, perhaps, conclude this subsection by noting that there are, of course, ways to test a ‘becoming adapted’ hypothesis directly using phylogenetic methods. One way to do this is by using Bayesian phylogenetic methods to reconstruct ancestral states, since this allows us to test the order in which two variables change in an evolutionary tree. There are, however, two important caveats. One is that, at best, our estimates of ancestral states are statistical guesses: they rest on the assumption of minimum parsimony, assume that traits are under tight genetic control and that particular models of evolutionary change hold. This is not always justified, and can lead to conclusions that are at odds with other evidence [e.g. the implausible claim that monogamy is the ancestral state for both apes and primates more generally made by Kappeler & Pozzi (2019)]. Second, the method only works if our phylogenies are fine-grained enough to allow a detectable lag between changes in the two variables to be identified. Pérez-Barbería *et al.* (2007) showed that while there was sufficient lag in the co-evolution of brain size and sociality in both carnivores and ungulates to establish a meaningful causal sequence (a switch to sociality was more likely to occur before a change in brain size in both orders, but a change in brain size did occasionally precede a change in sociality in carnivores), this was not the case in primates. In primates, the co-evolutionary ratchet is so tight that it is never possible to say which variable changed first: they always appear to change together. Contrasts of this kind between different taxonomic groups may not be unusual. Smaers *et al.* (2012) found similar differences in the co-evolution of brain and body size in different mammalian orders. This does not mean that there is no causal relationship involved in the primate case. It may simply mean that it happens too fast in geological time for a rather crude phylogenetic timescale to detect.

(2) Simpson’s Paradox

Most comparative analyses assume that they are dealing with simple unitary cause–effect relationships. However, Simpson (1951) pointed out that if there are grades in the data that reflect the influence of a third variable, then treating the data as a single homogenous distribution can give very misleading results. This is known as Simpson’s Paradox or the Yule–Simpson Effect, and is a version of the Ecological Fallacy. Figure 12A illustrates the problem: failure to take the existence of grades within the data into account yields a significant negative relationship ($r = -0.498$, $P = 0.05$) when the data quite obviously have a positive form (mean correlation for the two grades, $r = 0.984$). This problem was discussed at some length in the context of comparative analyses during the 1980s (Mace *et al.*, 1981; Harvey & Clutton-Brock, 1985; Harvey & Pagel, 1988, 1991), although mostly in respect of taxonomic grades of the kind originally identified by Jerison (1977).

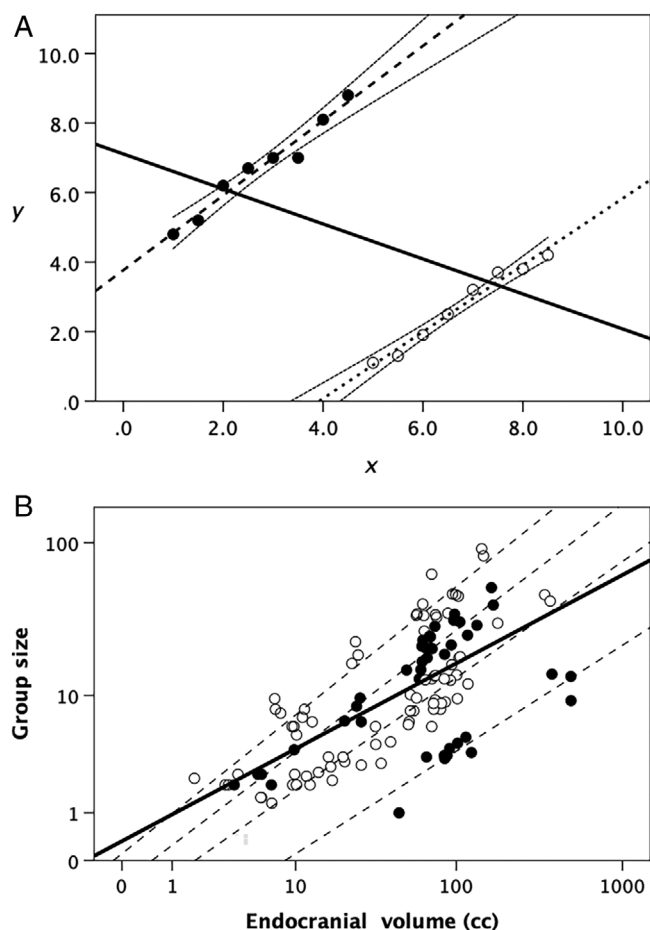


Fig. 12. (A) Yule-Simpson Effect. When there are distinct grades in a data set due to the influence of a third variable, a simple linear regression applied to the whole data set can yield a relationship that is diametrically opposite to the true relationship. The thick solid line is the overall ordinary least squares (OLS) regression; the dashed lines are the slopes for the separate grades (with the 95% CIs indicated in each case by the dotted lines). (B) Effects of grades in the primate social brain data, with species mean group size plotted against endocranial volume (ECV), both on a \log_{10} scale. The data have the classic tubular distribution characteristic of a data set with grades. A k -means cluster analysis reveals four grades (indicated by alternate unfilled and filled symbols), with least squares regressions fitted to individual grades. The solid line gives the overall least squares regression set through the full data set. The goodness of fit for the overall regression is $r^2 = 0.447$; the mean fit taking grades into account is $r^2 = 0.925$. Group size data are from Dunbar *et al.* (2018b); cranial volume data are from Isler *et al.* (2008). [Correction added on 14 April 2023, after first online publication: Figure 12 has been replaced.]

Of course, most real-world cases are not as extreme as that shown in Fig. 12A. A more common pattern is that in Fig. 12B, which shows the grades that are actually present in the primate social brain data. Dunbar & Shultz (2021a) showed, using k -means cluster analysis of five independent

brain data sets, that the social brain data consistently partition into four distinct clusters that form a set of parallel grades with very tight distributions (as indicated by the alternating black and white symbols in Fig. 12B). The ordinary least squares (OLS) regressions for the individual grades (the dashed lines) differ in intercept, but not in slope. On a double- \log_{10} plot, their mean slope is $b = 0.950$ (range 0.924–0.979, $0.851 \leq r^2 \leq 0.958$). The heavy line running across the grades is the OLS regression set through the whole data set ignoring the grades. All five regressions are significant ($P \ll 0.0001$), but the overall regression has a significantly shallower slope ($b = 0.617$; $t_{124} = 5.6$, $P < 0.0001$) than any of the individual grades, with a much poorer goodness-of-fit ($r^2 = 0.501$ without grades *versus* $r^2 = 0.925$ with grades). The reason there are grades in the data is not hard to see: in effect, the grades represent a series of glass ceilings on group size. When a taxon hits the upper limit on group size for the grade it is on, it has two choices: to push group size a little higher at the expense of losing group coherence, or move sideways onto the next grade by increasing brain size so as to allow new cognitive strategies that create more deeply bonded groups, thereby allowing further increases in group size (Dunbar & Shultz, 2021a; Dunbar, 2023).

If our interest is simply in establishing whether brain volume is a significant predictor of group size, this may not matter too much: at worst OLS regression provides a conservative test (it reduces the risk of Type I errors, albeit at the cost of increasing the risk of Type II errors). We are therefore very unlikely to conclude that there is a relationship present if there actually is not one there. However, the shallower slope on this regression will be much more problematic in a multiple regression, because the residuals to the OLS line will be much greater than they should be. As a result, the effect size for this relationship will be radically underestimated relative to that for any variable that exhibits no grade effect (e.g. diet, which seems to be grade-free), giving the false impression that the latter variable makes a disproportionately large contribution. We may be misled into concluding that there is no functional relationship at all for group size when, in fact, there is a very strong one. Both DeCasien *et al.* (2017) and Powell *et al.* (2017) unwittingly fell foul of this trap. The lesson is that it is always wise to inspect the graphical distribution of data before running any statistical test – and to be sufficiently attuned to both data distributions and real-life biology to recognise subtle patterns. Far too many researchers seem to rely on the statistics package print-outs without bothering to look at their data first. That may be fine if you really know for certain exactly what form the data have. But if you only *think* you know, nature will not spare your blushes.

Note, by the way, that we have not corrected for phylogeny in any of these analyses. This is because it is only necessary to do so when there is a significant phylogenetic signal, such that the degrees of freedom will be artificially inflated (leading to an elevated risk of Type I errors – falsely rejecting the null hypothesis). A strong phylogenetic signal for brains

and behaviour may well be characteristic of birds and some mammalian orders (see Logan *et al.*, 2018; Hooper *et al.*, 2022), but this is not true of all taxa – and especially primates, where the phylogenetic signals for all behavioural indices, including the social brain data, are low or non-existent (Kamilar & Cooper, 2013). No study, at least of primate data, has produced qualitatively different results by using phylogenetic methods. Aristide *et al.* (2016) and Hassler *et al.* (2022), for example, analysed the same data set with and without phylogenetic correction and obtained identical results. In short, phylogenetic methods should only be used when there is a demonstrable phylogenetic signal, and hence a risk that degrees of freedom will be inflated. Including phylogeny when it is not necessary is, at best, a form of virtue signalling whose effect is simply to reduce statistical power: in effect, it is a statistical version of the Zahavi Handicap Principle ('See how strong my result is – even adding unnecessary variables will not destabilise it'). It is important to remember that phylogenetic methods, in and of themselves, do not test selection hypotheses; the best they do is tell us whether a particular mode of neutral genetic evolution (drift) explains the observed data (essentially an ontogeny or *what?* question, not a *why?* question, once again raising the spectre of Tinbergen confounds).

(3) When a regression is not the right regression

Although regression analysis forms the backbone of comparative analysis, it seems not to be widely appreciated that regression actually constitutes a family of methods that differ in the assumptions made about the data. These assumptions mainly relate to how the residuals to the line of best fit are calculated, and to a requirement that the data are bivariate normal in form. In respect of the residuals, the main options are to take them against the y -axis, against the x -axis, perpendicular to the line of best fit, or from both x and y axes simultaneously (i.e. the area of the triangle to the line of best fit subtended by the datapoint). Which of these is best to use depends largely on the ratio of the error variances on the two axes.

OLS regression (the most commonly used model) assumes that the values on the x -axis are measured without error. This method was originally developed for use in experimental studies of the typical dose–response kind where the experimenter determines the values on the x -axis variable (e.g. by giving experimental subjects different carefully titrated quantities of some drug). When this is the case, and the data are bivariate normal, the calculation of the parameter values and statistical moments can be simplified by minimising just the residuals on the y -axis (the axis with all the error variance). In the rare cases where y -axis values are measured without error but there is significant error variance on the x -axis, major axis regression is the appropriate technique (it minimises residuals on the x -axis). When both variables are based on observational data, there is likely to be significant error variance on both axes. The presence of grades in the data will only exacerbate this because the data will likely be bivariate uniform (in effect, the data have a more tubelike distribution, as in Fig. 12B; see also Dunbar & Shultz, 2021a)

rather than being bivariate normal, thereby invalidating the central assumption for OLS regression. In both cases, this will have the effect of lowering the slope of the OLS regression equation (Fig. 12B).

Kendall & Stuart (1979) have shown that, if the error variances on the two axes are equal or unknown, then reduced major axis regression (RMA, or model II regression of Sokal & Rohlf, 2012) gives the maximum likelihood estimate of the true functional relationship (Fig. 13, solid line). RMA minimises the residuals on both axes simultaneously, and is equivalent to the geometric mean of the conventional OLS regression of y on x (long-dashed line in Fig. 13) and its converse (x regressed on y : short-dashed line in Fig. 13). Rayner (1985) recommended RMA regression when the error variances are unknown or there is error variance on both axes, because it is the only regression method that is independent of the error correlation. Its only disadvantage is that it is difficult to assign significance values to the regression coefficients, although Rayner (1985) does give a method for calculating 95% CIs for the slope.

As a matter of simple practice, however, the consensus has been that OLS regressions should only be used when the goodness of fit $r^2 > 0.95$, since OLS and RMA methods converge when the fit is high (see Martin, 1990). When $r^2 < 0.95$, RMA regression is recommended (although if $r^2 < 0.60$, even this will lose power: Jolicoeur, 1990). Note, by the way, that although Smith (2009) is often cited as grounds for not using RMA regression in comparative analyses, his justification for claiming this is mathematically spurious.

