How the flow and processing of information shapes the cerebrum ☆

Marc H.E. de Lussanet[a,*]

[a]*Psychology and Sport Science, Westf. Wilhelms-Universität Münster, Fliednerstraße 21, 48149 Münster, Germany*

Abstract

The cerebrum of mammals spans a vast range of sizes and yet has a very regular structure. The amount of folding of the cortical surface and the proportion of white matter gradually increase with size, but the underlying mechanisms remain elusive. Here, two laws are derived to fully explain these cerebral scaling relations. The two general laws are derived from the notion that total processing power of the cortex is determined by the total cortical surface (i.e., the number of neurons), whereas the most efficient over-all flow of information is governed by the size of local networks (cortical columns). Since information is transferred by axonal connections which have a definite volume, a trade-off can be formulated from theoretical considerations between local, inter-gyral information transfer and long-range information transfer. It can be shown that this trade-off is governed by a single parameter describing the size of local networks, t_{local}. Despite having just one free parameter, the first law fits the mammalian cerebrum better than any existing function, both across species and within humans. According to the second law, the scaling of white matter volume is also determined by the information principles. It follows that large cerebrums have much local processing and little global information flow. Moreover, paradoxically, a further increase in long-range connections would decrease the efficiency of information flow. These theoretical scaling principles help to compare the cerebrums across mammals regardless their size.

Keywords: Scaling law, Power relation, Evolution, Brain, Volume surface ratio, grey matter volume, white matter volume

1. Introduction

One of the most successful classical concepts of motor control is the speed-accuracy trade-off. It was introduced in the 1950s by Paul Fitts who could explain his results quantitatively from the hypothesis that the trade-off between speed and accuracy of any movement is governed by the maximal rate of information that can flow to control the movement (Fitts, 1954, Fitts and Petterson, 1964). The rate of information is expressed in bits per second (bit/s), so that Fitts proposed that any movement can be expressed with an index of difficulty (bit).

If information flow is a crucial factor in shaping motor control, it is likely to be a central factor too in the evolution of the shape of the central nervous system. Even if this notion may seem trivial, I am not aware it has ever been formulated explicitly. The goal of this work therefore is to test this idea. A most suitable structure to test seems to be the mammalian cerebrum, because its anatomy and function have been extremely well studied, whereas the laws that govern its macroscopic structure are unknown.

The mammalian cerebrum is a highly regular structure, having a cortex of grey matter on its surface which is wrapped around a core of white matter. The cortex is built up of regular layers, which have specific afferent or efferent long-distance connections (Shipp, 2005). The white matter contains long-range axonal connections. The white color comes from myelinization of these axons.

Despite this regular structure, the mammalian cerebrums span a tremendous size range from 11 mm^3 (11 μl, pigmy shrew) to $2.5 \cdot 10^6$ mm^3 (2.5 l, elephant) (Zhang and Sejnowski, 2000), with the brain of the sperm whale being even 60% larger (Marino, 1998). Given these extreme size differences, measures of the cerebrum are usually plotted on a double-logarithmic graph.

There exists a vast literature on interspecific scaling relationships of any thinkable measure of the vertebrate body. What becomes clear on a glance is that such relationships usually show a large variance, whereas narrow distributions are rare. The reason for this is obvious: most scaling relationships are governed by many factors and parameters and thus lead to a trend with much scatter (typically in the order of magnitudes). On the other hand, if one does find a narrow distribution, the chances are good that its singular cause is a simple mechanism.

2

Interestingly, a number of such narrow distributions with little variability around the general trend are known for the mammalian cerebrum (Harrison et al., 2002). These are usually described by a power relation (i.e., a linear relation on a double-logarithmic scale). However, not every
relation that is well described by a linear regression on a double logarithmic scale underlies a power-relation, as pointed out,for example, by Zhang and Sejnowski (2000). These authors investigated the power-relationships between the volume of the cortex, the white matter and the grey matter. Each of the three relationships between these parameters followed a simple power over the entire size range and with remarkably little deviation. These volumetric measures, but also the outer surface and the total cortical surface (including the inward folded surfaces) show power-relations to the brain volume, and the latter also to the cortical volume (Hofman, 1989). However, as the sum of two different powers (e.g., $x^1 + x^{1.5}$ or $x + 1$) is never
a power relation, at least some of the above relations do not truly reflect a power law (Zhang and Sejnowski, 2000). Similar relations with little variability from the trend have been reported for the number of neurons in the cerebrum, the cerebellum and the rest of the brain (Barton and Venditti, 2013, Gabi et al., 2010, Herculano-Houzel, 2011, Herculano-Houzel et al., 2010).

Two most prominent neuro-anatomic properties of the mammalian cerebrum is the division in grey and white matter and the convoluted (folded) surface, which increases gradually with brain size. Whereas the smallest mammalian brains possess a lissencephalic (smooth) cerebral surface, larger cerebrums have ridges (gyri) and folds (sulci). The number and depth of the sulci increase monotonically with brain size. Thus, since the outer shape changes with size, the mammalian cerebrums do not scale isometrically. According to Gross (1999, p. 90), already Franz Joseph Gall (1758-1828) asked why the cerebral cortex is convoluted and proposed that the folds conserve space. Many studies and a number of theories has been proposed since the early approaches (Changizi, 2001, Le Gros Clark, 1945, Mota and Herculano-Houzel, 2012, Prothero and Sundsten, 1984, Van Essen, 1997, Zhang and Sejnowski, 2000). Unfortunately these theories make no or extremely vague quantitative predictions, are circular, or contain serious errors so the problem remains unsolved (see Appendix).

The goal of the present work is to derive scaling laws for the cerebral surface-volume relation and for the white-grey matter volume relation. The laws are derived on the basis of a simple hypothesis, i.e., that the two central

3

functions of the cerebrum, processing and transmission of information, are the forces that shape the cerebrum. In the following it is shown that the volume-outer surface of the cerebrum scales isometrically. Subsequently, theoretical relations for the cerebral cortical surface (including the typical convolutions in large brains) to volume relation as well as for the volume of the white matter are derived.

2. Methods and theoretical background

2.1. Datasets

With the advent of neuroanatomy in the second half of the 19th century, methods have been developed to estimate the surface area of the cortex (Zilles et al., 2013). According to (Jerison, 1982) the first systematic approach was made already by Baillarger (1845), who did measurements of the cortical surface of a number of mammals and proposed a relation between total surface area and intelligence. To my knowledge, the first systematic and accurate measurements of a wide range of mammals have been developed and published by Brodmann and Henneberg (Brodmann, 1913, Henneberg-Neubabelsberg, 1910). A statistical approach for estimating the surface area was developed in the 1960-1970s, whereas modern methods are usually based on neuroimaging techniques.

The latter kind of measurements are likely to be the most accurate, because they are based on fresh, in-vivo measurements and can employ computerized triangulation methods. Probably the best database for the human cortex is by Toro et al. (2008). Unfortunately, similar databases for a range of mammals measured only the ratio of outer to total surface ("gyrification index") without retaining the true surface or volume of the cortex and are thus useless here. Thus, the older measurements had to be used here. The selection criteria for studies were as follows: 1. Systematic usage of a validated method; 2. Reporting of a range of mammals (not just humans); 3. Reporting of at least the cerebral surface area and the cerebral volume and/or the cerebral outer surface area from the same specimen. Seven studies applied to these criteria: (Brodmann, 1913, Elias and Schwartz, 1971, Haug, 1970, Mayhew et al., 1996, Schlenska, 1969, 1974). Finally, the measurements of reported total/outer surface ratio of 1.0 were added from Hofman (1985).

Exclusion criteria were applied as follows: 1. Malformed or diseased brains (Brodmann, 1913); 2. excessively deformed brains (Brodmann suspected excessive shrinkage of two brains of Ganese origin); 3. Estimates

4

(e.g., of cerebral volume) that were based on other specimens or taken from the literature; 4. Measures of cerebral volume that do not fit into a sphere of the reported total cortical surface[1]; in the case of a large enough sample of a single species, the statistical outliers were excluded.

