A critique of comparative studies of brain size

Susan D. Healy,*, and Candy Rowe

1 Institute of Evolutionary Biology, School of Biological Sciences, Kings Buildings, West Mains Road, Edinburgh EH9 3JT, UK
2 School of Biology and Psychology, Newcastle University, Henry Wellcome Building, Framlington Place, Newcastle upon Tyne NE2 4HH, UK

In recent years, there have been over 50 comparative analyses carried out in which social or ecological variables have been used to explain variation in whole brain size, or a part thereof, in a range of vertebrate species. Here, we review this body of work, pointing out that there are a number of substantial problems with some of the assumptions that underpin the hypotheses (e.g. what brain size means), with the data collection and with the ways in which the data are combined in the analyses. These problems are particularly apparent in those analyses in which attempts are made to correlate complex behaviour with parts of the brain that carry out multiple functions. We conclude that now is the time to substantiate these results with data from experimental manipulations.

Keywords: neocortex; brain evolution; cranial volume; behavioural complexity; whole brain size

1. INTRODUCTION

Why does brain size vary among species? Although this question can be approached in a number of different ways, evolutionary biologists assume that if natural selection has acted on the behaviour of a species it will also have directed the evolution of the underlying cognitive and neural architecture that control it. Behavioural ecologists have therefore been interested in which ecological factors and life-history traits might correlate with brain size, hoping to identify the most important selection pressures. In the last 10 years, the number of comparative analyses addressing the evolution of brain size has increased substantially. Indeed, there have been more than 50 recent studies correlating brain size with behavioural traits as diverse as migration, deception and female promiscuity (see table 1). However, continuing to add to this body of work will do relatively little to advance our understanding of either brain evolution or function. Instead, there is a need to re-evaluate both the methodology used in this work and the interpretation of the results. In this review, we outline the potential problems by continuing to pursue this research agenda, and then draw attention to methods that we believe will be more profitable for a better understanding of the function and evolution of the brain.

2. THE STUDY OF BRAIN SIZE IN BEHAVIOURAL ECOLOGY: LEARNING FROM THE PAST?

The current trend for using the comparative method to study the evolution of brain size is not new. In the late 1970s and early 1980s, there were a number of comparative studies that attempted to correlate variation in whole brain size with a range of ecological and life-history variables focused on birds, bats and non-human primates, the groups for which there were the most substantial data. There were some clear outcomes from these efforts. For example, diet appeared to explain the variation in brain size in non-human primates (Clutton-Brock & Harvey 1980; Harvey et al. 1980; Mace et al. 1980, 1981) and bats (Eisenberg & Wilson 1978), but not in birds (Bennett & Harvey 1985). A dietary correlate with increasing brain size in mammals is consistent with the speculation that the addition of meat to the diet of our ancestors led to an increase in human brain size (Foley & Lee 1991), although meat eating per se has not had this effect in non-human mammals as far as we can tell. Developmental mode, rather than diet, explained a significant proportion of the variation in whole brain size in birds, with altricial species having larger brains as adults than do precocial species (Bennett & Harvey 1985). This work suggested that different ecological or life-history parameters could explain differences in brain size across taxonomic groups.

However, despite some apparent success in correlating brain size with ecological and life-history variables, the enthusiasm for comparative analyses soon dried up. While this may, in part, have been because all possible correlations were extracted from the data, it also became clear that finding correlates of variation in whole brain size was not that helpful for understanding how brain structures mapped onto behaviour. One of the conclusions drawn was that 'large brains are a result of differential increases in size of parts of the brain, according to their importance' (Mace et al. 1980). Additionally, changes in different brain areas may occur at once, and there could be trade-offs between brain areas that cannot be measured in whole brain size. Therefore, while it was recognized that some ecological and life-history variables were...
Table 1. A chronological summary of the comparative studies published in the last 10 years that have looked for correlations between behavioural complexity and measures of brain size in (a) mammals and (b) birds.

