PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

# On variability in the density of corticocortical and thalamocortical connections

Jack W. Scannell, Simon Grant, Bertram R. Payne and Roland Baddeley

*Phil. Trans. R. Soc. Lond. B* 2000 **355**, 21-35
doi: 10.1098/rstb.2000.0547

THE ROYAL SOCIETY



# On variability in the density of corticocortical and thalamocortical connections

**Jack W. Scannell[1], Simon Grant[2], Bertram R. Payne[3] and Roland Baddeley[4]**

*[1]Neural Systems Group, Department of Psychology, University of Newcastle upon Tyne, Ridley Building, Newcastle upon Tyne NE1 7RU, UK (j.w.scannell@ncl.ac.uk)*

*[2]Department of Sensorimotor Sytems, Division of Neuroscience, Imperial College School of Medicine, Charing Cross Campus, Fulham Palace Road, London W6 8RF, UK*

*[3]Department of Anatomy and Neurobiology, Center for Advanced Biomedical Research, Boston University School of Medicine, 700 Albany Street, Boston, MA 02118, USA*

*[4]Department of Experimental Psychology, School of Biological Sciences, University of Sussex, Falmer, Brighton BN1 9RH, UK*

Variability is an important but neglected aspect of connectional neuroanatomy. The quantitative density of the 'same' corticocortical or thalamocortical connection may vary by over two orders of magnitude between different injections of the same tracer. At present, however, the frequency distribution of connection densities is unknown. Therefore, it is unclear what kind of sampling strategies or statistical methods are appropriate for quantitative studies of connectivity. Nor is it clear if the measured variability represents differences between subjects, or if it is simply a consequence of intra-individual differences resulting from experimental technique and the exact placement of tracers relative to local spatial and laminar variation in connectivity.

We used quantitative measurements of the density of a large number of corticocortical and thalamocortical connections from our own laboratories and from the literature. Variability in the density of given corticocortical and thalamocortical connections is high, with the standard deviation of density proportional to the mean. The frequency distribution is close to exponential. Therefore, analysis methods relying on the normal distribution are not appropriate. We provide an appendix that gives simple statistical guidance for samples drawn from exponentially distributed data.

For a given corticocortical or thalamocortical connection density, between-individual standard deviation is 0.85 to 1.25 times the within-individual standard deviation. Therefore, much of the variability reported in conventional neuroanatomical studies (with one tracer deposited per animal) is due to within-individual factors. We also find that strong, but not weak, corticocortical connections are substantially more variable than thalamocortical connections. We propose that the near exponential distribution of connection densities is a simple consequence of 'patchy' connectivity. We anticipate that connection data will be well described by the negative binomial, a class of distribution that applies to events occurring in clumped or patchy substrates. Local patchiness may be a feature of all corticocortical connections and could explain why strong corticocortical connections are more variable than strong thalamocortical connections. This idea is supported by the columnar patterns of many corticocortical but few thalamocortical connections in the literature.

**Keywords:** neuroinformatics; anterograde; retrograde; cats; cerebral cortex; thalamus

## 1. INTRODUCTION

Over the past 100 years, neuroanatomists have exerted a great deal of energy to trace the neural pathways that link different regions of the brain. This work has been motivated by the belief that an understanding of the connectional structure of the brain will lead to a better understanding of brain function (Meynert 1890). The last century has seen great advances in methods for tracing connections (e.g. Marchi & Algeri 1895; Nauta & Gygax 1954; Kristensson *et al.* 1971; Cowan *et al.* 1972; Gerfen & Sawchenko 1984) and in the application of these methods to the thalamus and cortex in several species (Le Gros Clark 1932, 1942; Rose & Woolsey 1948; Polyak 1927, 1933). These advances have resulted in an explosion in our knowledge of brain connectivity (e.g. Zeki & Shipp 1988; Felleman & Van Essen 1991; Young 1993; Scannell *et al.* 1995; Pandya & Yeterian 1985), but have contributed very little to our knowledge of the magnitude of, and variability in, individual brain connections. This is because quantification remains particularly laborious (but see Olson & Musil 1992; Musil & Olson 1988*a*,*b*, 1991; MacNeil *et al.* 1997; Hilgetag & Grant, this issue). Therefore, the vast majority of corticocortical and thalamocortical connection tracing studies still use a small number of individuals and report qualitative, rather than quantitative, measures of connection density.

Recently, MacNeil *et al.* (1997) published quantitative data on the strengths of cortical and thalamic projections

o the middle suprasylvian (MS) visual cortical area in he cat. This study used several different tracers, but was ery careful to minimize variability in the areal extent of he tracer deposit. MacNeil *et al.* (1997) included only ases where the tracer was confined to a particular reti- otopic region of MS cortex, where the tracer reached all ortical layers (but not the white matter), where the racer deposit was a reasonable size and where the white natter was not damaged by the injection. Their data how a low level of variability in the strength of indivi- ual thalamocortical connections, but a very high degree f variability in the strength of individual corticocortical onnections.