¹¹⁵ The dataset thus 141 samples from 56 species (incl. 29 humans) and 17 orders of mammals. The cerebral volume was available in 97 samples and the outer surface of 104 samples. For the fitting (Section 3.1) a single average per species was calculated: 45 species for the A_c-V_c fit and 51 species for the A_c-A_{out} fit. In either case, the A_c of a species was calculated over the ¹²⁰ samples that were available for the second parameter, so that the mean A_c for a given species could differ depending on the fit. The complete dataset as well as a commented listing of the excluded measurements is provided for downloading (see Supplement; de Lussanet, 2013).

The dataset of human data by Toro et al. (2008) was fitted individually. ¹²⁵ Furthermore, a dataset by Zhang and Sejnowski (2000) was used for fitting the grey and white matter volume.

2.2. Volume-surface relations and sphericity

The relation between the surface area and the volume of a sphere can be calculated from the standard formula:

$$
\begin{aligned}
V_{\text{sphere}} &= \tfrac{3}{4}\pi r^3 \\
S_{\text{sphere}} &= 4\pi r^2 \\
V_{\text{sphere}} &= \sigma A_{\text{sphere}}^{3/2}
\end{aligned}
\tag{1}
$$

where

$$
\sigma = \frac{V_{\text{sphere}}}{S_{\text{sphere}}^{3/2}} = \frac{\tfrac{4}{3}\pi}{(4\pi)^{3/2}} = \frac{1}{6\sqrt{\pi}} \approx 0.094
\tag{2}
$$

$$
\hat{\sigma} = \frac{S_{\text{sphere}}}{V_{\text{sphere}}^{2/3}} = \sqrt[3]{36\pi} \approx 4.84
$$

---

[1]This is a typical problem in the studies that made use of statistical approaches, which most probably led to underestimation of the surface measurements. In order to keep as many measurements as possible, I assumed that the ratio of inner and outer surface area is retained in these measurements.

5

130   The cerebral hemispheres deviate from a perfect sphere, so we can express the cerebral volume V_c as a function of the outer cerebral surface area, S_c:

$$V_c = \sigma s S_c^{3/2} \qquad (3)$$

where the sphericity factor s expresses how much the outer cerebral surface area differs from that of a sphere. Since the hemispheres are not exactly
135   spherical, $s < 1$. It can be computed directly from the outer cerebral surface, S_c, which is reported for part of the dataset:

$$s = \frac{V_c}{\sigma S_c^{3/2}} = 0.56 \pm 0.09 \text{ (mean} \pm \text{standard deviation)}$$

$$\hat{s} = \frac{S_c}{\hat{\sigma} V_c^{2/3}} = 1.51 \pm 0.16 \text{ (mean} \pm \text{standard deviation)}$$

with N = 49 species. However, the mean can only be used if S_c and V_c scale isometrically. To test this, a linear regression on the log-transformed data was performed, which gives $V_c = -1.32 \, S_c^{1.52}$ ($R^2 = 0.998$). Since
140   the power is very close to the $\frac{3}{2}$ power predicted for isometric scaling, and not significantly different (P= 0.16), the sphericity factor s is indeed scale-invariant (the same is true for $\hat{s}$).

2.3. The concept of equivalent thickness

A central observation of the gross anatomy of the cerebrum is that the
145   thickness of the grey-matter cortex tends to saturate with cerebral size. This has typically been quantified by simply dividing the grey matter volume by the cortical surface area (Harman, 1947, Hofman, 1988). This measure is not the average thickness, because it neglects that the cerebrum is convex so that the surface area of the grey-white matter border is smaller than the
150   outer surface (i.e., the equivalent thickness is smaller than the geometric thickness).

The equivalent thickness is a measure of the dimensionality of the cortical surface. For a sphere, the equivalent thickness is proportional to the radius. For a basket filled with laundry, the total surface of the laundry is
155   proportional to the volume of the basket, because the [equivalent] thickness of the laundry is independent of the volume of the basket. Thus, the observation that the thickness of the grey matter seems to saturate with brain

6

size, is an indication that the cortical surface tends to scale proportionally to the cerebral volume for large cerebrums.

This can be expressed formally as:

$$V_c = T_c A_c \tag{4}$$

where T_c is the equivalent cerebral thickness (i.e., the sum of grey and white matter), and A_c the total cortical surface area (i.e., including the cortical area that is located inside the sulci). In this relation, $T_c = {}^{V_c}/{}_{A_c}$ indeed saturates for large cerebrums. (see Fig. S1 of Appendix).

Similarly, the equivalent thickness in terms of the outer surface S_c is:

$$T_s = \frac{V_c}{S_c} = \frac{V_c}{\hat{\sigma}\hat{s}V_c^{2/3}} = \frac{\sqrt[3]{V_c}}{\hat{\sigma}\hat{s}} \tag{5}$$

2.4. Neurons in the cortex

The total cerebral surface is composed of the cortex of grey matter. The cortex is the region where the vast majority of cerebral neurons are located. Since the computational power of the cortex will depend largely on the number of neurons, the distribution of neurons is highly relevant for scaling relations of the cerebrum.

The number of neurons per unit cerebral cortical surface is constant, 10^5 neurons/mm^2. This was first estimated by Bok (1929), and confirmed using cell counts (Carlo and Stevens, 2013, Rockel et al., 1980). In both latter studies, cell counts (neurons and glia cells) were performed for five regions on the cerebral cortex that differ strongly in thickness and composition of the layers. Moreover, these samples were taken from four mammals of very different brain size (mouse, rat, cat, monkey). Both studies found not only the same number of 10^5 neurons per mm^2 for the different samples of each species, but also for each of the four species tested. The average cortical surface per neuron is thus $k = 10^{-5}$ mm^2, irrespective of the cerebral size and the location of the cortical surface.

2.5. Information conduction and axonal diameter

The diameter of axons in the white matter increases with cerebral size (Harrison et al., 2002). It is known that both, the average spike rate and the transmission velocity increase linearly with axonal diameter (Perge et al., 2009, 2012). Based on theoretical considerations, an increase of the signal band width will not increase the information rate in a proportional manner

7

(Balasubramanian et al., 2001, Shannon, 1948). Measurements on the optic
tract have indicated that the higher spiking rate in wide axons also leads to
a stronger correlation of subsequent spikes (Koch et al., 2006). The latter
study estimated that the information rate per axon is independent of its
diameter. Thus, increasing the diameter leads to a quadratic increase of its
cross-sectional area, but only a linear increase in its transduction velocity.

Assuming that the axonal diameter is scaled such transmission time
is approximately size-independent, means that the information rate in a
processing loop will decrease quadratically with the length of the connection.

3. Results

On the basis of the above background, a law can be derived to predict
the scaling of the cerebrum using just a single free parameter. The relations
address the surface-volume relation with respect information flow and local
processing. A second part addresses the grey and white matter volume.

3.1. Scaling law for the cortical surface

The two central functions of the cerebrum are computational power and
information transfer. Assuming that the number of neurons in the cortex
determines computational power, and given that the number of neurons
per cortical surface area is constant, the scaling parameter for the cerebral
volume V_c is cortical surface area, A_c.