(4) The art of prediction

In the previous subsection, we pointed out that if we underestimate the proportion of variance explained by a particular

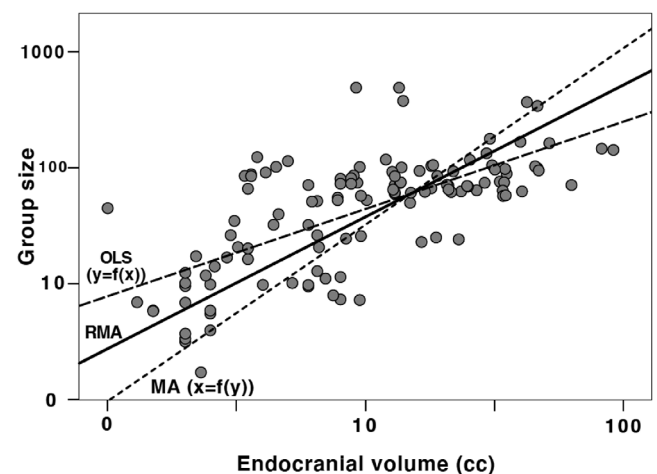


Fig. 13. Group size plotted against Endocranial volume (ECV) for primates, with both axes on a log scale. The long-dashed line is the ordinary least squares (OLS) regression; the solid line is the reduced major axis (RMA) regression for the same data; the short-dashed line is the equivalent major axis (MA) regression equation for x regressed on y . Group sizes from Dunbar *et al.* (2018b); ECV from Powell *et al.* (2017).

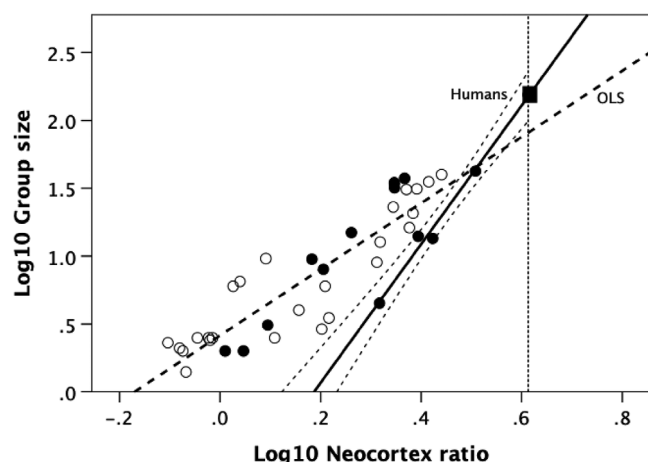


Fig. 14. Mean species group size plotted against neocortex ratio. Neocortex ratios are from Stephan *et al.* (1981) and group sizes are from Dunbar *et al.* (2018b). Grades (indicated by alternating black and white symbols) were identified using a *k*-means clustering analysis (see Dunbar & Shultz, 2021a). The ordinary least squares (OLS) regression line (solid line, with 95% CIs as light dashed lines) for the hominoids-only grade is shown. The overall OLS regression line for the whole data set (ignoring grades) is indicated by the heavy dashed line. For the full data set, $r^2 = 0.978$ taking grades into account; for the hominoids-only grade, $r^2 = 0.989$. For both graphs, the vertical dotted line demarcates neocortex ratio for modern humans, and the square symbol indicates the observed mean social group size for humans (153.7, based on 23 estimates of personal social network size and the size of small-scale communities; Dunbar, 2020).

predictor variable, then we inevitably overestimate the contribution by any other variable whose slope is estimated more accurately. There is, in addition, a second reason why we should worry about this. Comparative analysis is not simply about testing causal relationships. Increasingly, we want to use the relationships we find to predict values for particular taxa. This has been especially common in palaeontology for more than half a century, where unknown traits are commonly estimated from skeletal proxies based on cross-species

samples (e.g. see Pearce, Stringer & Dunbar, 2013; Dingwall *et al.*, 2013). This approach has also been used to predict extant species' responses to climate change (e.g. Dunbar, 1998a; Korstjens, Lehmann & Dunbar, 2010; Lehmann, Korstjens & Dunbar, 2010). Another important use is in reverse engineering exercises, where the residuals between observed and regression-predicted values are used both to estimate the selection pressure that a species is under at a particular point in its evolutionary history (Dunbar, 2009, 2014, 2022b; Bannan, Bamford & Dunbar, 2023) and to ask whether we might have missed any important factors when constructing our hypotheses (i.e. how much of the variance does our model *not* explain?). The use of reverse engineering to identify time points where lineages have been forced to undergo a phase transition by introducing some new adaptive trait or, conversely, to identify an environmental factor that might have triggered the emergence of an adaptive shift is a technique likely to prove of increasing value. Although we might get away with using OLS regression when we only want to know whether or not we have a bivariate correlation, we cannot afford to be so cavalier when using a regression line to make a prediction, especially when (i) the data point we want to predict lies beyond the range of the data on which the estimate of the slope is based and (ii) the axes are log-transformed (both of which will exaggerate the prediction error).

To illustrate the problem, consider the case where we might want to predict a value for group size in humans, based on the primate social brain relationship. When Lindenfors, Wartel & Lind (2021) did this, they found that the values predicted by their regression equations were *much* lower than both the predicted value given by Dunbar (1993a) and the actual observed value. In addition, the confidence intervals around their predictions were so wide that almost any number would fit, making a reliable prediction impossible. How, then, is it possible for two analyses of essentially the same data to come to radically different conclusions? The answer should, perhaps, be obvious: a combination of Simpson's Paradox and OLS regression. To see why, we plot the social brain data and the relevant regressions in Fig. 14. (We use the neocortex ratio data here, but in fact any index

Table 3. Regression equations for the data in Fig. 14 and predicted mean social group size for humans.

Regression model	Equation	r^2	Predicted group size for humans
Observed mean human group sizes ($N = 24$ online and offline samples, worldwide)**			153.7 ± 35.9
Lindenfors <i>et al.</i> (2021): neocortex ratio or neocortex plus rest of brain (mean of $N = 4$)			59.2^{\dagger}
ECV plus body mass (mean of $N = 4$)			51.6^{\dagger}
All data: OLS	$\log_{10}(\text{Group}) = 0.42 + 2.44 \times \log_{10}(\text{Neocortex ratio})$	0.729	82.4
All data: RMA	$\log_{10}(\text{Group}) = 0.22 + 3.11 \times \log_{10}(\text{Neocortex ratio})$	—	139.0
Hominoids grade only: OLS	$\log_{10}(\text{Group}) = -0.95 + 5.11 \times \log_{10}(\text{Neocortex ratio})$	0.989	152.2
Apes only: OLS	$\log_{10}(\text{Group}) = -0.90 + 4.97 \times \log_{10}(\text{Neocortex ratio})$	0.958	139.8

ECV, endocranial volume; OLS, ordinary least squares; RMA, reduced major axis.

**Data from Dunbar (2020).

† OLS regressions, using brain volumes based on Navarrete *et al.* (2018), and ECV, body mass and group sizes from DeCasien *et al.* (2017). For problems with both these data sets, see Dunbar & Shultz (2021a).

of brain size will yield the same result; see Dunbar & Shultz, 2021a). Table 3 provides the regression statistics for the different regression equations.

The overall OLS regression through the entire data set (essentially Lindenfors *et al.*'s analysis) has a slope ($b = 2.44$) that is highly significant (dashed line in Fig. 14; $P < 0.001$). It is, however, considerably shallower than that for the equivalent overall RMA regression ($b = 3.11$), and both are considerably shallower than the OLS regressions for the individual grades (averaged across the four grades: $b = 3.88$; for the hominoid-only grade: $b = 5.11$; for apes only: $b = 4.97$). The goodness of fit for the conventional overall OLS regression is a respectable $r^2 = 0.729$, which would usually be considered very acceptable. However, the overall goodness-of-fit for an OLS regression taking the grades into account is $r^2 = 0.978$, a *very* significant improvement. The goodness-of-fit for the hominoid-only grade on its own is $r^2 = 0.989$, that for the apes alone is $r^2 = 0.958$. In other words, the slopes steepen and sharpen up as the sample focusses increasingly on the appropriate grade.

Because the regression slopes vary so widely, the predictions for human group size are equally variable (Table 3). The empirically determined mean human group size, as shown by the filled square, is ~ 154 [range 72–250 for $N = 24$ samples (Dunbar, 2020); with the largest sample to date (61 million individual Facebook pages) giving a mean egocentric network size of exactly 149 (Bond *et al.*, 2012)]. Lindenfors *et al.* (2021) give predicted values, based on four different overall regression equations (two conventional OLS and two Bayesian) each for neocortex volume and for ECV, whose individual predictions range between 16.4 and 108.6. (Note that, implausibly, the lower prediction gives a value for mean group size that is smaller than the actual mean group size of a third of all non-human primate species, which ought to alert us to the fact that there must be a problem with the regression analysis.) The overall OLS regression shown in Fig. 14 predicts a value of 82.4 (well within the range of estimates given by Lindenfors *et al.*, 2021). By comparison, an overall RMA regression does considerably better with a prediction of 139. The grade-specific equation does better still, however: the hominoid-only OLS regression (the regression equation that, incidentally, has been used in all similar analyses since 1993) predicts a value of 152.2, which is indistinguishable from the observed value of 153.7.

Lindenfors *et al.* (2021) offer a second reason for not being able to predict a value for humans reliably: the confidence intervals on the predicted value(s) are so wide (2–520 across their eight regressions) that almost any value would confirm the prediction. Notice, however, that the confidence intervals they give are much wider than those generated by the regression equations for Fig. 14. In fact, what Lindenfors *et al.* (2021) report are confidence intervals when they should be giving prediction intervals. Although, somewhat confusingly, both are often referred to as confidence intervals, the two are, in fact, conceptually quite different: one is based on the scatter in the data and estimates the range within which all *individual* values (known and as yet unknown) will lie; the other

is based on the range within which the slope parameter varies, and hence gives the range within which predictions for a *mean* value should lie. The second is inevitably much narrower than the first. In effect, these parallel the difference between standard deviations and standard errors. The 95% *prediction* interval on the estimate of the population mean for the overall OLS regression line in Fig. 14 is {55.0–120.2}, whereas the 95% *confidence* interval (all possible individual cases) is {26.0–251.2}. If we are concerned with predicting the *mean* value for humans, not the likely range of all possible individual values, then we are only interested in the first. The observed mean value clearly falls well outside the prediction interval for the overall OLS line {55.0–120.2}, but well within the prediction interval for the hominoid grade OLS regression {70.8–195.0}. As it happens, the likely range in individual values is actually a very good fit to the observed 95% range of 58–238 for human personal social network sizes (Hill & Dunbar, 2003).

The question we should, perhaps, ask in these contexts is not how wide the CIs are, but rather the Bayesian question of how closely the observed value matches the predicted value. Applying Bayes' Theorem, with likelihoods estimated from the prediction intervals, the observed value of 154 is clearly a very good fit indeed to the value predicted by the equation for the correct social brain grade (for the hominids-only regression: $p_{\text{posterior}} = 0.949$; for the ape-only regression, $p_{\text{posterior}} = 0.946$), whereas it is a very poor fit to the prediction from the overall OLS regression equation ($p_{\text{posterior}} = 0.079$) or any of the Lindenfors *et al.* (2021) regressions. Note, by the way, that interpolating human neocortex ratio into regressions set to the other three grades in the social brain relationship predicts rather closely the values for the layers in the fractal structure of both human egocentric social networks and social groupings (Dunbar, 2020); indeed, this is also the case for those primates that have fractally structured multilevel social systems (Dunbar & Shultz, 2021a).

VI. DISCUSSION

A number of recent studies have claimed that analyses of data for some comparative questions (notably on brain evolution) are unstable and generate contradictory results. We have argued, however, that most of these inconsistencies are actually due to the way the analyses have been carried out, not to the underlying phenomena themselves or to inadequacies in the data, as has commonly been claimed. We identified four broad categories of error: (i) conflating different Tinbergen Questions (i.e. ignoring the fact that biological processes are intrinsically systemic); (ii) failure to appreciate the significance of Dobzhansky's Dictum when testing evolutionary hypotheses; (iii) poorly chosen hypotheses and proxy variables; and (iv) inadvertently testing a different hypothesis to the one intended. We suggest that most of these sources of error could have been avoided if a more biological,

systems-based approach had been adopted. Using path analysis instead of multiple regression would, for example, have allowed more careful consideration to be given to alternative relationships between variables than the overly simplistic single-cause/single-effect causality that characterises so many analyses.