<table>
<thead>
<tr>
<th>Authors</th>
<th>Year</th>
<th>Taxon</th>
<th>Brain Measurement Variable</th>
<th>Proposed Explanatory Variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) mammals</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Barton</td>
<td>1996</td>
<td>primates</td>
<td>neocortex volume</td>
<td>social group size, activity period, frugivory</td>
</tr>
<tr>
<td>Joffe &amp; Dunbar</td>
<td>1997</td>
<td>primates</td>
<td>brain volume, neocortex, area V1, LGN, amygdala volumes</td>
<td>social group size</td>
</tr>
<tr>
<td>Barton</td>
<td>1998</td>
<td>primates</td>
<td>brain, V1, LGN, neocortex masses, and number of neurons in LGN</td>
<td>visual specialization and social group size</td>
</tr>
<tr>
<td>Dunbar &amp; Bever</td>
<td>1998</td>
<td>carnivores, insectivores</td>
<td>neocortex volume</td>
<td>group size</td>
</tr>
<tr>
<td>Pawloski et al.</td>
<td>1998</td>
<td>primates</td>
<td>neocortex volume</td>
<td>social skills, mating success</td>
</tr>
<tr>
<td>Iwaniuk</td>
<td>2001</td>
<td>ground squirrels</td>
<td>cranial volume</td>
<td>mating system</td>
</tr>
<tr>
<td>Kudo &amp; Dunbar</td>
<td>2001</td>
<td>primates</td>
<td>brain and neocortex volume</td>
<td>coalition size, group size, sex differences</td>
</tr>
<tr>
<td>Hutcheon et al.</td>
<td>2002</td>
<td>bats</td>
<td>brain mass, auditory nuclei, olfactory bulb and hippocampus volumes</td>
<td>foraging ecology</td>
</tr>
<tr>
<td>Pellis &amp; Iwaniuk</td>
<td>2002</td>
<td>mammals</td>
<td>non-visual cortex, amygdala</td>
<td>play</td>
</tr>
<tr>
<td>Reader &amp; Laland</td>
<td>2002</td>
<td>primates</td>
<td>executive brain ratio</td>
<td>innovation, social intelligence and tool use</td>
</tr>
<tr>
<td>Changizi</td>
<td>2003</td>
<td>mammals</td>
<td>brain mass</td>
<td>behavioural repertoire size, muscle number</td>
</tr>
<tr>
<td>Barton</td>
<td>2004</td>
<td>primates</td>
<td>brain, neocortex, and visual brain volumes, LGN size</td>
<td>stereoscopic convergence</td>
</tr>
<tr>
<td>Byrne &amp; Corp</td>
<td>2004</td>
<td>primates</td>
<td>brain volume, neocortex volume</td>
<td>tactical deception</td>
</tr>
<tr>
<td>Lefebvre et al.</td>
<td>2004</td>
<td>primates</td>
<td>executive brain size</td>
<td>innovation</td>
</tr>
<tr>
<td>Marino et al.</td>
<td>2004</td>
<td>toothed whales</td>
<td>cranial volume</td>
<td>invasion of the aquatic environment, echolocation</td>
</tr>
<tr>
<td>Lindenfors</td>
<td>2005</td>
<td>primates</td>
<td>brain volume, neocortex volume</td>
<td>female group size (social intelligence)</td>
</tr>
<tr>
<td>Perez-Barberia &amp; Gordon</td>
<td>2005</td>
<td>ungulates</td>
<td>cranial volume</td>
<td>gregariousness (group size), diet, gestation length</td>
</tr>
<tr>
<td>Safi &amp; Dechmann</td>
<td>2005</td>
<td>bats</td>
<td>brain mass, inferior colliculi, auditory nuclei, olfactory bulb and hippocampus volumes</td>
<td>complexity of foraging habitat</td>
</tr>
<tr>
<td>(b) birds</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lefebvre et al.</td>
<td>1997</td>
<td>passerines</td>
<td>forebrain mass</td>
<td>innovation frequency</td>
</tr>
<tr>
<td>Nicolakakis &amp; Lefebvre</td>
<td>2000</td>
<td>birds</td>
<td>forebrain mass</td>
<td>nesting and feeding innovations</td>
</tr>
<tr>
<td>Sol &amp; Lefebvre</td>
<td>2000</td>
<td>birds</td>
<td>forebrain volume, brain volume</td>
<td>behavioural flexibility, invasion success</td>
</tr>
<tr>
<td>Timmermans et al.</td>
<td>2000</td>
<td>birds</td>
<td>hyperstriatum ventrale, neostriatum</td>
<td>feeding innovation, behavioural flexibility</td>
</tr>
<tr>
<td>Iwaniuk &amp; Nelson</td>
<td>2001</td>
<td>waterfowl</td>
<td>cranial volume</td>
<td>foraging mode and diet</td>
</tr>
<tr>
<td>Madden</td>
<td>2001</td>
<td>bowerbirds</td>
<td>cranial volume</td>
<td>bower complexity</td>
</tr>
<tr>
<td>Garamszegi et al.</td>
<td>2002</td>
<td>birds</td>
<td>brain mass</td>
<td>prey capture method, activity period, amphibious feeding tool use</td>
</tr>
<tr>
<td>Lefebvre et al.</td>
<td>2002</td>
<td>birds</td>
<td>brain mass, cranial volumes, neostriatum volumes</td>
<td>behavioural flexibility, invasion success</td>
</tr>
<tr>
<td>Sol et al.</td>
<td>2002</td>
<td>birds</td>
<td>brain mass, cranial volumes</td>
<td>developmental mode</td>
</tr>
<tr>
<td>Iwaniuk &amp; Nelson</td>
<td>2003</td>
<td>birds</td>
<td>brain mass, cranial volume</td>
<td>species richness, extinction risk</td>
</tr>
<tr>
<td>Nicolakakis et al.</td>
<td>2003</td>
<td>birds</td>
<td>brain mass</td>
<td></td>
</tr>
</tbody>
</table>