There are two components to the information transfer in a neuron: long-
range afferent axonal connections and efferent dendritic trees. The length of
axonal connections in the cerebrum has a typical, gamma-like distribution,
and increases with the size of the cerebrum (Jones and Wise, 1977, Kaiser
et al., 2009). Dendritic trees on the other hand, have a finite size. That
this is indeed so, is reflected by the fact that the number of neurons per cor-
tical surface area is constant (cf. section 2.4), and the established concept
of the cortical column (Mountcastle, 1997, Roerig and Chen, 2002). These
two properties have been described as long-range scale-free connectedness
(Eguíluz et al., 2005, Markov et al., 2013) and small-scale, small-world prop-
erties (Stoop et al., 2013, Watts and Strogatz, 1998). In other words, the
network of the cerebrum has two different scales: the small scale is man-
ifested on the level the cortical columns and the limited size of dendritic
trees, whereas the large scale is manifested in the long-range axonal con-
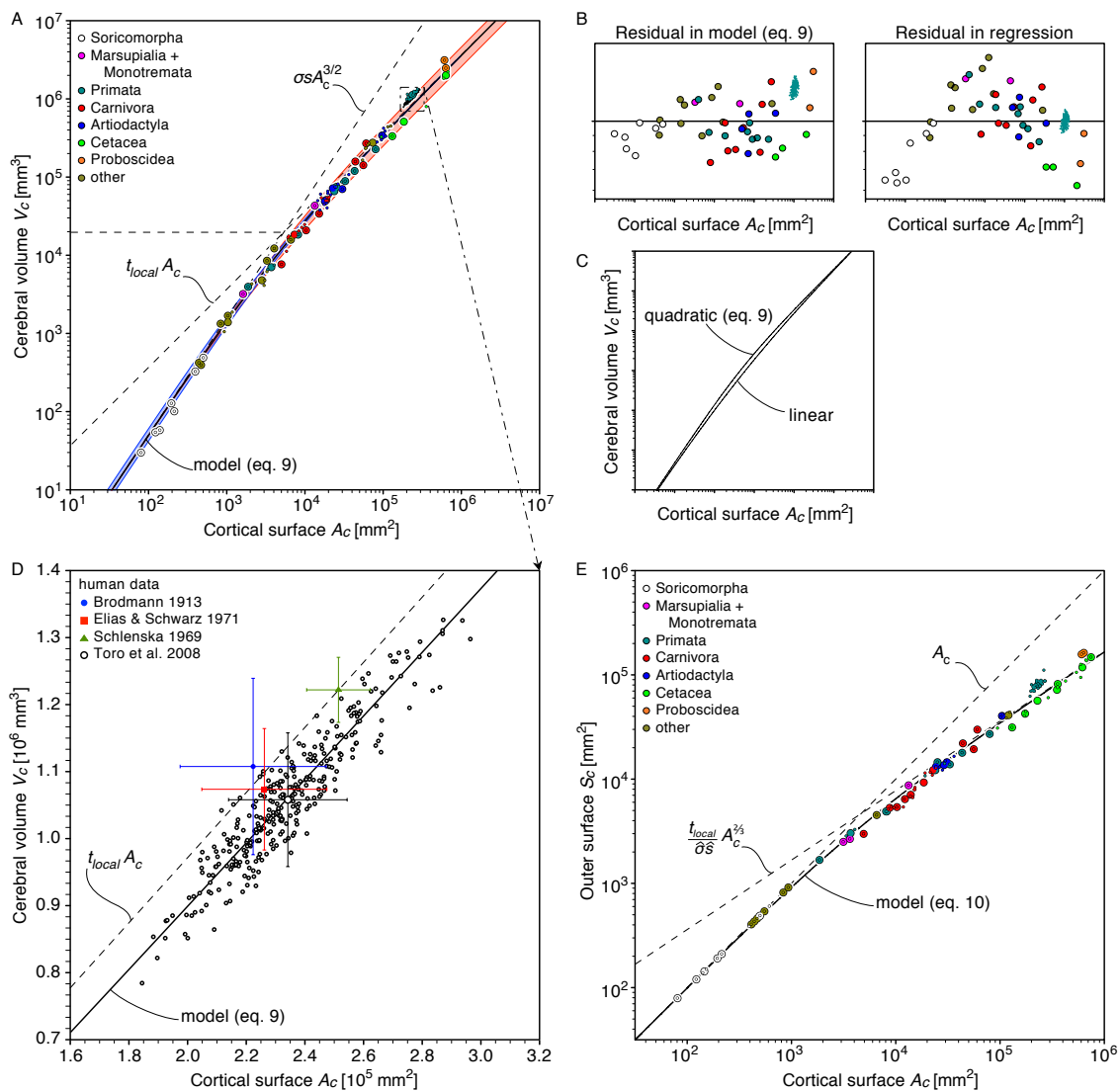nections.

8

Figure 1: (a) The model fit of t_{local} (eq. (7)) on the mean values per species. Continuous line: the model fit; dashed lines: the asymptotic relations with power slopes of 1 (cf. eq. (4)) and 1.5 (cf. eq. (3)). Large dots: species means; small dots: single measurements. Blue-shaded range: the limits of $s \pm 1$ standard deviation (SD); red-shaded range: $t_{local} \pm 1$ mm. (b) Comparison of the distribution of errors for the single-parameter fit of panel a (right) and a conventional two-parameter regression fit (left). The vertical scale measures ± 0.25 order of magnitude. (c) Comparison of the model, equation (7), and a model based on a linear relation between information resistance and channel length. (d) The model fit of t_{local} on the human data of Toro et al. (2008) (also indicated as black dots in panel a). (e) Model fit of t_{local} on the mean values per species, using the outer surface values (eq. (8)).

9

The volume to a given cortical surface area is thus determined by the flow information processed by the neurons, and the transfer of this information on the local- and global-scale. This can easiest be expressed in terms of *information resistance*: how the flow of information [bit/s] divides among local and global networks depends on how well each route transfers information. Information resistance is directly related to the inverse of the *channel capacity* which expresses the rate of information that can be reliably transmitted by a channel (Balasubramanian et al., 2001, Shannon, 1948). Assuming that the resistance to information increases with the squared length of the connection (Section 2.5), it becomes clear that the relative contribution of the small-scale and large-scale connections depends directly on the size of the cerebrum (T_s: eq. (5)). Thus, expressing information resistance in terms of equivalent connection length, and subsequently substituting equation (5) we get:

$$\frac{1}{R_{I,total}} = \frac{1}{R_{I,global}} + \frac{1}{R_{I,local}}$$
$$\frac{1}{T_c^2} = \frac{1}{T_s^2} + \frac{1}{t_{local}^2}$$
$$= \left(\frac{\hat{\sigma}\hat{s}}{\sqrt[3]{V_c}}\right)^2 + \frac{1}{t_{local}^2} \tag{6}$$

where the constant t_{local} expresses the scale-invariant information resistance of the local networks [m]. By substitution of the equivalent thickness (T_c, equation (4)), the volume to cortical surface is obtained as:

$$A_c^2 = \left(\hat{\sigma}\hat{s}V_c^{2/3}\right)^2 + \left(\frac{1}{t_{local}}V_c\right)^2 \tag{7}$$

A remarkable property of this surface-volume relation is that it has just a single free parameter: t_{local}.

The fit of equation (7) on the log-transformed data (Section 2.1) resulted in $t_{local} = 3.6$ mm (N = 45 species, $R^2 = 0.995$). This result is presented in Figure 1a. Even though the model has just a single free parameter, it describes the data much better than a conventional linear regression on the log-transformed data (which yields: $\log(V_c) = 0.67 + 1.24\log(A_c)$, with an $R^2 = 0.98$).

10

Another manner to analyze the quality of the model is by the trend in the errors. These are illustrated in Figure 1b. Whereas the linear regression shows a clear trend in the form of an inverted U, the model does not show any trend in the errors.

For part of the dataset, the volume data are not present or not reliable (see Section 2.1), but the outer surface is given in these data. Given that the sphericity parameter is invariant (Section 2.2), we can rewrite the model (eqn. (7)) as a function of the outer surface S_c, by substitution of equation (3):

$$A_c^2 = S_c^2 + \left(\frac{\sigma s}{t_{local}}\right)^2 S_c^3 \tag{8}$$

(given that $(\hat{\sigma}\hat{s})^3 = (\sigma s)^{-2}$).

The fit of relation is as good as the first ($t_{local} = 3.4$ mm, $R^2 = 0.996$, N = 51 species). This is shown in Figure 1d.