Perhaps the most serious casualty of this has been our understanding of anthropoid primate sociality, but we would argue that it has also impeded our understanding of mammal and bird sociality more generally. Most of the problems we have examined arise from failing to appreciate that primate sociality is actually in a very different league from the kinds of sociality we find in most (but not all) birds and mammals. Anthropoid primate social systems are based on bonded relationships, mediated by social grooming. These relationships depend on high-order cognitive abilities, many of which are unique to the anthropoid primates and depend on brain regions that are found only in this taxon (Passingham & Wise, 2012). One consequence is that primate groups exhibit a degree of multilevel organisation based on a distinctive fractal structure, and a degree of coherence and stability, that is uniquely characteristic of this taxon. This raises an important evolutionary question: why have primates (and those few other taxa that have similar social systems) gone down this route? Why did all the other species not do so?

It is important to remember that the Social Brain Hypothesis asks two separate questions: (i) why do some taxonomic groups (e.g. primates) have larger brains than other taxa of similar size (say, felids or sciurids) and (ii) within the primates, why do some species have larger brains than others? Almost no comparative brain analysis explains why primates need so much neural computing power to deal with foraging decisions that most other mammalian orders successfully solve with much smaller brains. Indeed, the decisions that cursorial carnivores make in stalking and pursuing prey are far more complex than any decision that a foraging primate makes (the disorganised chaos of chimpanzee hunts notwithstanding), yet felid brain size has undergone very little increase over this suborder's entire evolutionary history (Shultz & Dunbar, 2010a).

While primate sociality unquestionably has unique features, it is important to remember that a number of other mammalian orders (notably the equids, tylopods, delphinids, elephants, and perhaps others) also have stable, bonded social groups, and this may also be true of some avian taxa. Among the birds, likely examples include guinea fowl (*Acryllium vulturinum*), babblers (Timaliidae), mousebirds (Coliidae), woodpeckers (Picidae) and parrots (Psittaciformes), many of which have stable social groups (Papageorgiou *et al.*, 2019), albeit at the small end of the primate group size distribution. Note, however, that some of these are cooperative breeders with a single breeding pair (e.g. babblers; Nelson-Flower *et al.*, 2011), and may thus more closely resemble callitrichid primates who, uniquely among the anthropoid primates, lack some of the key brain regions that underpin bonded sociality in the other anthropoid primates (Passingham & Wise, 2012) and, as a result, have groups with a very different kind of social

style that are socially more fluid (Lukas & Clutton-Brock, 2018; Dunbar & Shultz, 2021a).

There are two issues here. First, we need to be cautious of assuming that just because small-brained birds have multilevel groups they do this in the same way, using the same cognition, as primates. Prosimian primates and some artiodactyl ungulates also live in small stable groups, but these are not based on the same cognition that underpins anthropoid primate social groups (Dunbar & Shultz, 2021a, 2022). As Tinbergen (1963) reminds us, the same functional outcome can be achieved by exploiting different mechanisms as a result of different evolutionary pathways. Second, the issue in all these cases is not whether the Social Brain Hypothesis is wrong because it does not apply in its quantitative form to all taxa, but rather what is different about taxa that do not exhibit such a relationship – and, in evolutionary terms, why?

The quantitative version of the Social Brain Hypothesis that we find in primates is simply one solution to the problem created by the scalar stresses of living in large groups (Dunbar & Shultz, 2021b). Bonded social groups are, however, cognitively very expensive (Dàvid-Barrett & Dunbar, 2013; Lewis *et al.*, 2017), and modelling suggests that only in a very small corner of the environmental state space does the balance between the costs and benefits make it worth a taxon's while opting for this strategy rather than less costly alternatives – with predicted frequencies that are very close to those actually observed (Sutcliffe, Dunbar & Wang, 2016).

If we are to understand why some lineages have opted for one solution and others for another, we need to develop a better understanding of the social dynamics of these species so as to determine how, when and why lineages are forced to switch into different strategic pathways in order to cope with the environmental stressors they encounter (a reverse engineering issue). Failure to do so risks overlooking aspects of the biological world that are both in need of explanation precisely because they stand out as puzzling exceptions and are, at the same time, potentially the most illuminating for understanding the grand sweep of adaptation. As the classical ethologists reminded us, nothing is more important than immersing ourselves in the daily lives of our study species – so as to be able to see the world from their point of view with all its cognitive limitations. It is the animals' own behaviour that should inform our hypotheses, not our theoretical preconceptions of how the world ought to be. Theories are tools for exploring the world, not inviolable truths (Dunbar, 1995). To this may be added the importance of not limiting this immersion to a single study species from one taxonomic group. There is no substitute for firsthand knowledge. We might add that our views have benefitted from the fact that both of us have undertaken field work (and, in some cases, experimental studies) on trees, insects (notably dragonflies), birds, ungulates, carnivores, primates and humans. That breadth of taxonomic perspective is what has allowed us to appreciate and understand the complexity and richness of what we have to explain.

In short, a much more nuanced approach is required that views the quantitative form of the Social Brain Hypothesis, as

we find it in primates, as being just one way that a set of universal biological principles play themselves out in particular biological contexts. Species do not arrive at a particular environmental space as ‘blank slates’. They do so with constraints imposed by their evolutionary histories, and these may predispose them to certain kinds of solutions because the alternatives are too costly to evolve – the reason, as Davies (1978) memorably reminded us, why butterflies never evolved machine guns. As with all behavioural ecological phenomena, the answer lies in a combination of species’ inherited biological constraints, the phenotypic flexibility that a species is capable of exhibiting, the nature of the scalar stresses that it faces from living in groups of different size, and the evolutionary trade-offs that all individuals are forced to make in their attempts to maximise fitness.

We need to know how flexible the structural and behavioural aspects of sociality are in different species (see also Strier *et al.*, 2014; Socias-Martínez & Peckre, 2023), and the extent to which adopting a particular social trajectory makes it difficult for species to back-track to alternatives when circumstances change. For example, the adoption of pairbonded monogamy by a number of primate lineages (mainly the smaller cebids and the gibbons) appears to have necessitated cognitive adaptations to support lifelong relationships that seem to be difficult to reverse (Opie *et al.*, 2013). This may well also be true of other mammalian orders and birds. More importantly, Pérez-Barbería *et al.* (2007) found that, in contrast to carnivores and ungulates, reversals in brain size never occur in primates, suggesting that whatever cognitive changes were introduced by increases in primate brain size are too difficult to unpick should there be selection against large group size at a later time. At the same time, we need to beware of assuming that an evolutionary approach consists simply of showing that behaviour is genetically determined. Brain size and structure might well be genetically determined (although probably much less so than is often assumed: Maguire *et al.*, 2000), but the point of having a large brain is to buffer the species against environmental stressors by being able to adjust behaviour without needing to undergo immediate genetic evolution (the classic Baldwin Effect). The cercopithecine monkeys offer a particularly germane example in this respect. They seem, as a taxon, to be unusually adaptable – more so even than the apes – and this may account for their remarkable ability to colonise an unusually wide range of habitats. We need to know much more about the extent to which animals can facultatively adjust aspects of their behaviour and biology (Strier *et al.*, 2014).

VII. CONCLUSIONS

(1) Comparative analyses are the mainstay of evolutionary hypothesis-testing. However, they have sometimes fuelled surprisingly partisan disputes. This been particularly true in respect of attempts to understand the evolution of large brains and smart cognition (the Social Brain Hypothesis).

(2) We argue that these conflicts are largely a consequence of poorly thought out hypothesis-testing rather than anything to do with either the theories or the data. In many cases, it seems to be a consequence of adopting a psychological (or mechanisms) approach to hypothesis-testing rather than a biological systems-based one. When we approach the problem in a more biological way, the results are robust and consistent and make sense of all contradictory findings.

(3) We identify four main sources of error, many of which are well-known logical fallacies. We particularly identify: confounding Tinbergen Four Questions, confusing ‘being adapted’ explanations with ‘becoming adapted’ ones (Dobzhansky’s Dictum), poorly chosen proxies for use in hypotheses-testing, and inappropriate statistical designs (notably falling foul of Simpson’s Paradox).

(4) These errors often seem to reflect a naïve understanding of animal (but especially anthropoid primate) sociality and the cognition that underpins it, creating a risk that we lose sight of the wider picture of mammalian (and perhaps avian) social and cognitive evolution.

(5) There is a pressing need for those who undertake comparative analyses to have a better understanding of the natural history of the species they study.

(6) Far too many analyses of brain evolution ignore the wealth of neuropsychological evidence on brain anatomy and function, and we recommend that greater attention is paid to this literature.

(7) We urge a more careful approach to comparative analyses that takes proper account of the biological differences between different taxa and a more systems-based approach to hypothesis-testing.

VIII. REFERENCES

- AIELLO, L. C. & WHEELER, P. (1995). The expensive-tissue hypothesis: the brain and the digestive system in human and primate evolution. *Current Anthropology* **36**, 199–221.
- ALBON, S. D., MITCHELL, B. & STAINES, B. W. (1983). Fertility and body weight in female red deer: a density-dependent relationship. *Journal of Animal Ecology* **52**, 969–980.
- ALTMANN, J. (1974). Observational study of behavior: sampling methods. *Behaviour* **49**, 227–266.
- ALTMANN, S. A. (1991). Diets of yearling female primates (*Papio cynocephalus*) predict lifetime fitness. *Proceedings of the National Academy of Sciences of the United States of America* **88**, 420–423.
- ALTMANN, S. A. (1998). *Foraging for Survival: Yearling Baboons in Africa*. University of Chicago Press, Chicago.
- AMICI, F., AURELI, F. & CALL, J. (2008). Fission-fusion dynamics, behavioral flexibility, and inhibitory control in primates. *Current Biology* **18**, 1415–1419.
- ANCRENAZ, M., LACKMAN-ANCRENAZ, I. & MUNDY, N. (1994). Field observations of aye-ayes (*Daubentonia madagascariensis*) in Madagascar. *Folia Primatologica* **62**, 22–36.
- ANDREWS-HANNA, J. R., REIDLER, J. S., SEPULCRE, J., POULIN, R. & BUCKNER, R. L. (2010). Functional-anatomic fractionation of the brain’s default network. *Neuron* **65**, 550–562.
- ARCHIE, E. A., ALTMANN, J. & ALBERTS, S. C. (2012). Social status predicts wound healing in wild baboons. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 9017–9022.
- ARENDTS, J. C., CHEUNG, M. Y. C., BARRACK, M. T. & NATTIV, A. (2012). Restoration of menses with nonpharmacologic therapy in college athletes with menstrual disturbances: a 5-year retrospective study. *International Journal of Sport Nutrition and Exercise Metabolism* **22**, 98–108.
- ARISTIDE, L., DOS REIS, S. F., MACHADO, A. C., LIMA, I., LOPES, R. T. & PEREZ, S. I. (2016). Brain shape convergence in the adaptive radiation of New