(Continued.)
<table>
<thead>
<tr>
<th>authors</th>
<th>year</th>
<th>taxon</th>
<th>brain measurement variable</th>
<th>proposed explanatory variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beauchamp &amp; Fernández-Juricic</td>
<td>2004</td>
<td>birds</td>
<td>forebrain size</td>
<td>social complexity (flock size), propensity to flock</td>
</tr>
<tr>
<td>Garamszegi &amp; Eens</td>
<td>2004</td>
<td>birds</td>
<td>brain mass, telencephalon and hippocampus volume</td>
<td>food hoarding</td>
</tr>
<tr>
<td>Iwaniuk &amp; Arnold</td>
<td>2004</td>
<td>birds</td>
<td>brain mass, cranial volume</td>
<td>co-operative breeding</td>
</tr>
<tr>
<td>Iwaniuk et al</td>
<td>2004</td>
<td>birds</td>
<td>brain mass, cranial volume</td>
<td>flightlessness</td>
</tr>
<tr>
<td>Lefebvre et al</td>
<td>2004</td>
<td>birds</td>
<td>neostriatum and hyperstriatum ventrale volume</td>
<td>innovation</td>
</tr>
<tr>
<td>Winkler et al</td>
<td>2004</td>
<td>birds</td>
<td>skull height</td>
<td>migration, behavioural flexibility</td>
</tr>
<tr>
<td>Day et al</td>
<td>2005</td>
<td>bowerbirds</td>
<td>brain mass, telencephalon, hippocampus and cerebellum volume</td>
<td>bower complexity</td>
</tr>
<tr>
<td>Garamszegi et al</td>
<td>2005a</td>
<td>birds</td>
<td>brain mass</td>
<td>extra-pair paternity, feeding innovation</td>
</tr>
<tr>
<td>Garamszegi et al</td>
<td>2005b</td>
<td>birds</td>
<td>brain mass</td>
<td>food storing and winter temperature</td>
</tr>
<tr>
<td>Garamszegi &amp; Lucas</td>
<td>2005</td>
<td>birds</td>
<td>hippocampus volume</td>
<td>development, locomotor behaviour, prey capture, cognitive ability</td>
</tr>
<tr>
<td>Iwaniuk &amp; Hurd</td>
<td>2005</td>
<td>birds</td>
<td>cerebrotype</td>
<td></td>
</tr>
<tr>
<td>Kalisinska</td>
<td>2005</td>
<td>ducks</td>
<td>brain mass, optic tectum, cerebellum, brain stem, hemispheres</td>
<td>foraging mode, diet composition</td>
</tr>
<tr>
<td>Møller et al</td>
<td>2005</td>
<td>birds</td>
<td>brain mass</td>
<td>immune defence organ size</td>
</tr>
<tr>
<td>Schultz et al</td>
<td>2005</td>
<td>birds</td>
<td>brain mass</td>
<td>population declines, resource specialization, complex behaviours</td>
</tr>
<tr>
<td>Sol et al</td>
<td>2005a</td>
<td>passerines</td>
<td>brain mass, cranial volume</td>
<td>migration and innovation</td>
</tr>
<tr>
<td>Sol et al</td>
<td>2005b</td>
<td>birds</td>
<td>brain mass</td>
<td>innovation propensity, invasion success</td>
</tr>
<tr>
<td>Sultan</td>
<td>2005</td>
<td>birds</td>
<td>cerebellar lobular length</td>
<td>visually guided beak behaviour</td>
</tr>
<tr>
<td>Iwaniuk et al</td>
<td>2006a</td>
<td>birds</td>
<td>mesencephalics lateralis, pars dorsalis</td>
<td>echolocation, vocal learning, auditory localization</td>
</tr>
<tr>
<td>Iwaniuk et al</td>
<td>2006b</td>
<td>birds</td>
<td>cerebellar foliation</td>
<td>developmental mode, cognitive ability</td>
</tr>
</tbody>
</table>

Notes:
- a, b, c: Used as a substitute measure for the neostriatum/hyperstriatum
- d: Estimated by Harvey et al. (1987) by several regression analyses converting brain mass to cranial capacity to actual brain tissue volume and then into neocortex ratio. Separate regressions are required for the two suborders for which this was done (strepsirrhine and anthropoid).
- e: Brain mass was calculated from formalin fixed birds, and brain mass was measured. From e.g. Crile & Quiring (1940). Armstrong & Bergeron (1985) and Boire (1989).
- g: Data collected from museum specimens, or taken from Cobb (1968), Boire (1989) and Barker (2000). Where brain mass was not available, it was calculated from cranial volumes or estimated from telencephalon measurements using linear regression.
- i: Supplementary data including full techniques available at: www.current-biology.com/cgi/content/full/15/17/491/DC1.
significantly correlated with variation in whole brain size, there was no clear idea either about what whole brain size actually meant biologically or how changes in size resulted in changes in behaviour.

This realization led to analyses that focused on variation in the size of specific parts of the brain to which a particular function could be ascribed. This work dealt almost exclusively with the hippocampus in birds (e.g. Krebs et al. 1989; Sherry et al. 1989; Healy et al. 1996; Reboreda et al. 1996) and the visual and olfactory areas in primates (Barton et al. 1995; Barton 1998; Barton & Harvey 2000). More substantially, these analyses led to experimental investigation of what cognitive advantages might be conferred by the enlargement of a specific part of the brain. This allowed testing of specific hypotheses, such as the role of the size of the hippocampus in spatial learning abilities in birds (e.g. Hampton & Shuttleworth 1996; Biegler et al. 2001), and heralded a movement away from imprecise terminology and rather vague hypotheses based upon general ‘foraging method’ or ‘social structure’ (e.g. Clutton-Brock & Harvey 1980).

We are therefore surprised that against this historical background, there has been a recent deluge of comparative analyses of variation in brain size in relation to the evolution of ‘complex’ behaviours. New, larger datasets have been brought to bear on brain size variation across a much wider taxonomic range than before, including birds, bats, primates, cetaceans, ungulates, ground squirrels, mammals more generally, and even insects (e.g. Egan & Funk 2006). This reawakened enthusiasm is concomitant with recent suggestions that organisms possess much more highly developed cognitive abilities than previously thought, e.g. episodic-like memory in scrub jays (Aphelecomia coerulescens; Clayton & Dickinson 1998), tool manufacture in crows (Hunt & Gray 2004) and teaching in ants (Franks & Richardson 2006). It is perhaps this work that has prompted interest in discerning relationships between brain size and behavioural complexity, and has led to the now bewildering array of correlations between brain size and behavioural traits. We find them bewildering in the sense that there seem to be almost as many hypotheses as there are studies, with little or no attempt to integrate the diverse results into a coherent scientific framework.