The second fit involves 7 species of cetaceans (rather than just three species in the first fit). A fit of the 15 samples collected from these 7 species of cetaceans resulted in $t_{local} = 3.1$ mm, $R^2 = 0.92$, which was thus close to the ensemble-fitting result.

Since the wide availability of neuroimaging facilities, it has become possible to measure the gross anatomy of large samples of cerebrums with standardized and automized procedures. Such a data set of human subjects (Pausova et al., 2007) was analyzed by Toro et al. (2008). These data (314 subjects, 164 females and 150 males of 12-20 years old) are presented in Figure 1c. The fit resulted in $t_{local} = 4.8$ mm ($R^2 = 0.96$), which was slightly better than the linear regression on the log-transformed data (which resulted in a power slope of 1.0, intercept 0.54 and $R^2 = 0.94$). Thus, the model not only describes the inter- but also the intraspecific variation with high accuracy.

To test for phylogenetic trends, t-tests were performed on the fitting errors, using $t_{local} = 3.6$ mm (i.e., the result of the fit on all species). Orders for which more than 5 species (and species with more than 10 samples) were available were included, i.e., Carnivora (N=8), Primata without humans (N = 8), Soricomorpha (N=7), and humans (N=315). The test only gave a significant result for the human data (t(314)= 133, P< 0.001, Bonferroni-corrected). To test for systematic differences between the five literature sources of the database (Brodmann, Haug, Hofman, Mayhew et al., Schlenska) t-tests were performed on the non-averaged data (again with $t_{local} = 3.6$ mm). This test gave no significant results (Bonferroni-corrected).
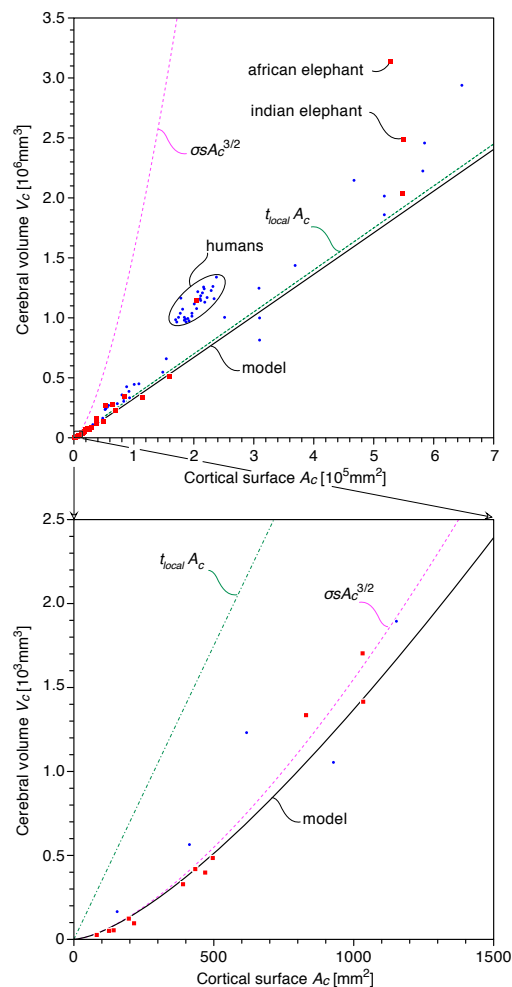
11

Figure 2: The data of Figure 1a on a linear scale. Top panel shows the entire data range, whereas the bottom panel shows the lissencephalic range on a much enlarged scale.

3.1.1. Sensitivity

As shown in Figure 1b, the fit of the scaling law (eq. (7)) has no systematic errors. Two main causes can underly the variable errors: measurement errors and true variations of the model parameters. The measurement errors are difficult to estimate due to the generally low number of samples measured per species. Judging from the human data (Fig. 1d) and taking the data of (Toro et al., 2008) as a reference, it seems that systematic errors may be considerable, depending on the author.

The sphericity parameter, s had a standard deviation (SD) of 0.09. The

12

variations in the sphericity parameter mainly affect the model predictions for the small range of the cerebrums (Fig. 1a). Variations in t_{local}, on the other hand affect the model predictions in the range of large cerebrums (Fig. 1a).

The model was based on a quadratic relation between brain size and information resistance for long-range axonal connections (see section 2.5), but there are no reliable data to tell whether this really is the correct relation. The effect of assuming a linear relation is show in Figure 1c. The linear assumption gives a slightly less good description of the data, but the effect is small. A fit of the linear model resulted in $t_{local} = 6.1$ mm ($R^2 = 0.995$).

Even if double logarithmic plots have huge merits, they can be misleading. Therefore, the data are plotted on linear scales as well (Fig. 2). The figures confirm that the fits are good over most of the range. On the linear scale it is much more clear that the human data are systematically off the Cortical surface–cerebral volume model (eqn. (8)). Also, the indian and african elephants may fit better with a larger t_{local}, but more measurements would be needed to confirm this.

3.2. Relation of grey and white matter

In 2000, Zhang and Sejnowski discovered that there is a remarkably regular relationship between the volumes of grey and white matter in the mammalian cerebrum, as reproduced in Figure 3. They found that the relation between the grey (G) and white (W) volume almost perfectly describes a straight line on a double-logarithmic scale, over a range of more than five magnitudes of cerebral volume. A straight line on a double-logarithmic scale can always be described by a simple power of the form

$$y = ax^b$$

with constants a and b.

Interestingly, the same was true for the relations between the total volume (V_c) with the grey volume, but also with the white volume. As Zhang and Sejnowski acknowledged, it is not possible that all three relations really are simple power relations, because $W + G = V_c$. Zhang and Sejnowski tried to derive a power law for the empirical relation, but their equations were circular (see Appendix). Moreover, the empirical power of 1.23 was significantly lower than the expected natural scaling power of 4/3. Thus, the relation between white and grey matter volume is not explained by a simple power law. Moreover, a large family of relations are approximately
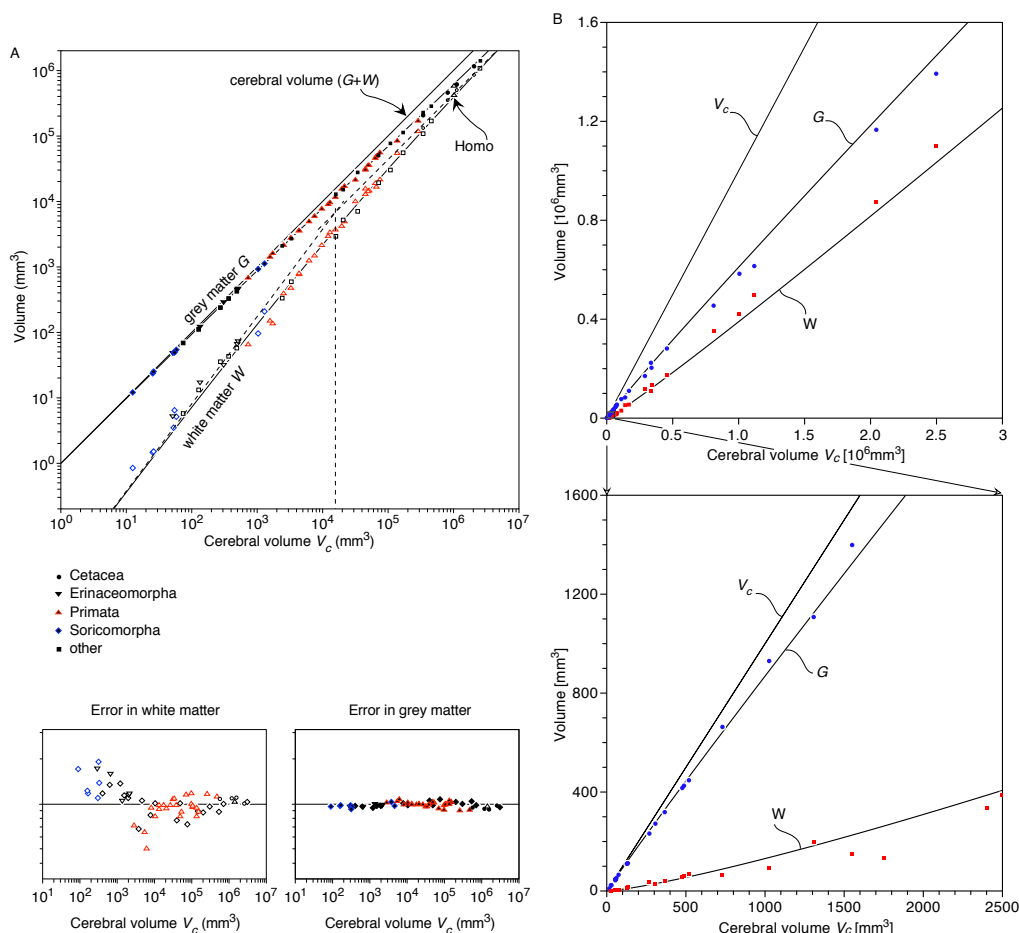
13