- World monkeys. *Proceedings of the National Academy of Sciences of the United States of America* **113**, 2158–2163.
- ASHTON, B. J., KENNEDY, P. & RADFORD, A. N. (2020). Interactions with conspecific outsiders as drivers of cognitive evolution. *Nature Communications* **11**, 4937.
- BANNAN, N., BAMFORD, J. & DUNBAR, R. I. M. (2023). The evolution of gender dimorphism in the human voice: the role of octave equivalence. *Current Anthropology* (in press).
- BARTON, R. A. & HARVEY, P. H. (2000). Mosaic evolution of brain structure in mammals. *Nature* **405**, 1055–1058.
- BEARDER, S. K. (2008). Lorises, bushbabies, and tarsiers: diverse societies in solitary foragers. In *Primate Societies* (eds B. B. SMUTS, D. L. CHENEY, R. M. SEYFARTH, R. W. WRANGHAM and T. T. STRUHSACKER), pp. 11–24. University of Chicago Press, Chicago.
- BEAUCHAMP, G. & FERNÁNDEZ-JURICIC, E. (2004). Is there a relationship between forebrain size and group size in birds? *Evolutionary Ecology Research* **6**, 833–842.
- BENSON-AMRAM, S., DANTZER, B., STRICKER, G., SWANSON, E. M. & HOLEKAMP, K. E. (2016). Brain size predicts problem-solving ability in mammalian carnivores. *Proceedings of the National Academy of Sciences of the United States of America* **113**, 2532–2537.
- BERGHÄNEL, A., SCHÜLKE, O. & OSTNER, J. (2015). Locomotor play drives motor skill acquisition at the expense of growth: a life history trade-off. *Science Advances* **1**, e1500451.
- BETTRIDGE, C., LEHMANN, J. & DUNBAR, R. I. M. (2010). Trade-offs between time, predation risk and life history, and their implications for biogeography: a systems modelling approach with a primate case study. *Ecological Modelling* **221**, 777–790.
- BICKART, K. C., HOLLENBECK, M. C., BARRETT, L. F. & DICKERSON, B. C. (2012). Intrinsic amygdala-cortical functional connectivity predicts social network size in humans. *Journal of Neuroscience* **32**, 14729–14741.
- BICKART, K. C., WRIGHT, C. I., DAUTOFF, R. J., DICKERSON, B. C. & BARRETT, L. F. (2011). Amygdala volume and social network size in humans. *Nature Neuroscience* **14**, 163–164.
- BOND, R. M., FARISS, C. J., JONES, J. J., KRAMER, A. D., MARLOW, C., SETTLE, J. E. & FOWLER, J. H. (2012). A 61-million-person experiment in social influence and political mobilization. *Nature* **489**, 295–298.
- BRAUNSDORF, M., FRECHES, G. B., ROUMAZEILLES, L., EICHERT, N., SCHURZ, M., UITHOL, S., BRYANT, K. L. & MARS, R. B. (2021). Does the temporal cortex make us human? A review of structural and functional diversity of the primate temporal lobe. *Neuroscience & Biobehavioral Reviews* **131**, 400–410.
- BRENT, L. J. N., RUIZ-LAMBEDES, A. & PLATT, M. L. (2017). Family network size and survival across the lifespan of female macaques. *Proceedings of the Royal Society, London* **284B**, 20170515.
- BURNHAM, D., BEARDER, S., CHEYNE, S., DUNBAR, R. I. M. & MACDONALD, D. (2012). A taste of predation past: is nocturnal behaviour in primates explained by the threat of predation by cats? *Folia Primatologica* **83**, 236–251.
- BYRNE, R. W. & CORP, N. (2004). Neocortex size predicts deception rate in primates. *Proceedings of the Royal Society, London* **271B**, 1693–1699.
- BYRNE, R. W. & WHITEN, A. (1988). Machiavellian intelligence. In *The Machiavellian Intelligence Hypothesis* (eds A. WHITEN and R. W. BYRNE), pp. 1–23. Oxford University Press, Oxford.
- BZDOK, D. & DUNBAR, R. I. M. (2020). The neurobiology of social distance. *Trends in Cognitive Sciences* **24**, 717–734.
- BZDOK, D. & DUNBAR, R. I. M. (2022). Social isolation and the brain in the pandemic era. *Nature Human Behaviour* **6**, 1333–1343.
- CALHIM, S., SHI, J. & DUNBAR, R. I. M. (2006). Sexual segregation among feral goats: testing between alternative hypotheses. *Animal Behaviour* **72**, 31–41.
- CAMERON, E. Z., SETSAAS, T. H. & LINKLATER, W. L. (2009). Social bonds between unrelated females increase reproductive success in feral horses. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 13850–13853.
- CARLISI, C. O., MOFFITT, T. E., KNOTT, A. R., HARRINGTON, H., IRELAND, D., MELZER, T. R., POULTON, R., RAMRAKHA, S., CASPI, A., HARIRI, A. R. & VIDING, E. (2020). Associations between life-course-persistent antisocial behaviour and brain structure in a population-representative longitudinal birth cohort. *Lancet Psychiatry* **7**, 245–253.
- CARNE, C., SEMPLE, S. & LEHMANN, J. (2015). Investigating constraints on the survival of orangutans across Borneo and Sumatra. *Tropical Conservation Science* **8**, 940–954.
- CARRINGTON, S. J. & BAILEY, A. J. (2009). Are there theory of mind regions in the brain? A review of the neuroimaging literature. *Human Brain Mapping* **30**, 2313–2335.
- CASTLES, M., HEINSOHN, R., MARSHALL, H. H., LEE, A. E., COWLISHAW, G. & CARTER, A. J. (2014). Social networks created with different techniques are not comparable. *Animal Behaviour* **96**, 59–67.
- CHAPMAN, C. A. & CHAPMAN, L. J. (2000). Determinants of group size in primates: the importance of travel costs. In *On the Move: How and Why Animals Travel in Groups* (eds S. BOINSKI and P. A. GARBER), pp. 24–42. University of Chicago Press, Chicago.
- CHARNOV, E. L. (1993). *Life History Invariants*. Oxford University Press, Oxford.
- CHARNOV, E. L. & BERRIGAN, D. (1993). Why do female primates have such long lifespans and so few babies? Or life in the slow lane. *Evolutionary Anthropology* **1**, 191–194.
- CHENEY, D. L., SILK, J. B. & SEYFARTH, R. M. (2016). Network connections, dyadic bonds and fitness in wild female baboons. *Royal Society Open Science* **3**, 160255.
- CLUTTON-BROCK, T. H. (ed.) (1988). *Reproductive Success: Studies of Individual Variation in Contrasting Breeding Systems*. University of Chicago Press, Chicago.
- CLUTTON-BROCK, T. H., GUINNESS, F. E. & ALBON, S. D. (1983). The costs of reproduction to red deer hinds. *Journal of Animal Ecology* **52**, 367–384.
- CLUTTON-BROCK, T. H. & HARVEY, P. H. (1977). Primate ecology and social organization. *Journal of Zoology* **183**, 1–39.
- CLUTTON-BROCK, T. H. & HARVEY, P. H. (1980). Primates, brains and ecology. *Journal of Zoology* **190**, 309–323.
- COLLINS, C. E., AIREY, D. C., YOUNG, N. A., LEITCH, D. B. & KAAS, J. H. (2010). Neuron densities vary across and within cortical areas in primates. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 15927–15932.
- CREELE, S., DANTZER, B., GOYMAN, W. & RUBENSTEIN, D. R. (2013). The ecology of stress: effects of the social environment. *Functional Ecology* **27**, 66–80.
- CRITTENDEN, B. M., MITCHELL, D. J. & DUNCAN, J. (2015). Recruitment of the default mode network during a demanding act of executive control. *eLife* **4**, e06481.
- CRUWYS, T., DINGLE, G. A., HASLAM, C., HASLAM, S. A., JETTON, J. & MORTON, T. A. (2013). Social group memberships protect against future depression, alleviate depression symptoms and prevent depression relapse. *Social Science & Medicine* **98**, 179–186.
- CUNDIFF, J. M. & MATTHEWS, K. A. (2018). Friends with health benefits: the long-term benefits of early peer social integration for blood pressure and obesity in midlife. *Psychological Science* **29**, 814–823.
- DÁVID-BARRETT, T. & DUNBAR, R. I. M. (2013). Processing power limits social group size: computational evidence for the cognitive costs of sociality. *Proceedings of the Royal Society, London* **280B**, 20131151.
- DAVIES, N. B. (1978). Territorial defence in the speckled wood butterfly (*Pararge aegeria*): the resident always wins. *Animal Behaviour* **26**, 138–147.
- DEACON, T. W. (1990). Problems of ontogeny and phylogeny in brain-size evolution. *International Journal of Primatology* **11**, 193–282.
- DEANER, R. O., ISLER, K., BURKART, J. & VAN SCHAIK, C. P. (2007). Overall brain size, and not encephalization quotient, best predicts cognitive ability across non-human primates. *Brain, Behavior and Evolution* **70**, 115–124.
- DEANER, R. O. & NUNN, C. L. (1999). How quickly do brains catch up with bodies? A comparative method for detecting evolutionary lag. *Proceedings of the Royal Society, London* **266B**, 687–694.
- DECASIEN, A. R., WILLIAMS, S. A. & HIGHAM, J. P. (2017). Primate brain size is predicted by diet but not sociality. *Nature Ecology and Evolution* **1**, 0112.
- DEVAINE, M., SAN-GALLI, A., TRAPANESE, C., BARDINO, G., HANO, C., JALME, M. S., BOURE, S., MASI, S. & DAUNIZEAU, J. (2017). Reading wild minds: a computational assay of theory of mind sophistication across seven primate species. *PLoS Computational Biology* **13**, e1005833.
- DINGWALL, H. L., HATALA, K. G., WUNDERLICH, R. E. & RICHMOND, B. G. (2013). Hominin stature, body mass, and walking speed estimates based on 1.5 million-year-old fossil footprints at Ileret, Kenya. *Journal of Human Evolution* **64**, 556–568.
- DOBZHANSKY, T. (1973). Nothing in biology makes sense except in the light of evolution. *American Biology Teacher* **35**, 125–129.
- DUNBAR, R. I. M. (1976). Some aspects of research design and their implications for the observational study of behaviour. *Behaviour* **58**, 78–98.
- DUNBAR, R. I. M. (1982). Adaptation, fitness and the evolutionary tautology. In *Current Problems in Sociobiology* (ed. KING'S COLLEGE SOCIOBIOLOGY GROUP), pp. 9–28. Cambridge University Press, Cambridge.
- DUNBAR, R. I. M. (1983). Life history tactics and alternative strategies of reproduction. In *Mate Choice* (ed. P. P. G. BATESON), pp. 423–433. Cambridge University Press, Cambridge.
- DUNBAR, R. I. M. (1988). *Primate Social Systems*. Chapman & Hall, London.
- DUNBAR, R. I. M. (1989). Reproductive strategies of female gelada baboons. In *Sociobiology of Sexual and Reproductive Strategies* (eds A. RASA, C. VOGEL and E. VOLAND), pp. 74–92. Chapman & Hall, London.
- DUNBAR, R. I. M. (1992). Neocortex size as a constraint on group size in primates. *Journal of Human Evolution* **22**, 469–493.
- DUNBAR, R. I. M. (1993a). Coevolution of neocortex size, group size and language in humans. *Behavioral and Brain Sciences* **16**, 681–735.
- DUNBAR, R. I. M. (1993b). Seeing biology through Aristotle's eyes. *New Scientist* **137**(1861), 39–42.
- DUNBAR, R. I. M. (1995). *The Trouble With Science*. Faber & Faber, London.
- DUNBAR, R. I. M. (1998a). Impact of global warming on the distribution and survival of the gelada baboon: a modelling approach. *Global Change Biology* **4**, 293–304.
- DUNBAR, R. I. M. (1998b). The social brain hypothesis. *Evolutionary Anthropology* **6**, 178–190.
- DUNBAR, R. I. M. (2003). Evolution of the social brain. *Science* **302**, 1160–1161.
- DUNBAR, R. I. M. (2009). Why only humans have language. In *The Prehistory of Language* (eds R. BOTHA and C. KNIGHT), pp. 12–35. Oxford University Press, Oxford.
- DUNBAR, R. I. M. (2014). *Human Evolution*. Pelican, London and Oxford University Press, New York.