We feel that it is now time to review these studies critically. First, we examine the validity and value of using whole brain size, or large, multi-functional parts of the brain, in comparative analyses. We question, both how these kinds of studies can help us understand the links between the evolution of behaviour and the brain, as well as the methods that are commonly used to assess brain size. Second, we assess what is meant by ‘behavioural complexity’, how it has been measured and review the correlations that have been found between aspects of complex behaviour and brain size. Some of these correlations complement each other, while others are contradictory, but there has not been any attempt to resolve these issues. Finally, we evaluate the successes of using the comparative method in studying the evolution of brain, and where this current body of research is taking us. Our aim is to highlight the potential pitfalls of using the comparative method to address questions concerning variation in brain size and to suggest some future directions.

3. BRAIN SIZE
(a) What is it?

The implicit assumption behind attempts to explain the existing variation in brain size among species is that any increase in size must provide some increase in function. Most of the comparative studies that have been published in the last 10 years are attempts to correlate various complex cognitive behaviours with variation in whole brain size (see table 1 for summary), not least because whole brain size appears relatively easy to measure in many animals. However, the apparent ease of acquiring relatively large datasets seems to have allowed a degree of complacency in our thinking about the evolution of the brain, and what it means to have a larger brain. This lack of rigour is highlighted by the use of vague terminology such as ‘brain space’ (Garamszegi & Ens 2004a), ‘neural space’ (Garamszegi et al. 2005a) and ‘cognitive potential’ (Byrne & Corp 2004). What do any of these terms actually mean? Even if we could easily measure whole brains, how is whole brain size in any way a valid and useful measure of behavioural complexity? These key questions need to be addressed in order to attempt to understand the results of comparative studies correlating variation in brain size and behavioural traits.

Although there is debate over the degree of modularity of function within the brain, most would agree that the brain has many different functional areas, some dedicated purely to sensory processing, others to motor control, to cognition and so on. There are also at least two models for the way in which these brain parts may evolve: (i) mosaic evolution, in which parts of the brain that share a function change together and those parts that differ in function change independently (Barton & Harvey 2000), and (ii) coordinated size change, in which multiple, functionally unrelated parts change together, as a result of the way in which the brain develops (Finlay & Darlington 1995). Importantly, it is far from guaranteed that significant changes to one, or even several, part(s) of the brain can be detected by measuring the whole brain, and yet it is with the functionality of these parts that we are concerned and it is their size on which natural selection appears to act. A clear example of this is a recent experimental finding that while Higher Vocal Centre (HVC) size (a brain region that plays a significant and integral role in song production in passerines) does correlate with song complexity, measurements of whole brain size do not (Spencer et al. 2005). In sum, we (and others, e.g. Rogers 2004) do not believe that whole brain size alone is a useful variable to measure when considering the role of selection in the relationship between brain and behaviour. Furthermore, it cannot help cognitive ecologists and neuroethologists identify relevant regions of the brain underpinning the behaviour in which they are interested.

However, there is considerable merit in examining variation in size of parts of the brain, since at least some of those parts can be readily assigned a function. Perhaps the clearest examples of this are the sensory and motor regions, and in fact the most convincing evidence for natural selection impacting on brain structures has come from the work correlating ecological variables and sensory structures (e.g. Barton 1998, 2004, 2006; Iwaniuk et al. 2006a).
There are, however, substantially more studies investigating possible relationships between complex behaviours (e.g. innovation, social intelligence, social complexity, tool use and female promiscuity) and brain areas associated with ‘higher’ cognitive function, such as the neocortex (now called the isocortex), the cerebellum and the forebrain (e.g. Lefebvre et al. 1997; Reader & Laland 2002; Pitnick et al. 2006; Shultz & Dunbar 2006, see table 1). These kinds of studies are certainly a significant advance on those using whole brain size, since they start to focus on areas of the brain likely to be specifically involved in the production of a particular behaviour. However, even this is not as straightforward as it may seem as often these higher cognitive function areas (e.g. the neocortex/isocortex) are involved in the production of multiple behaviours, not just those variables targeted in a particular analysis. Additionally, it is rare for there to be any direct evidence linking these structures to the behaviour in question. In table 1, we list a number of studies in which a correlation was predicted between variation in the size of a particular part of the brain with variation in a different behaviour, many of which are based on the same brain data. In most cases, significant correlations have been found, leading to the implication that each of these complex behaviours has independently driven significant changes in a particular brain region. While this may, indeed, be the case, some effort should be expended on reconciling these diverse results. For example, neocortex volume in primates has been correlated with social group size, social skills, mating success, tactical deception, social intelligence and more, while in birds, the volume of the forebrain has been correlated with innovation frequency, invasion success, social complexity, food hoarding, bower complexity and population decline (see table 1). In the meantime, considerable caution should be exercised when interpreting correlations between such multifunctional brain regions and complex behaviours, not least owing to the problems inherent in attributing a single function to such a region.