The work of MacNeil *et al.* (1997) inspired us to pool uantitative connection data from our laboratories with ublished studies to investigate variability more system- tically. This effort is important because variability has t least four serious implications. First, the degree and ature of variability have practical consequences for xperimental design in neuroanatomical studies and for he way that results are reported. Most single studies se small samples from which it is impossible to make a easonable estimate of the distribution from which the ample came. However, any form of statistical inference, ven one as simple as calculating the standard eviation, has to make assumptions about the likely istribution of the data. By pooling data from a number f studies we can obtain a reasonable picture of this istribution, that can then be 'assumed' by other esearchers. If our pooled data show that variability is igh, then sample size and random sampling error ecome important issues.

Second, variability can also have implications for the ay that results are reported by anatomists and nterpreted by other researchers. For example, given very ariable connection densities, no single tracer injection is kely to produce a very 'typical' pattern of labelling in he rest of the cortex and all individual results are likely o depart substantially from the average or most epresentative case. This presents challenges for those sing connection information for data analytic studies e.g. Young 1993; Scannell 1997; Stephan, Zilles & Kötter, his issue) or synthetic modelling studies (e.g. Kötter & ommer, this issue).

Third, individual differences in connection densities ay have great functional importance. For example, they ay have a causal role in shaping individual differences n behaviour. However, to our knowledge, no attempt has et been made to untangle the contributions of between- ndividual variability from within-individual variability. Variability in the results of connection tracing experi- ents in different animals will have several sources. ources include within-animal factors, which would be resent even if the same experiment could be repeated n the same animal (e.g. experimental error, within-area eterogeneity in projection patterns), and factors that eflect systematic differences between animals (e.g. inter- nimal differences in connectivity). In the study of MacNeil *et al.* (1997), a single tracer substance was eposited at a single location within each single animal. However, repeated measures within a subject are ecessary to estimate within-individual variability. deally, such measures should be made with tracers that possess virtually identical uptake, transport and visibility characteristics. While MacNeil *et al.* made efforts the minimize variability in the spatial extent of tracer deposits (e.g. by making large deposits to avoid labelling only certain subcompartments of MS), in the absence of repeated measures, their results cannot distinguish between-individual variability from within-individual variability. Therefore the magnitude of between-individual differences is unclear.

To help resolve this situation, for the rest of this paper we make a strict distinction between within-individual and inter-individual cases. 'Within-individual' refers to the results that would be obtained if repeated injections were made in the same cortical area of the same individual. Within-individual variability will be due to random experimental error and to local differences in connectivity or tracer uptake within a cortical area. 'Inter-individual' refers to the results that would be obtained if single injections were made in the same cortical area of different individuals. Inter-individual variability will be due to systematic differences between individuals plus within-individual variability. The difference between inter- and within-individual variability should let us estimate the proportion of variance that is due to within-animal factors and the proportion of variance that is due to between-animal factors.

Fourth, the distribution of connection densities may provide important insights into the organization of cortico-cortical and thalamocortical connections. We illustrate this point with a simplistic, and almost certainly incorrect, model. The model makes the following assumptions: (i) there are a very large number of neurons in area A; (ii)each neuron has a small and fixed probability of projecting to area B; (iii) neurons in area A project to area B independently of each other; (iv) the probability of a neuron projecting from A to B is equal across all of area A and area B; (v) we make identical tracer injections in different individuals.

The model that we have just outlined describes a Poisson process. If it were true, we would expect the distribution of connection densities for any particular connection to be given by equation (1). Here, $\mu$ is the mean number of labelled neurons in area A following a tracer injection in area B, and $p(r)$ is the probability of finding the number, $r$, of labelled neurons in area A.

$$p(r) = \frac{\mu^r e^{-\mu}}{r!}. \qquad (1)$$

For all but the weakest connections, the model predicts normally distributed connection densities, where the standard deviation is equal to the square root of $\mu$. Deviations from the model would show that other processes contribute to variability in connections. Different models of cortical organization predict different kinds of variability. For example, highly variable distributions are common in biology where local processes are Poisson, but where the mean of the process varies from site to site or from individual to individual (e.g. Solomon 1983; Shaw *et al.* 1998; Stear *et al.* 1998). So, given patchy connections between cortical areas, we would expect a highly variable distribution of connection densities (Montero 1981; Raczkowski & Rosenquist 1983; Symonds & Rosenquist 1984; De Yoe

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

t Van Essen 1985; Sherk 1986; Zeki & Shipp 1988, 1989; hipp & Grant 1991).