- DUNBAR, R. I. M. (2015). Lag effects in primate brain size evolution: a re-evaluation. *BioRxiv* <https://doi.org/10.1101/355453>.
- DUNBAR, R. I. M. (2018). The anatomy of friendship. *Trends in Cognitive Sciences* **22**, 32–51.
- DUNBAR, R. I. M. (2019). *Evolution: What Everyone Needs To Know*. Oxford University Press, New York.
- DUNBAR, R. I. M. (2020). Structure and function in human and primate social networks: implications for diffusion, network stability and health. *Proceedings of the Royal Society, London* **476A**, 20200446.
- DUNBAR, R. I. M. (2022a). Female dispersion is necessary, but not sufficient, for pairbonded monogamy in mammals. *Frontiers in Ecology and Evolution* **10**, 905296.
- DUNBAR, R. I. M. (2022b). Laughter and its role in the evolution of human social bonding. *Philosophical Transactions of the Royal Society of London* **289B**, 20210176.
- DUNBAR, R. I. M. (2023). Structural and cognitive solutions to prevent group fragmentation in primates. *BioRxiv* <https://doi.org/10.1101/2022.12.13.520310>.
- DUNBAR, R. I. M. & BEVER, J. (1998). Neocortex size predicts group size in carnivores and some insectivores. *Ethology* **104**, 695–708.
- DUNBAR, R. I. M., CHEYNE, S., LAN, D., KORSTJENS, A. H., LEHMANN, J. & COWLISHAW, G. (2019). Environment and time as constraints on the biogeographical distribution of gibbons. *American Journal of Primatology* **81**, e22940.
- DUNBAR, R. I. M. & DUNBAR, P. (1980). The pairbond in klipspringer. *Animal Behaviour* **28**, 251–263.
- DUNBAR, R. I. M., KORSTJENS, A. H. & LEHMANN, J. (2009). Time as an ecological constraint. *Biological Reviews* **84**, 413–429.
- DUNBAR, R. I. M. & MACCARRON, P. (2019). Group size as a trade-off between fertility and predation risk: implications for social evolution. *Journal of Zoology* **308**, 9–15.
- DUNBAR, R. I. M., MACCARRON, P. & ROBERTSON, C. (2018a). Tradeoff between fertility and predation risk drives a geometric sequence in the pattern of group sizes in baboons. *Biological Letters* **14**, 20170700.
- DUNBAR, R. I. M., MACCARRON, P. & SHULTZ, S. (2018b). Primate social group sizes exhibit a regular scaling pattern with natural attractors. *Biology Letters* **14**, 20170490.
- DUNBAR, R. I. M., McADAM, M. & O'CONNELL, S. (2005). Mental rehearsal in great apes and humans. *Behavioral Processes* **69**, 323–330.
- DUNBAR, R. I. M. & SHI, J. (2008). Sex differences in feeding activity results in sexual segregation of feral goats. *Ethology* **114**, 444–451.
- DUNBAR, R. I. M. & SHULTZ, S. (2007). Understanding primate brain evolution. *Philosophical Transactions of the Royal Society of London* **362B**, 649–658.
- DUNBAR, R. I. M. & SHULTZ, S. (2017). Why are there so many explanations for primate brain evolution? *Philosophical Transactions of the Royal Society, London* **244B**, 201602244.
- DUNBAR, R. I. M. & SHULTZ, S. (2021a). Social complexity and the fractal structure of social groups in primate social evolution. *Biological Reviews* **96**, 1889–1906.
- DUNBAR, R. I. M. & SHULTZ, S. (2021b). The infertility trap: the fertility costs of group-living in mammalian social evolution. *Frontiers in Ecology and Evolution* **9**, 634664.
- DUNBAR, R. I. M. & SHULTZ, S. (2022). The evolution of self-control. *BioRxiv* <https://doi.org/10.1101/2020.10.26.354852>.
- DUNBAR, R. I. M. & SOSIS, R. (2018). Optimising human community sizes. *Evolution and Human Behaviour* **39**, 106–111.
- DUNCAN, J. (2001). An adaptive coding model of neural function in prefrontal cortex. *Nature Reviews Neuroscience* **2**, 820–829.
- DUNCAN, J., ASSEM, M. & SHASHIDHARA, S. (2020). Integrated intelligence from distributed brain activity. *Trends in Cognitive Sciences* **24**, 838–852.
- DUNCAN, J., SEITZ, R. J., KOLODNY, J., BOR, D., HERZOG, H., AHMED, A., NEWELL, F. N. & EMSLIE, H. (2000). A neural basis for general intelligence. *Science* **289**, 457–460.
- DZIURA, S. L. & THOMPSON, J. C. (2014). Social-network complexity in humans is associated with the neural response to social information. *Psychological Science* **25**, 2095–2101.
- ELLIS, S., FRANKS, D. W., NATTRASS, S., CANT, M. A., WEISS, M. N., GILES, D., BALCOMB, K. C. & CROFT, D. P. (2017). Mortality risk and social network position in resident killer whales: sex differences and the importance of resource abundance. *Proceedings of the Royal Society, London* **284B**, 20171313.
- ESCRIBANO, D., DOLDÁN-MARTELLI, V., CRONIN, K. A., HAUN, D. B. M., VAN LEEUWEN, E. J. C., CUESTA, J. A. & SÁNCHEZ, A. (2022). Chimpanzees organize their social relationships like humans. *Scientific Reports* **12**, 16641.
- FARRIS, S. M. (2016). Insect societies and the social brain. *Current Opinion in Insect Science* **15**, 1–8.
- FEDOROVA, N., EVANS, C. L. & BYRNE, R. W. (2017). Living in stable social groups is associated with reduced brain size in woodpeckers (Picidae). *Biology Letters* **13**, 20170008.
- FINLAY, B. L. & DARLINGTON, R. B. (1995). Linked regularities in the development and evolution of mammalian brains. *Science* **268**, 1578–1584.
- FINLAY, B. L., DARLINGTON, R. B. & NICASTRO, N. (2001). Developmental structure in brain evolution. *Behavioral and Brain Sciences* **24**, 263–278.
- FITZPATRICK, J. L., ALMBRO, M., GONZALEZ-VOYER, A., HAMADA, S., PENNINGTON, C., SCANLON, J. & KOLM, N. (2012). Sexual selection uncouples the evolution of brain and body size in pinnipeds. *Journal of Evolutionary Biology* **25**, 1321–1330.
- FOX, K. C., MUTHUKRISHNA, M. & SHULTZ, S. (2017). The social and cultural roots of whale and dolphin brains. *Nature Ecology & Evolution* **1**, 1699–1705.
- FRÈRE, C. H., KRÜTZEN, M., MANN, J., CONNOR, R. C., BEJDER, L. & SHERWIN, W. B. (2010). Social and genetic interactions drive fitness variation in a free-living dolphin population. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 19949–19954.
- FRIEDMAN, L. & WALL, M. (2005). Graphical views of suppression and multicollinearity in multiple linear regression. *American Statistician* **59**, 127–136.
- GESQUIERE, L. R., ALTMANN, J., ARCHIE, E. A. & ALBERTS, S. C. (2018). Interbirth intervals in wild baboons: environmental predictors and hormonal correlates. *American Journal of Physical Anthropology* **166**, 107–126.
- GITTLEMAN, J. L. (1986). Carnivore brain size, behavioral ecology, and phylogeny. *Journal of Mammalogy* **67**, 23–36.
- GONZALEZ-VOYER, A., WINBERG, S. & KOLM, N. (2009). Distinct evolutionary patterns of brain and body size during adaptive radiation. *Evolution* **63**, 2266–2274.
- GRABOWSKA, M., KOPPERUD, B. T., TSUBOIB, M. & HANSEN, M. T. (2023). Both diet and sociality affect primate brain-size evolution. *Systemic Biology* (in press).
- GRAFEN, A. (1991). Modelling in behavioural ecology. In *Behavioural Ecology: An Evolutionary Approach* (Volume 3, eds J. R. KREBS and N. B. DAVIES), pp. 5–31. Blackwells, Oxford.
- GRUETER, C. C., BISSONNETTE, A., ISLER, K. & VAN SCHAIK, C. P. (2013). Grooming and group cohesion in primates: implications for the evolution of language. *Evolution and Human Behavior* **34**, 61–68.
- HAGER, R., LU, L., ROSEN, G. D. & WILLIAMS, R. W. (2012). Genetic architecture supports mosaic brain evolution and independent brain-body size regulation. *Nature Communications* **3**, 1–5.
- HAMILTON, M. J., MILNE, B. T., WALKER, R. S., BURGER, O. & BROWN, J. H. (2007). The complex structure of hunter-gatherer social networks. *Proceedings of the Royal Society, London* **274B**, 2195–2202.
- HAMILTON, W. D. (1964). The genetical theory of social behaviour. I, II. *Journal of Theoretical Biology* **7**, 1–52.
- HAMPTON, W. H., UNGER, A., VON DER HEIDE, R. J. & OLSON, I. R. (2016). Neural connections foster social connections: a diffusion-weighted imaging study of social networks. *Social Cognitive and Affective Neuroscience* **11**, 721–727.
- HARDIE, J. L. & COONEY, C. R. (2023). Sociality, ecology and developmental constraints predict variation in brain size across birds. *Journal of Evolutionary Biology* **36**, 144–155.
- HARVEY, P. H. & CLUTTON-BROCK, T. H. (1985). Life history variation in primates. *Evolution* **39**, 559–581.
- HARVEY, P. H. & PAGEL, M. D. (1988). The allometric approach to species differences in brain size. *Human Evolution* **3**, 461–472.
- HARVEY, P. H. & PAGEL, M. D. (1991). *The Comparative Method in Evolutionary Biology*. Oxford University Press, Oxford.
- HASSLER, G. W., GALLONE, B., ARISTIDE, L., ALLEN, W. L., TOLKOFF, M. R., HOLBROOK, L. J., BAELE, G., LEMEY, P. & SUCHARD, M. A. (2022). Principled, practical, flexible, fast: a new approach to phylogenetic factor analysis. *Methods in Ecology and Evolution* **13**, 2181–2197.
- HATCHWELL, B. J., SHARP, S. P., SIMEONI, M. & MCGOWAN, A. (2009). Factors influencing overnight loss of body mass in the communal roosts of a social bird. *Functional Ecology* **23**, 367–372.
- HERCULANO-HOUZEL, S., COLLINS, C. E., WONG, P. & KAAS, J. H. (2007). Cellular scaling rules for primate brains. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 3562–3567.
- HERMANN, E., CALL, J., HERNÁNDEZ-LLOREDA, M. V., HARE, B. & TOMASELLO, M. (2007). Humans have evolved specialized skills of social cognition: the cultural intelligence hypothesis. *Science* **317**, 1360–1366.
- HILL, R. A., BENTLEY, A. & DUNBAR, R. I. M. (2008). Network scaling reveals consistent fractal pattern in hierarchical mammalian societies. *Biology Letters* **4**, 748–751.
- HILL, R. A. & DUNBAR, R. I. M. (1998). An evaluation of the roles of predation rate and predation risk as selective pressures on primate grouping behaviour. *Behaviour* **135**, 411–430.
- HILL, R. A. & DUNBAR, R. I. M. (2003). Social network size in humans. *Human Nature* **14**, 53–72.
- HILL, R. A. & LEE, P. C. (1998). Predation risk as an influence on group size in cercopithecoid primates: implications for social structure. *Journal of Zoology* **245**, 447–456.
- HOF, P. R., CHANIS, R. & MARINO, L. (2005). Cortical complexity in cetacean brains. *Anatomical Record A* **287**, 1142–1152.
- HOLT-LUNSTAD, J., SMITH, T. B. & LAYTON, J. B. (2010). Social relationships and mortality risk: a meta-analytic review. *PLoS Medicine* **7**, e1000316.