(b) Measuring brain size
Authors of comparative analyses of variation in brain size not only face the issue of defining brain size functionally, they also have to contend with the problem of measuring brain size. Although a handful of the recent studies rely on datasets used in earlier studies on brain size (e.g. Lefebvre et al. 1997; Reader & Laland 2002), there has been a significant increase in the amount of data available over the last 20 years, particularly for birds (e.g. Mlikovsky 1990; Madden 2001; Iwaniuk et al. 2002; Garamszegi & Eens 2004a; Day et al. 2005). However, there have also been significant additions for many mammalian taxa, including bats (Baron et al. 1996), cetaceans (Marino et al. 2006) and ungulates (Pérez-Barberia & Gordon 2005). While these new data have hugely increased the potential for testing adaptive hypotheses for variation in whole brain size and brain structures, it is of concern that they have been collected using a wide range of measurement methods.

For example, one method is to take cranial measurements and to calculate approximate brain mass from these dimensions. But even the ways to measure cranial volume differ. For example, cranial volumes have been estimated from calculations based on data from skulls filled with lead shot (e.g. Iwaniuk & Nelson 2001), from X-rays of skulls to ascertain the width, length and height of the skull (Madden 2001), as well as from CT scans on fossilized specimens (Marino et al. 2006). It is not clear how these measures correlate with each other, but also cranial volume is, at best, an indirect measure of brain mass, and it is debatable as to whether or not it is an accurate surrogate for measuring the whole brain (Iwaniuk & Nelson 2002).

More commonly, brain mass is measured directly. Although this is most accurately done when the brain of a near-dead animal is perfused, fixed in suitable solutions before slicing and the areas of the slices measured, this technique has been used in only one of the recent analyses (bowerbirds: Day et al. 2005). In other studies, brain mass has been measured from dead animals, for example, from frozen specimens (e.g. Garamszegi et al. 2002; Garamszegi & Eens 2004a; Møller et al. 2005) and from museum specimens where the brains have been re-hydrated in formalin (Iwaniuk & Arnold 2004). However, these methods are likely to introduce errors into the measurements, and the reliability of the data becomes questionable, particularly since freezing and dessication will change the water content of the brain tissue, which differs seasonally as it is (Yaskin 1984).

There are, therefore, two problems with this variation in the measurement of brain mass. The first is that the brains of different species or genera may shrink or re-hydrate to different extents, for example, due to the difference in surface area to volume ratio or from having been taken at different times of the year, irrespective of possible seasonal changes (e.g. Smulders et al. 1995). The second problem comes when authors combine the data produced from different studies using different methods, without clearly demonstrating that this is appropriate. This has been done on multiple occasions in order to increase the amount of data (e.g. Kudo & Dunbar 2001; Lefebvre et al. 2002, 2004; Sol et al. 2005b; Iwaniuk et al. 2006a; Shultz & Dunbar 2006). However, different ways of measuring whole brain size do not always result in the same outcome. For example, while cranial measurements correlate with bower complexity (Madden 2001), brain mass does not (Day et al. 2005). These problems are typically explained away by the argument, based on plausibility alone that any error introduced into the dataset will be random, rather than introducing any kind of bias. While this may be true, it seems curious that checking for potential biases is rarely, if ever, carried out, and that data sources are not statistically included in the comparative analysis.

Taken together, there appears to have been a lack of scientific rigour in assembling many of the datasets used in the tests of hypotheses as to why whole brain size varies. The ways in which data have been collected may introduce biases falsely leading to significant correlations, and make negative results difficult to interpret. This is exacerbated when sample sizes are small and datasets are combined from different sources. We need to understand more fully how these brain measurements are related and to identify the methods for brain measurement that provide the most relevant quantity for hypothesis testing.
4. MAKING THINGS MORE COMPLICATED: CORRELATING BEHAVIOURAL COMPLEXITY WITH BRAIN SIZE

Issues with determining the function of brain size (whole or parts thereof) and its measurement constitute only half the problem. The other half concerns the variables with which variation in brain size has been correlated. Almost all of the recent comparative studies have been focused on aspects of complex behaviour, under the assumption that to perform an unusual or cognitively difficult behaviour will require a larger brain. Although behavioural complexity is one of those tricky terms to define, it has been applied to a range of what might be considered cognitively impressive behaviours such as deception, innovation, tool use, elaborate bower building and responding to new environments. There is now an edifice built on comparative analyses that suggests that behavioural complexity is reliably correlated with the size of the whole brain or large multifunctional areas within it. There is, however, a danger that these correlations are then used to reinforce the original implicit assumption that complex cognition necessarily drives an increase in whole brain size, not least because when one examines the correlations in more detail, there are a number of unexplained inconsistencies among them.

Variation in the size of the whole brain, as well as several brain areas, have now been shown to correlate with some form of behavioural complexity (summarized in table 1). Although some physiological and sensory variables, such as diet (Iwaniuk & Nelson 2001) and stereoscopic convergence (Barton 2004) have also been investigated, we restrict our focus here to the literature on behavioural convergence (Barton 2004) have also been investigated, we restrict our focus here to the literature on behavioural complexity, considering these in the following section. Given the wide array of explanatory variables, we have classified them into three main groups of hypotheses: social complexity and intelligence; innovation or behavioural flexibility; and sexual selection and mate choice.

(a) Social complexity and intelligence

Social complexity has been measured in number of ways. Some authors have considered the number of animals living together in a group as a surrogate for social complexity. These include the degree of co-operative breeding in birds (Iwaniuk & Arnold 2004), flock size outside the breeding season/proptensity to flock (Beauchamp & Fernandez-Juricic 2004), group size in primates (Barton et al. 1996; Lindenfors 2005) and ‘gregariousness’ in ungulates (Pérez-Barberia & Gordon 2005). The assumption appears to be that complexity arises when an increasing number of individuals are gathered together. The other method for measuring social complexity is data-mining the literature for mentions of complex behaviour. This has been used predominantly for the primate analyses, for instance with respect to tactical deception (Byrne & Corp 2004) or innovation and tool use (Reader & Laland 2002). This method of data collection runs the risk of publication bias, through the availability or popularity of certain study species, but does at least attempt to capture the behaviour of each species.