Figure 3: (*a*) White and grey mater volumes from Zhang and Sejnowski (2000). The white matter volumes (open symbols) and the model fit (eqn. (9)). (*b*) The same data on a linear scale. Top panel shows the entire data range, whereas the bottom panel shows the lissencephalic range on a much enlarged scale.

linear on double logarithmic scale.

Again the information transfer can be used as a starting point for deriving a general scaling theory. Above has become clear that two scales play a role in the scaling: the local, intra-gyral and the global, extra-gyral scale. On the local, intra-gyral scale the information transfer is scale-invariant, so the white matter volume, W, should scale proportionally to the total volume, V_c.

On the global scale the W should also scale with V_c. However, in addition

14

the distances scale with cerebral size, so that W should also scale with the radius. Since the scaling due to information transfer is governed by the *information resistance* according to equation (6), the scaling of the white matter volume should also be governed by the inverse sum of the local and global scales. Thus we obtain the following equation:

$$\frac{1}{W} = \frac{1}{arV_c} + \frac{1}{bV_c} \qquad (9)$$

with r as given by equation (1). Fitting yields $a = 0.0025$ mm^{-1} and $b = 0.50$ (R$^2_W = 0.995$ and R$^2_G = 0.9998$ for the white and grey matter volume respectively). The fit is shown in Figure 3. This fit is as good as the one by Zhang and Sejnowski (who obtained an R$^2 = 0.996$ for their model).

Both, the volume-surface model (eq. (7)) and the grey-white matter volume model (eq. (9)) are limit relations, indicated by dashed lines in Figures 1a and 3a. A measure for the transition between the limit relations is the point of intersection between them. As expected if the models are both driven by the same mechanism, the points of intersection are almost the same (at $1.9 \cdot 10^4$ and $1.6 \cdot 10^4$ mm^3 respectively).

4. Discussion

Compared to other scaling relations known for animals over a wide size range, the here treated scaling relations of the mammalian brain are remarkably regular, and seemingly linear, on a double-logarithmic scale. This has been known for a long time but the reasons have remained elusive, because the power relations that are obtained by such linear regressions on log-transformed data are incompatible with isometric scaling laws. Here, for the first time, a theory is developed that fully explains these narrow scaling relations. Moreover it predicts these relations with parameters that are physiologically meaningful, determined by the central functions of the cerebrum –processing and transmitting information– and that can in part be derived from independent measurements.

The obvious advantage of the current theory driven approach over mere regression fitting, is that it opens the possibility to analyze exceptions to the scaling. The first thing that can be noted here is that such exceptions are remarkably scarce. On the basis of the available data, there are no phylogenetic trends, either for the grey and white matter volumes nor for the volume-surface relations. This is remarkable, given recent claims for such effects, e.g., between Glires and Primates (Herculano-Houzel, 2011). In this

15

sense the discussion remains similar to that of a century ago when Brodmann proposed that the six-layered structure of the mammalian neocortex is a universal feature (Brodmann, 1909).

There are, however also clear indications of a divergence in the volume-surface relations for the largest cerebrums. The human data set is the most reliable, in having a t_{local} of almost a mm more than the general fit. The elephants appear to have a similar t_{local} to humans. Interestingly, the cetaceans align very well with the data of smaller mammals, although it is usually thought that cetacean cerebrums are deviant, i.e., have a unusually large cerebral surface. This latter conviction is an artifact of fitting the logarithmic data with a linear regression.

The scale-invariant sphericity s is assumed to be constant, though, as indicated by the standard deviation, it does shows considerable variation. The effect of such variations mainly affects the predictions for small cerebrums, as shown by the sensitivity analysis (cf. Fig. 1a). Overall, the volume of lissencephalic cerebrums is slightly underestimated by the model, which is due to the very small (statistically not significant) size dependency indicating that Ss increases slightly with brain size (cf. Section 2.2).

The only free parameter of the volume-surface model, $t_{local} \approx 3.6$ mm, expresses the information resistance on the local scale, but is also a length measure for the local networks. The measure expresses equivalent length (cf. Section 2.3). This is in the order of the size of the gyral lobes of convoluted brains. The largest cerebrums closely approximate the limit relation t_{local}. That means, such large cerebrums effectively function more like a communicating cluster of local processing centers than as a global processor. The model is only a first order approximation in so far, that axonal lengths show a wide distribution within a cerebrum (Jones and Wise, 1977, Kaiser et al., 2009). Further the regional properties are likely to vary (Shipp, 2005). The deviation of humans and elephants indicates a change in the local network properties, so that the local processing is increased. As a result, the current result indicates that the large number of cortical areas in the human brain is a true and remarkable property (Felleman and Van Essen, 1991). It remains to be tested whether this is a property of specific regions of the human cerebrum, such as the prefrontal cortex.

It was estimated that the information resistance increases quadratically with connective distance, but this may not be entirely accurate. As a comparison, the scaling law was also estimated with a linear scaling of information resistance with connective distance (Fig. 1c). As shown by the figure,

16

the predicted surface-volume relation is rather insensitive to such a seemingly grave difference. The theoretical, quadratic scaling of information resistance with distance describes the available data better than the linear example, which is still better than a linear regression.

Lastly, the principle of information processing can even be applied to the grey and white matter volumes. This last law is the best that is currently available, but it is less strong than the surface-volume model, because its two parameters (a and b) so-far do not have direct physiological meaning. The model confirms that local processing networks are very prominent in large cerebrums. The model seems to systematically underestimate the white matter volume in the very small cerebrums (cf. Fig. 3a). This may in part be due to systematic measurement errors in these very small cerebrums. In part it may be due to the deviation in sphericity parameter s indicating that small cerebrums and therefore have relatively longer connections than assumed by the model.

Several earlier studies have developed a theoretical approach to explain the empirical scaling relations (Changizi, 2001, Mota and Herculano-Houzel, 2012, Prothero and Sundsten, 1984, Toro and Burnod, 2005, Zhang and Sejnowski, 2000, see also Appendix). Different to these studies, the present approach started from the notion that the computational power and information transfer are the most likely candidates as basic parameters in shaping the cortex. Since computational power depends directly on the number of neurons, which is again constant per cortical surface, it follows that cortical surface is to be taken as the independent parameter.

Still, the relations are mainly based on relatively scarce data from heterogenous sources. It would therefore be highly valuable to improve the measurements, and with modern imaging techniques it is feasible to measure much larger samples within species in a much more automated manner (Manger et al., 2012, Pillay and Manger, 2007, Toro et al., 2008). Such databases with larger samples should also enable to test more specific patterns such as developmental and gender patterns (Paus and Toro, 2009) as well as local differences between regions of the cerebrum.

The result suggests that the structure of our cerebrum has landed us in a local optimum, because the wrinkled surface structure makes that many distances are longer than necessary (think of the two flanking sides of a sulcus). In a nuclear structure such as the birds' cerebrum this problem does probably not exist (The Avian Brain Nomenclature Consortium, 2005).