- HOOPER, R., BRETT, B. & THORNTON, A. (2022). Problems with using comparative analyses of avian brain size to test hypotheses of cognitive evolution. *PLoS One* **17**, e0270771.
- HORVÁTH, K., MARTOS, J., MIHALIK, B. & BÓDIZS, R. (2011). Is the social brain theory applicable to human individual differences? Relationship between sociability personality dimension and brain size. *Evolutionary Psychology* **9**, 244–256.
- HUMPHREY, N. K. (1976). The social function of intellect. In *Growing Points in Ethology* (eds P. P. G. BATESON and R. A. HINDE), pp. 303–317. Cambridge University Press, Cambridge.
- HUXLEY, J. (1942). *Evolution: The Modern Synthesis*. George Allen & Unwin, London.
- HYON, R., CHAVEZ, R. S., CHWE, J. A. H., WHEATLEY, T., KLEINBAUM, A. M. & PARKINSON, C. (2022). White matter connectivity in brain networks supporting social and affective processing predicts real-world social network characteristics. *Communications Biology* **5**, 1–11.
- ISAACS, E. B., FISCHL, B. R., QUINN, B. T., CHONG, W. K., GADIAN, D. G. & LUCAS, A. (2010). Impact of breast milk on intelligence quotient, brain size, and white matter development. *Pediatric Research* **67**, 357–362.
- ISLER, K., KIRK, E. C., MILLER, J. M., ALBRECHT, G. A., GELVIN, B. R. & MARTIN, R. D. (2008). Endocranial volumes of primate species: scaling analyses using a comprehensive and reliable data set. *Journal of Human Evolution* **55**, 967–978.
- IWANO, T. (1991). An ecological and behavioral study of the aye-aye (*Daubentonia madagascariensis*). *African Studies Monographs* **12**, 19–42.
- JANSON, C. H. (1990). Ecological consequences of individual spatial choice in foraging groups of brown capuchin monkeys, *Cebus apella*. *Animal Behaviour* **40**, 922–934.
- JERISON, H. J. (1977). *Evolution of the Brain and Intelligence*. Academic Press, New York.
- JOFFE, T. H. & DUNBAR, R. I. M. (1997). Visual and socio-cognitive information processing in primate brain evolution. *Proceedings of the Royal Society, London* **264B**, 1303–1307.
- JOLICOEUR, P. (1990). Bivariate allometry: interval estimation of the slopes of the ordinary and standardized normal major axes and structural relationship. *Journal of Theoretical Biology* **144**, 275–285.
- JOLLY, A. (1966). Lemur social behavior and primate intelligence: the step from prosimian to monkey intelligence probably took place in a social context. *Science* **153**, 501–506.
- KAMILAR, J. M. & COOPER, N. (2013). Phylogenetic signal in primate behaviour, ecology and life history. *Philosophical Transactions of the Royal Society of London* **368B**, 20120341.
- KANAI, R., BAHRAMI, B., ROYLANCE, R. & REES, G. (2012). Online social network size is reflected in human brain structure. *Proceedings of the Royal Society, London* **279B**, 1327–1334.
- KAPPELER, P. M. & POZZI, L. (2019). Evolutionary transitions toward pair living in nonhuman primates as stepping stones toward more complex societies. *Science Advances* **5**, eaay1276.
- KENDALL, M. G. & STUART, A. (1979). *The Advanced Theory of Statistics. Vol. 2. Inference and Relationships*. Griffin, London.
- KIESOW, H., DUNBAR, R. I. M., KABLE, J. W., KALENSCHER, T., VOGLEY, K., SCHILBACH, L., MARQUAND, A. F., WIECKI, T. V. & BZDOK, D. (2020). 10,000 social brains: sex differentiation in human brain anatomy. *Science Advances* **6**, eaz1170.
- KIM, D. A., BENJAMIN, E. J., FOWLER, J. H. & CHRISTAKIS, N. A. (2016). Social connectedness is associated with fibrinogen level in a human social network. *Proceedings of the Royal Society, London* **283B**, 20160958.
- KING, A. J. & COWLISHAW, G. (2009). All together now: behavioural synchrony in baboons. *Animal Behaviour* **78**, 1381–1387.
- KINGMA, S. A., SANTEMA, P., TABORSKY, M. & KOMDEUR, J. (2014). Group augmentation and the evolution of cooperation. *Trends in Ecology & Evolution* **29**, 476–484.
- KLEIBER, M. (1961). *The Fire of Life. An Introduction to Animal Energetics*. Wiley, New York.
- KOKKO, H., JOHNSTONE, R. A. & CLUTTON-BROCK, T. H. (2001). The evolution of cooperative breeding through group augmentation. *Proceedings of the Royal Society, London* **268B**, 187–196.
- KORSTJENS, A. H., LEHMANN, J. & DUNBAR, R. I. M. (2010). Resting time as an ecological constraint on primate biogeography. *Animal Behaviour* **79**, 361–374.
- KORSTJENS, A. H., LEHMANN, J. & DUNBAR, R. I. M. (2018). Time constraints do not limit group size in arboreal guenons but do explain community size and distribution patterns. *International Journal of Primatology* **39**, 511–531.
- KOTRSCHAL, A., BUECHEL, S. D., ZALA, S. M., CORRAL-LOPEZ, A., PENN, D. J. & KOLM, N. (2015). Brain size affects female but not male survival under predation threat. *Ecology Letters* **18**, 646–652.
- KRONMAL, R. A. (1993). Spurious correlation and the fallacy of the ratio standard revisited. *Journal of the Royal Statistical Society* **156**, 379–392.
- KRUPENY, C., KANO, F., HIRATA, S., CALL, J. & TOMASELLO, M. (2016). Great apes anticipate that other individuals will act according to false beliefs. *Science* **354**, 110–114.
- KUDO, H. & DUNBAR, R. I. M. (2001). Neocortex size and social network size in primates. *Animal Behaviour* **62**, 711–722.
- KWAK, S., JOO, W., YOUM, Y. & CHEY, J. (2018). Social brain volume is associated with indegree social network size among older adults. *Proceedings of the Royal Society, London* **285B**, 20172708.
- LACK, D. (1954). *The Natural Regulation of Animal Numbers*. Clarendon Press, Oxford.
- LAKATOS, I. (1980). *The Methodology of Scientific Research Programmes*. Cambridge University Press, Cambridge.
- LANDE, R. (1979). Quantitative genetic analysis of multivariate evolution, applied to brain: body size allometry. *Evolution* **33**, 402–416.
- LARDEUX-GILLOUX, I. A. (1997). *Social intelligence and dynamics in group-living orang-utans, Pongo pygmaeus pygmaeus*. PhD Thesis: University of London.
- LEHMANN, J. & DUNBAR, R. I. M. (2009). Implications of body mass and predation for ape social system and biogeographical distribution. *Oikos* **118**, 379–390.
- LEHMANN, J., KORSTJENS, A. H. & DUNBAR, R. I. M. (2010). Apes in a changing world – the effects of global warming on the behaviour and distribution of African apes. *Journal of Biogeography* **37**, 2217–2231.
- LEHMANN, J., MAJOLO, B. & MCFARLAND, R. (2016). The effects of social network position on the survival of wild Barbary macaques, *Macaca sylvanus*. *Behavioral Ecology* **27**, 20–28.
- LEWIS, P. A., BIRCH, A., HALL, A. & DUNBAR, R. I. M. (2017). Higher order intentionality tasks are cognitively more demanding. *Social Cognitive and Affective Neuroscience* **12**, 1063–1071.
- LEWIS, P. A., REZAIE, R., BROWNE, R., ROBERTS, N. & DUNBAR, R. I. M. (2011). Ventromedial prefrontal volume predicts understanding of others and social network size. *NeuroImage* **57**, 1624–1629.
- LI, W., MAI, X. & LIU, C. (2014). The default mode network and social understanding of others: what do brain connectivity studies tell us. *Frontiers in Human Neuroscience* **8**, 74.
- LIAO, W. B., LOU, S. L., ZENG, Y. & KOTRSCHAL, A. (2016). Large brains, small guts: the expensive tissue hypothesis supported within anurans. *American Naturalist* **188**, 693–700.
- LINDENFORS, P., WARTEL, A. & LIND, J. (2021). ‘Dunbar’s number’ deconstructed. *Biology Letters* **17**, 20210158.
- LINKLATER, W. L., CAMERON, E. Z., MINOT, E. O. & STAFFORD, K. J. (1999). Stallion harassment and the mating system of horses. *Animal Behaviour* **58**, 295–306.
- LOGAN, C. J., AVIN, S., BOOGERT, N., BUSKELL, A., CROSS, F. R., CURRIE, A., JELBERT, S., LUKAS, D., MARES, R., NAVARRETE, A. F., SHIGENO, S. & MONTGOMERY, S. H. (2018). Beyond brain size: uncovering the neural correlates of behavioral and cognitive specialization. *Comparative Cognition & Behavior Reviews* **13**, 55–89.
- LUKAS, D. & CLUTTON-BROCK, T. H. (2018). Social complexity and kinship in animal societies. *Ecology Letters* **21**, 1129–1134.
- LUKAS, D. & HUCHARD, E. (2014). The evolution of infanticide by males in mammalian societies. *Science* **346**, 841–844.
- MACCARRON, P. & DUNBAR, R. I. M. (2016). Identifying natural grouping structure in gelada baboons: a network approach. *Animal Behaviour* **114**, 119–128.
- MACDONALD, D. W., CAMPBELL, L. A., KAMLER, J. F., MARINO, J., WERHAHN, G. & SILLERO-ZUBIRI, C. (2019). Monogamy: cause, consequence, or corollary of success in wild canids? *Frontiers in Ecology and Evolution* **7**, 341.
- MACE, G. M., HARVEY, P. H. & CLUTTON-BROCK, T. H. (1981). Brain size and ecology in small mammals. *Journal of Zoology* **193**, 333–354.
- MACKINNON, J. (1974). The behaviour and ecology of wild orang-utans (*Pongo pygmaeus*). *Animal Behaviour* **22**, 3–74.
- MACLEAN, E. L., HARE, B., NUNN, C. L., ADDESSI, E., AMICI, F., ANDERSON, R. C., AURELI, F., BAKER, J. M., BANIA, A. E., BARNARD, A. M., BOOGERT, N. J., BRANNON, E. M., BRAY, E. E., BRAY, J., BRENT, J. M., ET AL. (2014). The evolution of self-control. *Proceedings of the National Academy of Sciences of the United States of America* **111**, E2140–E2148.
- MAGUIRE, E. A., GADIAN, D. G., JOHNSRUDE, I. S., GOOD, C. D., ASHBURNER, J., FRACKOWIAK, R. S. J. & FRICK, C. D. (2000). Navigation-related structural change in the hippocampi of taxi drivers. *Proceedings of the National Academy of Sciences of the United States of America* **97**, 4398–4403.
- MAGURRAN, A. E. (1988). *Ecological Diversity and its Measurement*. Princeton University Press, Princeton.
- MAJOLO, B., DE BORTOLI VIZIOLI, A. & SCHINO, G. (2008). Costs and benefits of group living in primates: group size effects on behaviour and demography. *Animal Behaviour* **76**, 1235–1247.
- MARINO, L., CONNOR, R. C., FORDYCE, R. E., HERMAN, L. M., HOF, P. R., LEEFEBVRE, L., LUSSEAU, D., MCGOWAN, B., NIMCHINSKY, E. A., PACK, A. A., RENDELL, L., REIDENBERG, J. S., REISS, D., UHREN, M. D., VAN DER GUHT, E. & WHITEHEAD, H. (2007). Cetaceans have complex brains for complex cognition. *PLoS Biology* **5**, e139.
- MARS, R. B., FOXLEY, S., VERHAGEN, L., JBABDI, S., SALLET, J., NOONAN, M., NEUBERT, F.-X., ANDERSSON, J., CROXSON, P., DUNBAR, R. I. M., KHRAPITCHEV, A., SIBSON, N., MILLER, K. & RUSHWORTH, M. (2016). The extreme capsule fiber complex in humans and macaque monkeys: a comparative diffusion MRI tractography study. *Brain Structure and Function* **221**, 4059–4071.