Despite earlier authors finding no relationship between measures of social complexity and whole brain size (e.g. Clutton-Brock & Harvey 1980), a significant number have been reported in the recent studies. Social complexity is correlated with enlarged brains in primates (Kudo & Dunbar 2001; Byrne & Corp 2004; Barrett & Henzi 2005; Lindenfors 2005) and in ungulates (Pérez-Barberia & Gordon 2005; Shultz & Dunbar 2006), but does not correlate with variation either in whole brain size, or forebrain size, in birds (Beauchamp & Fernandez-Juricic 2004; Iwaniuk & Arnold 2004). There are several possible reasons for these disparate results. First, it may be that the social bonds or group structure are very different between mammals and birds. Second, it may be that some of these correlations are confounded by life-history and ecological traits that have not been included in the analyses. Third, it may be that the definitions of social complexity need closer examination and clearer definition. We, therefore, need to be cautious in interpreting these results.

There is also a related phenomenon that should be considered here, which is social learning, where animals learn from the behaviour of a conspecific. Social learning in primates is positively correlated with brain size or at least with variation in the ‘executive brain ratio’ (Reader & Laland 2002). This ratio is calculated by dividing the combined sum of the neocortex and striatum volumes by the sum of the volumes of the mesencephalon and the medulla oblongata. This ratio is thought to reflect the relative size of higher cognitive or executive areas in the brain, identifying areas that might be under selection to produce complex behaviours (although for experimental data that do not concur with this view, see Whiting & Barton 2003; Ramnani 2006). However, there is no empirical justification to link these particular brain regions to social learning, and to some degree at least, the choice of brain areas are likely to have been dictated by the data to be found in Stephan et al. (1981) and Zilles & Rehkämper (1988). Additionally, if as Clark et al. (2001) suggest, relative striatum size is invariant across mammals, it is not clear why this area comprises part of the executive brain ratio calculation. Thus, while this is a valid attempt to investigate the relative enlargement of subcomponents of the brain, it should be remembered that these large subcomponents of the brain are not responsible for producing single behaviours. Indeed, as we turn to consider innovation and behavioural complexity, we find that variation in size of these same regions is also correlated with other complex behaviours. We are faced, then, with two approaches that should be considered together, one, multiple targets for selection (i.e. multiple parts of the brain contributing to a single behaviour); and two, multiple sources of selection (i.e. multiple behaviours influencing a single part of the brain).

(b) Innovation and behavioural flexibility

The main method for collecting data on innovation and behavioural complexity is to collate descriptions from the literature of innovative behaviours. These are labelled as being ‘innovative’ if the authors of the original papers included descriptors such as ‘novel’, ‘never seen before’ and so on (beginning with Lefebvre et al. 1997). Of course, since innovation in all this work is concerned with foraging, these accounts might be biased by the visibility of the species, the species’ distribution in relation to human populations, and also the energetic needs of the species. Indeed, some of the innovations given as examples suggest hunger as a driving force (e.g. the first report of a rook eating frozen human vomit: Lefebvre et al. 1997). In birds, innovation frequency correlates with enlargement of the whole brain (Lefebvre et al. 2004), but more especially with
the forebrain (Lefebvre et al. 1997; Nicolakakis & Lefebvre 2000), and two subregions within the forebrain, the neostriatum and hyperstriatum ventrale (Timmermans et al. 2000). Although not defined as an executive brain region, it is often considered that the forebrain is where most of the complex computation occurs within a bird brain (e.g. Lefebvre et al. 2004). Higher rates of innovation are also correlated with enlarged executive brain size in primates (Reader & Laland 2002), although the frequency with which they innovate also correlates with the frequency of instances of social learning.

Innovative birds, with their relatively larger brains, are also much less likely to be migrants (Winkler et al. 2004; Sol et al. 2005b). This correlation has been interpreted as supporting the hypothesis that those species that cannot respond to changes in food conditions that occur at the end of the summer by being innovative are forced to leave the breeding grounds (Sol et al. 2005b). While an interesting idea, one might also think that migrants could also face significant changes in their diet and foraging mode as they travel between continents. Indeed, given the likely importance of spatial cognition in migration, one might have predicted that migrating birds should have larger brains. This would be consistent with the finding that spatial cognition involved in food storing appears to have driven an increase in whole brain size (Garamszegi & Eens 2004b). Quite what is going on with migration, spatial cognition and whole brain size is unclear: whole brain size changes were not correlated with food hoarding in earlier studies (Krebs et al. 1989; Sherry et al. 1989) and brain size decreases significantly in migrants as they age (Healy et al. 1996). Is it possible that the relationship between brain size and migration is due to older birds being included in the recent migrant data? Other, perhaps more parsimonious factors and constraints on migrants, such as the energetics of having a large brain and bill morphology (Winkler et al. 2004), should also be considered before concluding that these changes are due to behavioural flexibility.