In conclusion, the general basis for scaling of the cerebrum are the corti-

17

cal surface area, A_c, and the size of the local networks, as expressed by the t_{local} parameter. These local networks may well differ between regions of the cortex, because the expression of layers differs substantially, for example between granular and non-granular cortex. Moreover, the scaling of the cortex is only sensitive to variations in t_{local} for large cerebrums (Section 3.1.1).

On the basis of the dataset used, there is no evidence for phylogenetic trends in contrast to earlier claims for systematic phylogenetic differences. With respect to the t_{local} with one exception: Humans and elephants appear to have a significantly larger t_{local} parameter than all other mammals (except, possibly elephants), indicating that information flow within regional networks may play a larger role than in other mammals.

Third, the amount of white matter in the cerebrum scales according the current information approach.

Fourth, the relation between cortical surface and cortical volume could be fitted using just a single parameter, t_{local} = 3.6 mm. The underlying theory is based on efficient information processing and transmission. The t_{local} of 3.6 mm is consistent with an average gyral width of about 10 mm, and a maximal lissencephalic radius of 10 mm ($\approx 2 \cdot 10^2$ mm^3 volume).

The theoretical foundation of the scaling laws suggests that the functions of the cerebrum, i.e., processing and transmission of information also are central in defining its macroscopic shape. This in turn suggests that the local structures in the cerebrum are also strongly related to their function, which is consistent with the fact that primary sulci tent to be located at the primary sensory and motor regions, and that the borders of functional regions in the cerebrum usually follow the gyral folds.

Thus, the volume to surface relation does seem to be optimal. The optimization criterion is derived directly from the central function of the cerebrum: the processing and transfer of information. This result also has a direct and important consequence for the way the brain works, because it means that small cerebrums work fundamentally differently than large cerebrums, in the regionalization of information processing.

Acknowledgements

18

References

Baillarger, J. (1845). De l'étendue de la surface du cerveau. *Gazette des Hôpiteaux*, 18:179.

Balasubramanian, V., Kimber, D., and Berry, M. J. (2001). Metabolically efficient information processing. *Neural Comput.*, 13:799–815.

Barton, R. A. and Venditti, C. (2013). Human frontal lobes are not relatively large. *Proc. Natl. Acad. Sci. USA*, 110(22):9001–9006.

Bok, S. T. (1929). Der Einfluß der in den Furchen und Windungen auftretenden Krümmungen der Großhirnrinde auf die Rindenarchitektur. *Zeitschr. gesamte Neur. Psych.*, 121(1):682–750.

Brodmann, K. (1909). *Vergleichende Lokalisationslehre der Großhirnrinde; in ihren Prinzipien dargestellt auf Grund des Zellenbaues.* Joh. Ambrosius Barth, Leipzig.

Brodmann, K. (1913). Neue Forschungsergebnisse der Großhirnrindenanatomie. *Verhandlungen der Gesellschaft Deutscher Naturforscher und Ärzte*, 85(1):200–240.

Carlo, C. N. and Stevens, C. F. (2013). Structural uniformity of neocortex, revisited. *Proc. Natl. Acad. Sci. USA*, 110(4):1488–1493.

Changizi, M. A. (2001). Principles underlying mammalian neocortical scaling. *Biol. Cybern.*, 84(3):207–215.

de Lussanet, M. H. E. (2013). The human and mammalian cerebrum scale by computational power and information resistance. *arXiv*, q-bio.NC:1308.1252.

Eguíluz, V. M., Chialvo, D. R., Cecchi, G. A., Baliki, M., and Apkarian, A. V. (2005). Scale-free brain functional networks. *Phys. Rev. Lett.*, 94:018102.

Elias, H. and Schwartz, D. (1971). Cerebrocortical surface areas, volumes, lengths of gyri and their interdependence in mammals, including man. *Z. Säugetierk.*, 36:147–163.

Felleman, D. J. and Van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cereb. Cortex*, 1:1–47.

Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *J. Exp. Psych. Hum. Percept. Perform.*, 47:381–391.

Fitts, P. M. and Petterson, J. R. (1964). Information capacity of discrete motor responses. *J. Exp. Psych. Hum. Percept. Perform.*, 47:381–391.

Gabi, M., Collins, C. E., Wong, P., Torres, L. B., Kaas, J. H., and Herculano-Houzel, S. (2010). Cellular scaling rules for the brains of an extended number of primate species. *Brain Behav. Evol.*, 76:32–44.

Gross, C. G. (1999). *Brain, vision, memory: Tales in the history of neuroscience.* Bradford books.

Harman, P. J. (1947). On the significance of fissuration of the isocortex. *J. Comp. Neurol.*, 87(2):161–168.

Harrison, K. H., Hof, P. R., and Wang, S. S. H. (2002). Scaling laws in the mammalian neocortex: Does form provide clues to function? *J. Neurocytol.*, 31(3):289–298.

Haug, H. (1970). Der makroskopische Aufbau des Großhirns: Qualitative und quantitative Untersuchungen an den Gehirnen des Menschen, der Delphinoideae und des Elefanten. In *Ergebnisse der Anatomie und Entwicklungsgeschichte*, volume 43, Berlin. Springer.

19

Henneberg-Neubabelsberg, R. (1910). Messungen der Oberfläche der Großhirnrinde. *J. f. Psychol. u. Neurol.*, 17(3-4):144–158.

Herculano-Houzel, S. (2011). Not all brains are made the same: New views on brain scaling in evolution. *Brain Behav. Evol.*, 78:22–36.

Herculano-Houzel, S., Mota, B., Wong, P., and Kaas, J. H. (2010). Connectivity-driven white matter scaling and folding in primate cerebral cortex. *Proc. Natl. Acad. Sci. USA*, 107(44):19008–19013.

Hofman, M. A. (1985). Size and shape of the cerebral cortex in mammals. I. The cortical surface. *Brain Behav. Evol.*, 27:28–40.

Hofman, M. A. (1988). Size and shape of the cerebral cortex in mammals. II. The cortical volume. *Brain Behav. Evol.*, 32:17–26.

Hofman, M. A. (1989). On the evolution and geometry of the brain in mammals. *Prog. Neurobiol.*, 32(2):137–158.

Jerison, H. J. (1982). Allometry, brain size, cortical surface and convolutedness. In Armstrong, E. and Falk, D., editors, *Primate brain evolution*, pages 77–84, New York. Plenum Press.

Jones, E. G. and Wise, S. P. (1977). Size, laminar and columnar distribution of efferent cells in the sensory-motor cortex of monkeys. *J. Comp. Neurol.*, 175(2):391–438.

Kaiser, M., Hilgetag, C. C., and van Ooyen, A. (2009). A simple rule for axon outgrowth and synaptic competition generates realistic connection lengths and filling fractions. *Cereb. Cortex*, 19(12):3001–3010.

Koch, K., McLean, J., Segev, R., Freed, M. A., Berry II, M. J., Balasubramanian, V., and Sterling, P. (2006). How much the eye tells the brain. *Curr. Biol.*, 16(14):1428–1434.

Le Gros Clark, W. (1945). Deformation patterns on the cerebral cortex. In *Essays on growth and form*, pages 1–22, Oxford. Oxford University Press.

Manger, P. R., Prowse, M., Haagensen, M., and Hemingway, J. (2012). Quantitative analysis of neocortical gyrencephaly in african elephants (loxodonta africana) and six species of cetaceans: comparison with other mammals. *J. Comp. Neurol.*, 520(11):2430–2439.

Marino, L. (1998). A comparison of encephalization between odontocete cetaceans and anthropoid primates. *Brain Behav. Evol.*, 51(4):230–238.