- MARS, R. B., NEUBERT, F.-X., NOONAN, M. P., SALLET, J., TONI, L. & RUSHWORTH, M. F. S. (2012). On the relationship between the “default mode network” and the “social brain”. *Frontiers in Human Neuroscience* **6**, 189.
- MARTIN, R. D. (1990). *Primate Origins and Evolution*. Chapman and Hall, London.
- MATEJČ, J., KRATOCHVÍL, L., PAVELKOVÁ, Z., PAVELKOVÁ, R. & VOHRALÍK, V. & NEMEC, P. (2016). Absolute, not relative brain size correlates with sociality in ground squirrels. *Proceedings of the Royal Society, London* **283B**, 20152725.
- McFARLAND, R. & MAJOLO, B. (2013). Coping with the cold: predictors of survival in wild Barbary macaques, *Macaca sylvanus*. *Biology Letters* **9**, 20130428.
- MEGUERDITCHIAN, A., MARIE, D., MARAGIOTODI, K., ROTH, M., NAZARIAN, B., ANTON, J.-L. & CLAUDIÈRE, N. (2020). Baboons (*Papio anubis*) living in larger social groups have bigger brains. *Evolution and Human Behavior* **42**, 30–34.
- MINK, J. W., BLUMENSCHINE, R. J. & ADAMS, D. B. (1981). Ratio of central nervous system to body metabolism in vertebrates: its constancy and functional basis. *American Journal of Physiology* **241**, R203–R212.
- MOFFITT, T., CASPI, A., RUTTER, M. & SILVA, P. (2001). *Sex Differences in Antisocial Behaviour: Conduct Disorder, Delinquency, and Violence in the Dunedin Longitudinal Study*. Cambridge University Press, Cambridge.
- MOLERO SAMUELSON, Y., HODGINS, S., LARSSON, A., LARM, P. & TENGSTRÖM, A. (2010). Adolescent antisocial behavior as predictor of adverse outcomes to age 50: a follow-up study of 1,947 individuals. *Criminal Justice and Behavior* **3**, 158–174.
- MONTGOMERY, S. H., CAPELLINI, I., BARTON, R. A. & MUNDY, N. I. (2010). Reconstructing the ups and downs of primate brain evolution: implications for adaptive hypotheses and *Homo floresiensis*. *BMC Biology* **8**, 1–19.
- MOSS, C. J., CROZE, H. & LEE, P. C. (eds) (2011). *The Amboseli Elephants: A Long-Term Perspective on a Long-Lived Mammal*. University of Chicago Press, Chicago.
- NAVARRETE, A. F., BLEZER, E. L. A., PAGNOTTA, M., DE VIET, E. S. M., TODOROV, O. S., LINDENFORS, P., LALAND, K. N. & READER, S. M. (2018). Primate brain anatomy: new volumetric MRI measurements for neuroanatomical studies. *Brain, Behavior and Evolution* **91**, 109–117.
- NEKARIS, A. & BEARDER, S. K. (2007). The Loriform primates of Asia and mainland Africa. In *Primates in Perspective* (eds C. J. CAMPBELL, A. FUENTES, K. MACKINNON, M. PANGER and S. K. BEARDER), pp. 24–45. Oxford University Press, Oxford.
- NELSON-FLOWER, M. J., HOCKEY, P. A. R., O'RYAN, C., RAIHANI, N. J., DU PLESSIS, M. A. & RIDLEY, A. R. (2011). Monogamous dominant pairs monopolize reproduction in the cooperatively breeding pied babbler. *Behavioral Ecology* **22**, 559–565.
- NOONAN, M. P., MARS, R. B., SALLET, J., DUNBAR, R. I. M. & FELLOWS, L. K. (2018). The structural and functional brain networks that support human social networks. *Behavioural Brain Research* **355**, 12–23.
- NUÑEZ, C. M. V., ADELMAN, J. S. & RUBENSTEIN, D. I. (2015). Sociality increases juvenile survival after a catastrophic event in the feral horse (*Equus caballus*). *Behavioral Ecology* **26**, 138–147.
- O'DONNELL, S., CLIFFORD, M. & MOLINA, Y. (2011). Comparative analysis of constraints and caste differences in brain investment among social paper wasps. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 7107–7112.
- OELSCHLÄGER, H. H. (2008). The dolphin brain—a challenge for synthetic neurobiology. *Brain Research Bulletin* **75**, 450–459.
- OPIE, C., ATKINSON, Q., DUNBAR, R. I. M. & SHULTZ, S. (2013). Male infanticide leads to social monogamy in primates. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 13328–13332.
- OPIE, C., ATKINSON, Q., DUNBAR, R. I. M. & SHULTZ, S. (2014). Reply to Lukas & Clutton-Brock: infanticide still drives primate monogamy. *Proceedings of the National Academy of Sciences of the United States of America* **111**, E1675.
- PAPAGEORGIOU, D., CHRISTENSEN, C., GALL, G. E., KLAREVAS-IRBY, J. A., NYAGUTHI, B., COUZIN, I. D. & FARINE, D. R. (2019). The multilevel society of a small-brained bird. *Current Biology* **21**, R1120–R1121.
- PASSINGHAM, R. E. (2021). *Understanding the Prefrontal Cortex: Selective Advantage, Connectivity and Neural Operations*. Oxford University Press, Oxford.
- PASSINGHAM, R. E. & WISE, S. P. (2012). *The Neurobiology of the Prefrontal Cortex: Anatomy, Evolution and the Origin of Insight*. Oxford University Press, Oxford.
- PATTERSON, S. K., SANDEL, A. A., MILLER, J. A. & MITANI, J. C. (2014). Data quality and the comparative method: the case of primate group size. *International Journal of Primatology* **35**, 990–1003.
- PAWLOWSKI, B. P., LOWEN, C. B. & DUNBAR, R. I. M. (1998). Neocortex size, social skills and mating success in primates. *Behaviour* **135**, 357–368.
- PEARCE, E., STRINGER, C. & DUNBAR, R. I. M. (2013). New insights into differences in brain organisation between Neanderthals and anatomically modern humans. *Proceedings of the Royal Society, London* **280B**, 1471–1481.
- PEER, M., HAYMAN, M., TAMIR, B. & ARZY, S. (2021). Brain coding of social network structure. *Journal of Neuroscience* **41**, 4897–4909.
- PÉREZ-BARBERÍA, J., SHULTZ, S. & DUNBAR, R. I. M. (2007). Evidence for intense coevolution of sociality and brain size in three orders of mammals. *Evolution* **61**, 2811–2821.
- PETERS, R. H. (1986). *The Ecological Implications of Body Size*. Cambridge University Press, Cambridge.
- PILLEMER, S., HOLTZER, R. & BLUMEN, H. M. (2017). Functional connectivity associated with social networks in older adults: a resting-state fMRI study. *Social Neuroscience* **12**, 242–252.
- PITNICK, S., JONES, K. E. & WILKINSON, G. S. (2006). Mating system and brain size in bats. *Proceedings of the Royal Society, London* **273B**, 719–724.
- POLLEN, A. A., DOBBERFUHL, A. P., SCACE, J., IGULU, M. M., RENN, S. C., SHUMWAY, C. A. & HOFMANN, H. A. (2007). Environmental complexity and social organization sculpt the brain in Lake Tanganyikan cichlid fish. *Brain, Behavior and Evolution* **70**, 21–39.
- POPPER, K. (1962). *Conjectures and Refutations: The Growth of Scientific Knowledge*. Routledge, London.
- POWELL, J., LEWIS, P. A., ROBERTS, N., GARCÍA-FIÑANA, M. & DUNBAR, R. I. M. (2012). Orbital prefrontal cortex volume predicts social network size: an imaging study of individual differences in humans. *Proceedings of the Royal Society, London* **279B**, 2157–2162.
- POWELL, L. E., ISLER, K. & BARTON, R. A. (2017). Re-evaluating the link between brain size and behavioural ecology in primates. *Proceedings of the Royal Society, London* **284B**, 20171765.
- PRADO, E. L. & DEWEY, K. G. (2014). Nutrition and brain development in early life. *Nutrition Reviews* **72**, 267–284.
- RAYNER, J. M. V. (1985). Linear relations in biomechanics: the statistics of scaling functions. *Journal of Zoology* **286**, 415–439.
- REZNICK, D. & ENDLER, J. A. (1982). The impact of predation on life history evolution in Trinidadian guppies (*Poecilia reticulata*). *Evolution* **36**, 160–177.
- ROBINS, L. N. & RATCLIFF, K. S. (1978). Risk factors in the continuation of childhood antisocial behavior into adulthood. *International Journal of Mental Health* **7**, 96–116.
- ROGELL, B., DOWLING, D. K. & HUSBY, A. (2020). Controlling for body size leads to inferential biases in the biological sciences. *Evolution Letters* **4**, 73–82.
- ROSATI, A. G. (2017). Foraging cognition: reviving the ecological intelligence hypothesis. *Trends in Cognitive Sciences* **21**, 691–702.
- ROSENQUIST, J. N., FOWLER, J. H. & CHRISTAKIS, N. A. (2011). Social network determinants of depression. *Molecular Psychiatry* **16**, 273.
- ROUMAZEILLES, L., EICHERT, N., BRYANT, K. L., FOLLONI, D., SALLET, J., VIJAYAKUMAR, S., FOXLEY, S., TENDLER, B. C., JBABDI, S., REVELEY, C., VERHAGEN, L., DERSHOWITZ, L. B., GUTHRIE, M., FLACH, E., MILLER, K. L. & MARS, R. B. (2020). Longitudinal connections and the organization of the temporal cortex in macaques, great apes, and humans. *PLoS Biology* **18**, e3000810.
- ROUMAZEILLES, L., LANGE, F. J., BENN, R. A., ANDERSSON, J. L., BERTELSON, M. F., MANGER, P. R., KHRAPICHEV, A. A., BRYANT, K. L., SALLET, J. & MARS, R. B. (2022). Cortical morphology and white matter tractography of three phylogenetically distant primates: evidence for a simian elaboration. *Cerebral Cortex* **32**, 1608–1624.
- ROUMAZEILLES, L., SCHURZ, M., LOJKIEWIEZ, M., VERHAGEN, L., SCHÜFFELGEN, U., MARCHE, K., MAHMOUDI, A., EMBERTON, A., SIMPSON, K., JOLI, O., KAHMASSI, M., RUSHWORTH, M. F. S., MARS, R. B. & SALLET, J. (2021). Social prediction modulates activity of macaque superior temporal cortex. *Science Advances* **7**, eab2392.
- RUCKSTUHL, K. E. & KOKKO, H. (2002). Modelling sexual segregation in ungulates: effects of group size, activity budgets and synchrony. *Animal Behaviour* **64**, 909–914.
- RUCKSTUHL, K. E. & NEUHAUS, P. (2002). Sexual segregation in ungulates: a comparative test of three hypotheses. *Biological Reviews* **77**, 77–96.
- RUSHWORTH, M. F., MARS, R. B. & SALLET, J. (2013). Are there specialized circuits for social cognition and are they unique to humans? *Current Opinion in Neurobiology* **23**, 436–442.
- SAKAI, S. T., ARSZNOV, B. M., LUNDRIGAN, B. L. & HOLEKAMP, K. E. (2011). Brain size and social complexity: a computed tomography study in hyaenidae. *Brain, Behavior and Evolution* **77**, 91–104.
- SALLET, J., MARS, R. B., NOONAN, M. P., ANDERSON, J. L., O'REILLY, J. X., JBABDI, S., CROXSON, P. L., JENKINSON, M., MILLER, K. L. & RUSHWORTH, M. F. S. (2011). Social network size affects neural circuits in macaques. *Science* **334**, 697–700.
- SANDEL, A. A., MILLER, J. A., MITANI, J. C., NUNN, C. L., PATTERSON, S. K. & GARAMSZEI, L. Z. (2016). Assessing sources of error in comparative analyses of primate behavior: intraspecific variation in group size and the social brain hypothesis. *Journal of Human Evolution* **94**, 126–133.
- SANTINI, Z., JOSE, P., KOYANAGI, A., MEILSTRUP, C., NIELSEN, L., MADSEN, K., HINRICHSSEN, C., DUNBAR, R. I. M. & KOUSHEDE, V. (2021). The moderating role of social network size in the temporal association between formal social participation and mental health: a longitudinal analysis using two consecutive waves of the Survey of Health, Ageing and Retirement in Europe (SHARE). *Social Psychiatry and Psychiatric Epidemiology* **56**, 417–428.
- SHULTZ, S. & DUNBAR, R. I. M. (2006). Both social and ecological factors predict ungulate brain size. *Proceedings of the Royal Society, London* **273B**, 207–215.