One final extension of the behavioural flexibility and innovation hypothesis is that birds with relatively larger brains enjoy better survival if subject to environmental change. With regard to occurrences of translocation by humans, birds with relatively larger brains are more likely to be successful at establishing themselves in their new environment than are smaller-brained, less behaviourally flexible species (Sol & Lefebvre 2000; Sol et al. 2002, 2005a). There is a similar effect with man-made environmental change: populations of larger-brained species are apparently experiencing less steep declines in the UK compared with those of smaller-brained species (Shultz et al. 2005). A consistent pattern then? Well, not quite. In contrast, Nicolakakis et al. (2003) and Shultz et al. (2005) found that while enlarged brains and behavioural flexibility were correlated with species richness, there was no correlation between brain size and threat of extinction. This difference in outcome is rather surprising as it appears that both groups used the same brain data in their analyses. Thus, although these correlations are interesting, we would be cautious about interpreting these findings as evidence for selection for larger brains and behavioural complexity without further supporting empirical evidence.

Given this apparent relationship between innovation and enlarged brains, it comes as little surprise that there are also positive correlations between the neostriatum size or whole brain size and tool use in birds (Lefebvre et al. 2002), and a similar relationship between executive brain size and tool use in primates (Reader & Laland 2002). Descriptions of tool use for these analyses have been gleaned from the literature and often appear to be correlated with other complex behaviours (Reader & Laland 2002). Disentangling what behaviour and which brain areas might be relevant in evolutionary terms is again problematic.

A small twist in the tale comes with regard to which parts of the brain are involved in complex behaviour. Day et al.’s (2005) work on bower building showed a correlation between the complexity of the bowers built by bowerbirds (the complexity score based on the number and size of the walls or maypoles, and the number, diversity and orderliness of the arrangement of the decorations) and the size of the cerebellum, a part of the brain more usually associated with motor skills, rather than with complex cognition. This result leaves us with the entirely plausible possibility that the production of complex behaviours may rely on complex motor skills as well as complex cognition and clearly illustrates the problem of focusing on particular regions and particular behaviours in isolation.

(c) Sexual selection and mate choice

Our final ‘complexity’ comes from those studies in which sexual selection has been posited to have a role in brain enlargement. In bats, there appears to be a correlation between variation in brain size and mating system: intriguingly, species in which females show mate fidelity have larger brains than do those species in which the females are promiscuous (Pitnick et al. 2006), which they suggest is due to sexual antagonistic coevolution. This occurs because both sexes are ‘under selection to subvert the reproductive investment made by their sexual partners’. As the males of these latter species have larger testes, Pitnick et al. (2006) even suggest that there may be a trade-off between brain size and testes size in males. Interesting as this idea may be, these authors do not examine any other ecological and life-history variables in their analysis (such as diet), and we would be wary of putting too much faith in their result.

In birds, Garamszegi et al. (2004a) predicted that they would find a relationship between song complexity and relative brain size in males. Their prediction was not met, although they did confirm (using a smaller but different dataset) the earlier result of DeVoogd et al. (1993), who found that the volume of the HVC was larger in those species that typically have a larger song repertoire. However, as well as looking for correlations between song complexity and increased whole brain size within each sex, there are also several investigations into sexual dimorphism in whole brain size, with the prediction that the greater the degree of dimorphism, the greater the role of sexual selection. Møller et al. (2005) found a correlation between sexual dimorphism in the brain and relative spleen size, although not with heart or liver size. Although they postulate a role for immune function in the enlargement of the brain in birds, variation in spleen is also correlated with, among other things, sexual
 animals experiments that activational hormone levels, and cognitive areas of the brain in humans and other known that development affects the size of sensory, motor developmental history is not known. And yet, it is well table 1 use the brains from adult animals, for which the of brain size is that there appears to be a presumption that brain size may be traded increased testes size. However, yet another concern arises from analyses examining sexual dimorphism in the brain. As neural sexual dimorphism often occurs as a result of hormonal action during the animal’s lifetime, these data may be obscuring the fact that some parts of the brain are more plastic in males than females (or vice versa), rather than those regions being necessarily enlarged as a direct result of natural or sexual selection (Jacobs et al. 1990; Galea & McEwen 1999). Further comparative analyses will not help us to determine which of these is the case.

(d) Conceptual and methodological issues

Now we have reviewed the literature in more detail, where does this leave us? We see three main problems with this wealth of comparative data on behavioural complexity. First, behavioural complexity is difficult to measure, and authors need to be extremely clear when defining what they regard as behavioural complexity. The second problem is that the number of correlations is already bewildering, and yet there has been no attempt to understand how these factors correlate with each other, or with ecological variables that have already been shown to explain variation in brain size among species. For example, both diet in primates (Clutton-Brock & Harvey 1980; Harvey et al. 1980; Mace et al. 1980, 1981) and developmental mode in birds (Bennett & Harvey 1985; Iwaniuk & Nelson 2003) are reliably found to correlate with variation in brain size, and yet very few of the recent comparative studies include them as factors in the analyses. If this literature is to tell us anything about the potential selection pressures on brain evolution, it is clearly time for multiple factors to be considered in comparative studies, rather than examining a single variable per study of brain evolution. Third, it is important to think about why the brain should be bigger to perform a particular behaviour, and to generate some testable hypotheses. One example of this is in where paternity assurance is postulated to drive the evolution of the brain regions that control ‘sexually dimorphic, complex behaviour’ (Garamszegi et al. 2005b). Empirically this is meaningless, since we cannot identify either the brain region or any target behaviour. We look forward to studies that consider these issues, and consider more fully both their rationale and methodology.