Markov, N. T., Ercsey-Ravasz, M., Lamy, C., Ribeiro Gomes, A. R., Magrou, L., Misery, P., Giroud, P., Barone, P., Dehay, C., Toroczkai, Z., Knoblauch, K., Van Essen, D. C., and Kennedy, H. (2013). The role of long-range connections on the specificity of the macaque interareal cortical network. *Proc. Natl. Acad. Sci. USA*, 110(13):5187–5192.

Mayhew, T. M., Mwamengele, G. L., and Dantzer, V. (1996). Stereological and allometric studies on mammalian cerebral cortex with implications for medical brain imaging. *J. Anat.*, 189(1):177–184.

Mota, B. and Herculano-Houzel, S. (2012). How the cortex gets its folds: An inside-out, connectivity-driven model for the scaling of Mammalian cortical folding. *Front. Neuroanat.*, 6:3.

Mountcastle, V. B. (1997). The columnar organization of the neocortex. *Brain*, 120(4):701–722.

Paus, T. and Toro, R. (2009). Could sex differences in white matter be explained by g ratio? *Frontiers in Neuroanatomy*, 3:14.

20

Pausova, Z., Paus, T., Abrahamowicz, M., Almerigi, J., Arbour, N., Bernard, M., Gaudet, D., Hanzalek, P., Hamet, P., Evans, A. C., Kramer, M., Laberge, L., Leal, S. M., Leonard, G., Lerner, J., Lerner, R. M., Mathieu, J., Perron, M., Pike, B., Pitiot, A., Richer, L., Séguin, J. R., Syme, C., Toro, R., Tremblay, R. E., Veillette, S., and Watkins, K. (2007). Genes, maternal smoking, and the offspring brain and body during adolescence: Design of the Saguenay Youth Study. *Hum. Brain Mapp.*, 28(6):502–518.

Perge, J. A., Koch, K., Miller, R., Sterling, P., and Balasubramanian, V. (2009). How the optic nerve allocates space, energy capacity, and information. *J. Neurosci.*, 29(24):7917–7928.

Perge, J. A., Niven, J. E., Mugnaini, E., Balasubramanian, V., and Sterling, P. (2012). Why do axons differ in caliber? *J. Neurosci.*, 32(2):626–638.

Pillay, P. and Manger, P. R. (2007). Order-specific quantitative patterns of cortical gyrification. *Eur. J. Neurosci.*, 25(9):2705–2712.

Prothero, J. W. and Sundsten, J. W. (1984). Folding of the cerebral cortex in mammals: a scaling model. *Brain Behav. Evol.*, 24(2-3):152–167.

Rockel, A. J., Hiorns, R. W., and Powell, T. P. S. (1980). The basic uniformity in structure of the neocortex. *Brain*, 103(2):221–244.

Roerig, B. and Chen, B. (2002). Relationships of local inhibitory and excitatory circuits to orientation preference maps in ferret visual cortex. *Cerebral Cortex*, 12(2):187–198.

Schlenska, G. (1969). Messungen der Oberfläche und der Volumenanteile des Gehirnes menschlicher Erwachsener mit neuen Methoden. *Z. Anat. Entwickl. Gesch.*, 128(1):47–59.

Schlenska, G. (1974). Volumen- und Oberflächenmessungen an Gehirnen verschiedener Säugetiere im Vergleich zu einem errechneten Modell. *J. Hirnforsch.*, 15(15):401–408.

Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Tech. J.*, 27:379–423, 623–656.

Shipp, S. (2005). The importance of being agranular: A comparative account of visual and motor cortex. *Phil. Trans. R. Soc. B*, 360:797–814.

Stoop, R. L., Saase, V., Wagner, C., Stoop, B., and Stoop, R. (2013). Beyond scale-free small-world networks: Cortical columns for quick brains. *Phys. Rev. Lett.*, 110(10):108105.

The Avian Brain Nomenclature Consortium, . (2005). Avian brains and a new understanding of vertebrate brain evolution. *Nat. Rev. Neurosci.*, 6:151–159.

Toro, R. and Burnod, Y. (2005). A morphogenetic model for the development of cortical convolutions. *Cereb. Cortex*, 15(12):1900–1913.

Toro, R., Perron, M., Pike, B., Richer, L., Veillette, S., Pausova, Z., and Paus, T. (2008). Brain size and folding of the human cerebral cortex. *Cereb. Cortex*, 18(10):2352–2357.

Van Essen, D. C. (1997). A tension-based theory of morphogenesis and compact wiring in the central nervous system. *Nature*, 385(6614):313–318.

Watts, D. J. and Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature*, 393(6684):440–442.

Zhang, K. and Sejnowski, T. J. (2000). A universal scaling law between gray matter and white matter of cerebral cortex. *Proc. Natl. Acad. Sci. USA*, 97(10):5621–5626.

Zilles, K., Palomero-Gallagher, N., and Amunts, K. (2013). Development of cortical

21

folding during evolution and ontogeny. *Trends Neurosci.*, 36(5):275–284.

22

Appendix

Earlier theories

The theory of brain scaling has a long history. It is outside the scope to give a complete listing and I am not even sure to know all theories. A few are worth treating in in more detail though because they are cited and mentioned frequently.

Surface-volume relations

The theories for the convoluted surface area of the cerebrum seem to have to longest history Baillarger (1845), Gross (1999). For a review, see (Hofman, 1989). For a long time it has been thought that the folding is a result of a large cortical area being fitted inside a comparably small space inside the skull. This model apparently dates back to Le Gros Clark (1945), and has been modeled numerically in a slightly modified form (Toro & Burnod, 2005, Toro, 2012). However, for this model to work, one has to assume that the grey matter is incompressible, rubber-like in the dimensions parallel to the surface, and that the white matter is viscous. These are highly unlikely properties making the model as a whole questionable. Moreover, the model does not make any predictions as to the scaling relations of the cortical surface area.

A popular theory was developed by Prothero & Sundsten (1984) in a number of works. According to this model, the gyral width and height are governed by the white matter of the long-range axonal connections. The strongest prediction of the model is that there exists an upper limit to the size of the cerebral cortex. The strongest weakness of the model is that it approximates the cerebrum with a cubical core of white matter which has a surface of equally-shaped gyral ridges, separated by sulci of equal depth. In such a design, the gyri of a large cerebrum cannot contain any white matter. However, in a true cerebrum, the depth of neighboring sulci shows strong differences, so in practice this is unlikely to present a real design problem to the cerebrum.

According to one intuition the space required by the site matter is a shaping factor for the gyri (Prothero & Sundsten, 1984). According to a second notion, the anisotropic material properties of the grey and white matter are essential shaping factors. For example, according to an early proposal, the cortex behaves like a thick, lubricated rubber sheet that is packed inside a skull that is too small for its surface (Le Gros Clark, 1945).

23

It will fold (as Le Gros Clark demonstrated experimentally) because the sheet is hardy compressible in the tangential directions. This proposal has also been modeled numerically in a slightly modified form (Toro & Burnod, 2005). In this model, the cortical sheet did not grow against a skull, but against a compliant centripetal force.

A very interesting mechanism for the development of folding has been made by Van Essen (1997). According to Van Essen, the mechanical stiffness of axons make that the cerebral white matter is a highly anisotropic material (Peter & Mofrad, 2012). It seems plausible that the mechanism of Van Essen accounts for the development of the relationship between cerebral volume and cortical surface.

A general law for white and grey volume?

Zhang & Sejnowski (2000) also derived what they called "a general law" to predict the relation between the grey and white matter volume. Unfortunately, there was a circularity in their equations as I demonstrate here. To avoid confusion, I will use symbols that are consistent with the present work.

Zhang and Sejnowski started with two assumptions. (1) That there is a direct linear relation between the cortical surface area (A_c) and the physiological cross-sectional area (PCSA) of the axons. (2) That the global geometry minimizes the average length of the axonal fibers (L_{ax}). This latter assumption is somehow incomplete, because the shortest possible average length would obviously be zero.