- SHULTZ, S. & DUNBAR, R. I. M. (2007). The evolution of the social brain: anthropoid primates contrast with other vertebrates. *Proceedings of the Royal Society, London* **274B**, 2429–2436.
- SHULTZ, S. & DUNBAR, R. I. M. (2010a). Encephalisation is not a universal macroevolutionary phenomenon in mammals but is associated with sociality. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 21582–21586.
- SHULTZ, S. & DUNBAR, R. I. M. (2010b). Social bonds in birds are associated with brain size and contingent on the correlated evolution of life-history and increased parental investment. *Biological Journal of the Linnean Society* **100**, 111–123.
- SHULTZ, S. & DUNBAR, R. I. M. (2010c). Species differences in executive function correlate with hippocampus volume and neocortex ratio across non-human primates. *Journal of Comparative Psychology* **124**, 252–260.
- SHULTZ, S. & DUNBAR, R. I. M. (2022). Socioecological complexity in primate groups and its cognitive correlates. *Philosophical Transactions of the Royal Society of London* **377B**, 20210296.
- SHULTZ, S. & FINLAYSON, L. V. (2010). Large body and small brain and group sizes are associated with predator preferences for mammalian prey. *Behavioral Ecology* **21**, 1073–1079.
- SHULTZ, S., NOE, R., MCGRAW, S. & DUNBAR, R. I. M. (2004). A community-level evaluation of the impact of prey behavioural and ecological characteristics on predator diet composition. *Proceedings of the Royal Society, London* **271B**, 725–732.
- SHUMWAY, C. A. (2010). The evolution of complex brains and behaviors in African cichlid fishes. *Current Zoology* **56**, 144–156.
- SILK, J. B., ALBERTS, S. C. & ALTMANN, J. (2003). Social bonds of female baboons enhance infant survival. *Science* **302**, 1232–1234.
- SILK, J. B., BEEHNER, J. C., BERGMAN, T. J., CROCKFORD, C., ENGH, A. L., MOSCOVICE, L. R., WITTIG, R. M., SEYFARTH, R. M. & CHENEY, D. L. (2009). The benefits of social capital: close social bonds among female baboons enhance offspring survival. *Proceedings of the Royal Society, London* **276**, 3099–3104.
- SILK, J. B., BEEHNER, J. C., BERGMAN, T. J., CROCKFORD, C., ENGH, A. L., MOSCOVICE, L. R., WITTIG, R. M., SEYFARTH, R. M. & CHENEY, D. L. (2010). Strong and consistent social bonds enhance the longevity of female baboons. *Current Biology* **20**, 1359–1361.
- SIMPSON, E. H. (1951). The interpretation of interaction in contingency tables. *Journal of the Royal Statistical Society* **13**, 238–241.
- SINGLETON, I. & VAN SCHAİK, C. P. (2002). The social organisation of a population of Sumatran orang-utans. *Folia Primatologica* **73**, 1–20.
- SLATKIN, M. & HAUSFATER, G. (1976). A note on the activities of a solitary male baboon. *Primates* **17**, 311–322.
- SMAERS, J. B., DECHMANN, D. K., GOSWAMI, A., SOLIGO, C. & SAFI, K. (2012). Comparative analyses of evolutionary rates reveal different pathways to encephalization in bats, carnivores, and primates. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 18006–18011.
- SMEELE, S. Q. (2023). Using relative brain size as predictor variable: serious pitfalls and solutions. *BioRxiv* <https://doi.org/10.1101/2022.03.15.484414>.
- SMITH, A. R., SEID, M. A., JIMÉNEZ, L. C. & WCISLO, W. T. (2010). Socially induced brain development in a facultatively eusocial sweat bee *Megalopta genalis* (Halictidae). *Proceedings of the Royal Society, London* **277 B**, 2157–2163.
- SMITH, R. J. (2009). Use and misuse of the reduced major axis for line-fitting. *American Journal of Physical Anthropology* **140**, 476–486.
- SOCIAS-MARTÍNEZ, L. & PECKRE, L. (2023). Does sociality affect evolutionary speed? <https://doi.org/10.5281/zenodo.7693687>.
- SOKAL, R. R. & ROHLF, F. J. (2012). *Biometry*. WH Freeman, San Francisco.
- SPAGNA, A., DUFFORD, A. J., WU, Q., WU, T., ZHENG, W., COONS, E. E., HOF, P. R., HU, B., WU, Y. & FAN, J. (2018). Gray matter volume of the anterior insular cortex and social networking. *Journal of Comparative Neurology* **526**, 1183–1194.
- STAFF, R. T., MURRAY, A. D., AHEARN, T. S., MUSTAFA, N., FOX, H. C. & WHALLEY, L. J. (2012). Childhood socioeconomic status and adult brain size: childhood socioeconomic status influences adult hippocampal size. *Annals of Neurology* **71**, 653–660.
- STANBROOK, E., FITZPATRICK, J. L., BALSHINE, S. & SHULTZ, S. (2022). The evolution of monogamy in cichlids and marine reef fishes. *Frontiers in Ecology and Evolution*. <https://doi.org/10.3389/fevo.2022.1045383>.
- STEIN, Z. & SUSSER, M. (1975). Fertility, fecundity, famine: food rations in the Dutch famine 1944/5 have a causal relation to fertility, and probably to fecundity. *Human Biology* **47**, 131–154.
- STEPHAN, H., FRAHM, H. & BARON, G. (1981). New and revised data on volumes of brain structures in insectivores and primates. *Folia Primatologica* **35**, 1–29.
- STERLING, E. J. & MCCRELESS, E. E. (2006). Adaptations in the aye-aye: a review. In *Leurs: Ecology and Adaptation* (eds L. GOULD and M. L. SAUTHER), pp. 159–184. Springer, Berlin.
- STEVENS, J. R. (2014). Evolutionary pressures on primate intertemporal choice. *Proceedings of the Royal Society, London* **281B**, 20140499.
- STRIER, K. B., LEE, P. C. & IVES, A. R. (2014). Behavioral flexibility and the evolution of primate social states. *PLoS One* **9**, e114099.
- SUGARDJITO, J., TE BOEKHORST, I. J. A. & VAN HOOFF, J. A. R. A. M. (1987). Ecological constraints on the grouping of wild orang-utans (*Pongo pygmaeus*) in the Gunung Leuser National Park, Sumatra, Indonesia. *International Journal of Primatology* **8**, 17–41.
- SUTCLIFFE, A. G., DUNBAR, R. I. M., BINDER, J. & ARROW, H. (2012). Relationships and the social brain: integrating psychological and evolutionary perspectives. *British Journal of Psychology* **103**, 149–168.
- SUTCLIFFE, A. G., DUNBAR, R. I. M. & WANG, D. (2016). Modelling the evolution of social structure. *PLoS One* **11**, e0158605.
- TAMARIT, I., CUESTA, J., DUNBAR, R. I. M. & SÁNCHEZ, A. (2018). Cognitive resource allocation determines the organisation of personal networks. *Proceedings of the National Academy of Sciences of the United States of America* **115**, 1719233115.
- TAMARIT, I., SÁNCHEZ, A. & CUESTA, J. A. (2022). Beyond Dunbar circles: a continuous description of social relationships and resource allocation. *Scientific Reports* **12**, 1–11.
- TESTARD, C., BRENT, L. J., ANDERSSON, J., CHIOU, K. L., NEGRON-DEL VALLE, J. E., DECASINI, A. R., ACEVEDO-ITHIER, A., STOCK, M. K., ANTÓN, S. C., GONZALEZ, O., WALKER, C. S., FOXLEY, S., COMPO, N. R., BAUMAN, S., RUIZ-LAMBIDES, A. V., ET AL. (2022). Social connections predict brain structure in a multidimensional free-ranging primate society. *Science Advances* **8**, eabl5794.
- TINBERGEN, N. (1963). On aims and methods of ethology. *Zeitschrift für Tierpsychologie* **20**, 410–433.
- TREMBLAY, R. E., PIHL, R. O., VITARO, F. & DOBKIN, P. L. (1994). Predicting early onset of male antisocial behavior from preschool behavior. *Archives of General Psychiatry* **51**, 732–739.
- TSUBOI, M., HUSBY, A., KOTRSCHAL, A., HAYWARD, A., BUECHEL, S. D., ZIDAR, J., LÖVLIE, H. & KOLM, N. (2015). Comparative support for the expensive tissue hypothesis: big brains are correlated with smaller gut and greater parental investment in Lake Tanganyika cichlids. *Evolution* **69**, 190–200.
- VAN DER BIJL, W. & KOLM, N. (2016). Why direct effects of predation complicate the social brain hypothesis: and how incorporation of explicit proximate behavioral mechanisms might help. *BioEssays* **38**, 568–577.
- VAN HARMELEN, A. L., GIBSON, J. L., ST CLAIR, M. C., OWENS, M., BRODBECK, J., DUNN, V., LEWIS, G., CROUDACE, T., JONES, P. B., KIEVIT, R. A. & GOODYER, I. M. (2016). Friendships and family support reduce subsequent depressive symptoms in at-risk adolescents. *PLoS One* **11**, e0153715.
- VAN OVERWALLE, F. (2009). Social cognition and the brain: a metaanalysis. *Human Brain Mapping* **30**, 829–858.
- VAN SCHAİK, C. P. (1983). Why are diurnal primates living in groups? *Behaviour* **87**, 120–144.
- VAN SCHAİK, C. P. (2000). Social counterstrategies against infanticide by males in primates and other mammals. In *Primate Males: Causes and Consequences of Variation in Group Composition* (ed. P. KAPPELER), pp. 34–54. Cambridge University Press, Cambridge.
- VAN SCHAİK, C. P. & DUNBAR, R. I. M. (1990). The evolution of monogamy in large primates: a new hypothesis and some critical tests. *Behaviour* **115**, 30–62.
- VON DER HEIDE, R., VYAS, G. & OLSON, I. (2014). The social network-network: size is predicted by brain structure and function in the amygdala and paralimbic regions. *Social Cognitive and Affective Neuroscience* **9**, 1962–1972.
- WAKEFIELD, M. L. (2013). Social dynamics among females and their influence on social structure in an East African chimpanzee community. *Animal Behaviour* **85**, 1303–1313.
- WALLER, M. T. (2011). *The ranging behavior of bonobos in the Lomako forest*. PhD Thesis: University of Oregon.
- WARTEL, A., LINDENFORS, P. & LIND, J. (2019). Whatever you want: inconsistent results are the rule, not the exception, in the study of primate brain evolution. *PLoS One* **14**, e0218655.
- WATTS, J., HAMERSLAG, E., SPRUIES, C. & DUNBAR, R. I. M. (2022). Food storage facilitated the emergence of hunter-gatherer professional religious specialists. *Evolutionary Human Sciences* **4**, e17.
- WEST, B., CULBRETH, G., DUNBAR, R. I. M. & GRIGOLINI, P. (2023). Fractal structure of human and primate social networks provide optimum information flow. *BioRxiv*. <https://biorexiv.org/cgi/content/short/2023.02.23.529431v1>
- WEST, B., MASSARI, G. F., CULBRETH, G., FAILLA, R., BOLOGNA, M., DUNBAR, R. I. M. & GRIGOLINI, P. (2020). Relating size and functionality in human social networks through complexity. *Proceedings of the National Academy of Sciences of the United States of America* **117**, 18355–18358.
- WEST, R. J. (2014). The evolution of large brain size in birds is related to social, not genetic, monogamy. *Biological Journal of the Linnean Society* **111**, 668–678.
- WILLEMS, E. P. & HILL, R. A. (2009). A critical assessment of two species distribution models: a case study of the vervet monkey (*Cercopithecus aethiops*). *Journal of Biogeography* **36**, 2300–2312.
- WITTEMYER, G., DOUGLAS-HAMILTON, I. & GETZ, W. M. (2005). The socioecology of elephants: analysis of the processes creating multilayered social structures. *Animal Behaviour* **69**, 1357–1371.

- WITTIG, R. M., CROCKFORD, C., LEHMANN, J., WHITTEN, P. L., SEYFARTH, R. M. & CHENEY, D. L. (2008). Focused grooming networks and stress alleviation in wild female baboons. *Hormones and Behavior* **54**, 170–177.
- WITTIG, R. M., CROCKFORD, C., WELTRING, A., LANGERGRABER, K. E., DESCHNER, T. & ZUBERBÜHLER, K. (2016). Social support reduces stress hormone levels in wild chimpanzees across stressful events and everyday affiliations. *Nature Communications* **7**, 1–8.
- WRANGHAM, R. W. (1980). An ecological model of female-bonded primate groups. *Behaviour* **75**, 262–300.
- YANG, Y. C., BOEN, C., GERKEN, K., LI, T., SCHORPP, K. & HARRIS, K. M. (2016). Social relationships and physiological determinants of longevity across the human life span. *Proceedings of the National Academy of Sciences of the United States of America* **113**, 578–583.
- YOKOYAMA, C., AUTIO, J. A., IKEDA, T., SALLET, J., MARS, R. B., VAN ESSEN, D. C., GLASSER, M. F., SADATO, N. & HAYASHI, T. (2021). Comparative connectomics of the primate social brain. *NeuroImage* **245**, 118693.
- YOUNG, C., MAJOLO, B., HEISTERMANN, M., SCHÜLKE, O. & OSTNER, J. (2014). Responses to social and environmental stress are attenuated by strong male bonds in wild macaques. *Proceedings of the National Academy of Sciences of the United States of America* **111**, 18195–18200.
- ZHOU, W.-X., SORNETTE, D., HILL, R. A. & DUNBAR, R. I. M. (2005). Discrete hierarchical organization of social group sizes. *Proceedings of the Royal Society, London* **272B**, 439–444.

(Received 9 September 2022; revised and accepted 17 March 2023; published online 31 March 2023)