The final significant problem in studying the evolution of brain size is that there appears to be a presumption that brain size is invariant within a lifetime, ignoring any effects of development or experience. All the studies cited in table 1 use the brains from adult animals, for which the developmental history is not known. And yet, it is well known that development affects the size of sensory, motor and cognitive areas of the brain in humans and other animals (Thompson 1993). In addition, it is clear from animal experiments that activational hormone levels, environment, experience and age can affect the size of brain areas (e.g. Brenowitz et al. 1991; Clayton & Krebs 1994; Healy et al. 1996; Rosenzweig & Bennett 1996). More recently, Iwaniuk et al. (2006b) found that pesticide exposure can also lead to a reduction in avian brain size. It is clear, then, that there is a whole host of factors that influence brain size, not just evolutionary selection pressures, and we need to ensure that comparative analyses are not being confounded by developmental and experiential differences among species.

One concern that we have not addressed here but one that we believe should also receive careful review is the methods used for data analysis. Perhaps, the most important issues include the use of multivariate statistical techniques when there is considerable error in the variables that have been measured (a problem that becomes compounded when the predictor variables are themselves correlated), the lack of independent data sets, and the need to control the body size and the different techniques used to do this. The extent of the potential problems and the ways to deal with them are beyond the scope of this review, but there would be considerable merit in such an analysis being carried out, sooner rather than later. The lack of replicability in outcome on the bowerbird data, for example, is of considerable concern and while we have suggested some explanations for the disagreement, it is possible that the data analyses are the cause of the discrepancy.

5. CORRELATIONS BETWEEN BRAIN SIZE AND BEHAVIOURAL COMPLEXITY: WHERE NOW?

There are both benefits and costs to using the comparative method to help to understand how variation in brain size has come about. We have focused very much on the latter, but we do not mean to conclude that this approach is not a useful tool. Far from it, it allows us to address evolutionary questions that are impossible to ask using other methods. However, the relative ease of using this method, especially when applied to apparently readily measured traits such as brain size, can result in analyses being produced that rarely lead to substantive experimental investigation that might allow us to demonstrate causation.

For these comparative analyses to be useful they should form, at the very least, the basis for exploratory experimental analyses. For example, in the instance in which it is proposed that an enlarged forebrain is correlated with innovative propensity, one way of confirming some of these results is to provide experimental confirmation that all, or part of, the forebrain does in fact participate in this behaviour. There are many techniques for exploring this, among which lesions, both permanent and reversible, functional magnetic resonance imaging (e.g. the human left parietal cortex is active when the subject is specifically looking at tools: Chao & Martin 2000) and methods for looking at immediate early gene action come to mind (e.g. Goodson et al. 2005). This suggestion is not a new one since Mace et al. (1980) made it over 25 years ago. It is disappointing, however, that apart from some investigation into the function of the hippocampus in order to determine what an increase in hippocampal volume means in birds, there have been few instances in which Mace et al.’s good advice has been heeded. The paucity of experimental follow-up leads to,
even in the hippocampus case, significant dispute as to the very validity of the comparative (and evolutionarily) approach from psychologists and neuroscientists (e.g. Macphail & Bolhuis 2001), the very last thing that is needed to move our understanding of variation in brain size forward.

One major reason for wanting to substantiate the comparative analyses with experimental data is that many of the correlations are dependent on behavioural definitions that have a tendency to be rather slippery. The most obvious one is that of complexity, by and large meaning capacities we find impressive (e.g. complex social interactions, behaviours such as tool use and so on) but also those about which we find it difficult to be terribly precise. The correlation can all too easily become one in which increases in the size of a part of the brain that ‘deals with’ complex information is correlated with behaviour that is complex. At its least precise, enlargement of the whole brain is correlated with an increase in complex behaviour. This does not contribute much to a discussion of variation in brain size. While there are, of course, quite good mitigating circumstances for some interesting yet hard-to-define behaviours, more effort should be made to reduce the imprecision. After all, if it was generally agreed over 25 years ago that complexity was not a satisfying answer, surely it is not more so now.

We do not pretend to have all the answers to this problem. However, as in other parts of evolutionary biology, an increasingly interdisciplinary approach is likely to yield many new insights and in this case of brain size variation it is highly likely that the psychological and/or the neuroscience literature will provide helpful clues.

6. SUMMARY

In summary, there have been a lot of comparative studies published recently that purport to explain variation in brain size. Many of their authors appear to be unaware of results both from the last great flurry of such publications of 25 years ago, as well as the contemporaneous literature addressing the same kinds of issues (in some cases, exactly the same question). Additionally, many of the studies contain data that may not be adequate for the analyses in which they are included. Finally, the testing of a limited number of predicted correlations would appear to be the raison d'être for the work such that one need not go on to demonstrate causation, in spite of the availability of suitable techniques. We all know that correlation does not demonstrate causation but causation is the context in which the results are invariably interpreted. It would be a shame for evolutionary biologists to misappropriate the comparative method in this way when it can be such a valuable tool.

We would like to thank David Shuker and Alex Thiele for their very helpful discussion and comments on earlier drafts of this manuscript, those authors who kindly supplied reprints, Rob Barton and an anonymous referee for their thoughtful comments, and Julie Harris and Lance Thomson for lending us their house in the garden.

REFERENCES


Cobb, S. 1968 On size of auditory nuclei in some Apodiformes and Caprimulgiformes. Auk 85, 132–133.


Szekely, T., Catchpole, C. K., DeVoogd, A., Marchl, Z., & DeVoogd, T. J. 1996 Evolutionary changes in a song control area of the brain (HVC) are associated with evolutionary changes in song repertoire among European warblers (Sylviidae). *Proc. R. Soc. B* 263, 607–610.