The grey matter volume V_g is the surface A_c times the equivalent cortical thickness T_g (cf. eq. (4)):

$$V_g = T_g A_c \tag{s1}$$

Given assumption no. 1, the white matter volume V_w is:

$$V_w = c_1 A_c L_{ax} \tag{s2}$$

the constant $0.5 < c_1 < 1$ was introduced by Zhang and Sejnowski to acknowledge that axons have a start and an end, so that cortico-cortical connections would be counted double.

The crucial step is to obtain the relation between G and W. Zhang and Sejnowski postulated that the total grey matter volume V_g depends on the average axon length L_{ax}. To match the dimensions they assumed that $V_g = c_2 L^3$. They tested this postulate against a family of alternative postulates in which L^3 is replaced by $L^{3-n} X^n$ with $n < 2$ and X another

24

length measure such as the cortical thickness T or average gyrus length, etc. They then showed that L_{ax} is smallest for the original postulate, assuming that all alternative parameters $X < L$. Combining this postulate with equations (s1) and (s2) yields the "universal scaling law":

$$V_w = \frac{c_3}{T_g} G^{4/3} \tag{s3}$$

This formulation is misleading because according to equation (s1) V_g equals $T_g A_c$ so "the law" really states:

$$V_w = c_3 T_g^{1/3} A_c^{4/3} \tag{s4}$$

so, in fact, the equation does not directly link W and G at all. If we replace V_w again with equation (s2) we obtain again the postulation, so we have not learned anything. The circularity in the derivation of their law occurred in their equation [13], where they substituted the axonal length from their equation [2] (eq. (s2)), which already been substituted in their equation [5] (i.e., the "universal scaling law", eq. (s3)).
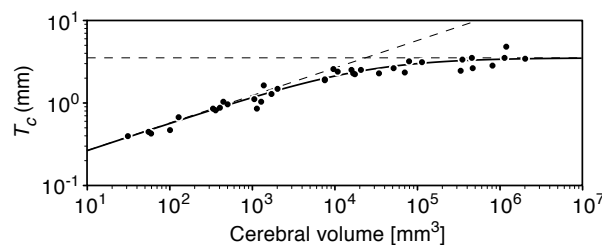
Other relations

Figure S1: Equivalent cerebral thickness ($T_c = V_c/A_c$). Computed from the data of (Hofman, 1989). Dashed lines have a power slope of 1/3 and 1; the curve is equation (6).

A popular measure of the property of the cortex is the equivalent thickness (see Section 2.3). The model also predicts the equivalent cerebral thickness T_c, in equation (6). This relation is presented in Figure S1. It shows again the very good fit of the model to the data.

The theoretical approach by Changizi (2001) is interesting for its attempt to explain a whole set of scaling relations of the mammalian cerebrum at once, and thus aiming at a general theory, instead of focussing on just one relation at the time. A weakness of his approach is that it is

25

based on the power fitted from double-logarithmic relations of a long, het-
erogenous list of studies, without any reference to the quality of these fits.
For example, he lists a power of 0.08-0.197 for the relation between cortical
thickness and grey matter volume although this relation deviates strongly
from a power relation (cf. Fig. S1). His model consists of two parts. The
first part predicts that the number of synapses per volume of dendritic tree
is a scale-invariant constant. However he fails to mention the implicit as-
sumption that the branching rate must be independent of dendritic length.
The second part builds upon this, and a number of far-reaching assump-
tions. For example, it is assumed that each region of the brain connects
to a limited, fixed fraction of other regions, independent of the number of
regions in the cerebrum. Not only is the fraction of connected regions as-
sumed constant, also the fraction of neurons connected to in each connected
area is assumed constant.

A recent modeling approach has been presented by Mota & Herculano-
Houzel (2012), to develop a computational framework for the data on cell
counts in the cerebrum and other brain regions in a large range of mammals
by Susana Herculano-Houzel et al. (Gabi *et al.*, 2010, Herculano-Houzel,
2009, Herculano-Houzel *et al.*, 2010, Herculano-Houzel, 2011, 2012). This
model treats the cerebral volume as the sum of grey and white matter. It has
neurons, axons and glia as major parameters, and is explorative in nature,
meaning that it is an attempt to list and identify probable parameters that
underly the empirical scaling laws. The main result is that it is a complex
matter.

References to Appendix

Hofman, M. A. 1989 On the evolution and geometry of the brain in mammals. *Prog. Neurobiol.*, **32**(2), 137–158. (doi:10.1016/0301-0082(89)90013-0)

Baillarger, J. 1845 De l'étendue de la surface du cerveau. *Gazette des Hópiteaux*, **18**, 179.

Gross, C. G. 1999 *Brain, vision, memory: Tales in the history of neuroscience.* Bradford books.

Le Gros Clark, W. 1945 Deformation patterns on the cerebral cortex. In *Essays on growth and form*, pp. 1–22. Oxford: Oxford University Press.

Toro, R. & Burnod, Y. 2005 A morphogenetic model for the development of cortical convolutions. *Cereb. Cortex*, **15**(12), 1900–1913. (doi:10.1093/cercor/bhi068)

Prothero, J. W. & Sundsten, J. W. 1984 Folding of the cerebral cortex in mammals: a scaling model. *Brain Behav. Evol.*, **24**(2-3), 152–167.

Van Essen, D. C. 1997 A tension-based theory of morphogenesis and compact wiring in the central nervous system. *Nature*, **385**(6614), 313–318. (doi:10.1038/385313a0)

26

Peter, S. J. & Mofrad, M. R. K. 2012 Computational modeling of axonal microtubule bundles under tension. *Biophysical J.*, **102**(4), 749–757. (doi:10.1016/j.bpj.2011.11.4024)

Zhang, K. & Sejnowski, T. J. 2000 A universal scaling law between gray matter and white matter of cerebral cortex. *Proc. Natl. Acad. Sci. USA*, **97**(10), 5621–5626. (doi:10.1073/pnas.090504197)

Changizi, M. A. 2001 Principles underlying mammalian neocortical scaling. *Biol. Cybern.*, **84**(3), 207–215. (doi:10.1007/s004220000205)

Mota, B. & Herculano-Houzel, S. 2012 How the cortex gets its folds: An inside-out, connectivity-driven model for the scaling of Mammalian cortical folding. *Front. Neuroanat.*, **6**, 3. (doi:10.3389/fnana.2012.00003)

Gabi, M., Collins, C. E., Wong, P., Torres, L. B., Kaas, J. H. & Herculano-Houzel, S. 2010 Cellular scaling rules for the brains of an extended number of primate species. *Brain Behav. Evol.*, **76**, 32–44. (doi:10.1159/000319872)

Herculano-Houzel, S. 2009 The human brain in numbers: A linearly scaled-up primate brain. *Front. Hum. Neurosci.*, **3**, 31. (doi:10.3389/neuro.09.031.2009)

Herculano-Houzel, S., Mota, B., Wong, P. & Kaas, J. H. 2010 Connectivity-driven white matter scaling and folding in primate cerebral cortex. *Proc. Natl. Acad. Sci. USA*, **107**(44), 19 008–19 013. (doi:10.1073/pnas.1012590107)

Herculano-Houzel, S. 2011 Not all brains are made the same: New views on brain scaling in evolution. *Brain Behav. Evol.*, **78**, 22–36. (doi:10.1159/000327318)

Herculano-Houzel, S. 2012 The remarkable, yet not extraordinary, human brain as a scaled-up primate brain and its associated cost. *Proc. Natl. Acad. Sci. USA*, **109**(S1), 10 661–10 668. (doi:10.1073/pnas.1201895109)

Toro, R. 2012 On the possible shapes of the brain. *Evol. Biol.*, **39**(4), 600–612. (doi:10.1007/s11692-012-9201-8)

27

Supplement

775 **Dataset**

The data used for validating the scaling laws are attached as an excel table. At first the included measurements are printed, followed by the excluded ones including the reason for exclusion, and finally the species' mean values used for fitting. Cells with equations are marked in color.

28