## 2. METHODS

### (a) *Quantitative connection data*

We used quantitative data from published retrograde tracing udies (Olson & Musil 1992; Musil & Olson 1988*a*,*b*, 1991) and om our own laboratories. The published studies used the fluores- ent retrograde tracers nuclear yellow (NY) and bisbenzimide Bb) to investigate the connections of medial prefrontal cortex PFCm), area 6m, and anterior and posterior cingulate areas CGa and CGp) of the cat. These studies are particularly useful ecause two tracers were frequently placed in the same cortical rea of the same individual. For the rest of this paper, we assume hat NY and Bb have very similar neuronal uptake and transport haracteristics, so that they sample the same sets of connections. his assumption lets us use these double-label studies to estimate ithin-individual variability in connection patterns.

The data from our laboratories were obtained from injec- ons of the retrograde tracers WGA-HRP, Fluorogold or hodamine-labelled latex microspheres (MacNeil *et al*. 1997; rant & Shipp 1991). The tracers were injected into either the iddle suprasylvian visual area (MacNeil *et al*. 1997; Grant & hipp 1991) or the posterolateral lateral suprasylvian visual rea (PLLS; Grant & Shipp 1991). Details of our methods are ublished elsewhere (MacNeil *et al*. 1997; Grant & Shipp 1991). o minimize variability due to spatial variation in the tracer eposits (Grant & Shipp 1991; MacNeil *et al*. 1997), we ccepted cases only if they met all the following criteria. First, racer deposits exposed a reasonable area of the cortex to the racer substance, with the aim of avoiding differential labelling n subsets of afferent neurons with small or patchy terminal rborizations (Sherk & Ombrellaro 1988; Payne *et al*. 1991; hipp & Grant 1991). Second, all cortical layers were exposed o tracer, so avoiding differential labelling in subsets of afferent eurons with different laminar terminations. Third, tracer had ot spread into the white matter. And fourth, the label had not pread into the adjacent sulcus or lateral suprasylvian areas. or the data from the laboratory of Dr Payne, tracer deposits ere also limited to a particular region of the visual field epresentation.

Retrogradely labelled neurons in cortical areas or thalamic uclei distant to the injection sites were counted. The propor- on, $S_i$, of labelled thalamic or cortical neurons in any given rea was then calculated by dividing the number of neurons in he area or nucleus, $R_i$, by the total number of counted neurons, $\mathcal{T}$, in the cortex or thalamus, respectively (equation (2)).

$$_i = \frac{R_i}{\mathcal{T}} \tag{2}$$

his strategy eliminates any potential differences in the efficacy f labelling of cortical and thalamic neurons that might exist. ossible contributing variables include differences in numbers, zes and concentrations of terminals along cortical and halamic axon arbours, and differences in neuronal transport apacities. Moreover, expression of labelling densities in the orm of proportions removes the variability in connection ensities that result from absolute differences in tracer uptake etween injections and emphasizes differences in the pattern of abelling. When we refer to connection strength or density lsewhere in this paper, it is these normalized values to which we efer. However, where possible we have repeated the analyses with the raw, unnormalized, cell counts from our own laboratories. The raw data yield very similar results.

Scaling preserves relative mean connection density across comparable injections, but causes systematic underestimation of the variability of connections. This is because the total number of labelled cells, $\mathcal{T}$, covaries with the number of labelled cells in each area or nucleus, $R_i$ (equation (2)).

$$\mathcal{T} = R_1 + R_2 + R_3 + \ldots . \tag{3}$$

The variance of the scaled total $(\mathcal{T}/\mathcal{T})$ is, by definition, zero. The covariation between $\mathcal{T}$ and $R_i$ is negligible for weak connections but substantial for strong connections. For example, given reasonable assumptions about the nature of the covariance, standard deviations calculated from scaled data on connections containing 50%, 25% and 10% of labelled neurons will be 0.5, 0.7 and 0.85 of their true values. Fortunately, it is possible to estimate and correct for this scaling bias (see § 3(e)).

### (b) *Inter- and intra-individual samples*

For each corticocortical or thalamocortical connection, we calculated the mean and standard deviation of the proportion of retrogradely labelled neurons following tracer injections. These sample statistics were computed from data for the same connection repeatedly measured within single published studies or single laboratories. We did not make any composite samples using data from different laboratories even when the same connections were measured. This is because any differences in the proportion of cortex searched for labelled neurons or differences in counting methods could introduce errors into our calculations.

To separate the within-individual from the inter-individual cases, we divided the results of studies where more than one tracer was injected into an animal, into two kinds of subsamples (table 1). The first inter-individual subsamples (table 1*a*) were arranged so that they did not contain more than one measure from any single animal. The variability of these samples provides an estimate of inter-individual variability. The within-individual subsamples (table 1*b*) consisted of paired injections in the same area of the same hemisphere of the same animal. These samples provide an estimate of within-individual variability.

## 3. RESULTS

By pooling data from our laboratories and from the literature, we were able to obtain quantitative connection data on the relative densities of 130 corticocortical and 54 thalamocortical connections. Quantitative connectional neuroanatomy is extremely laborious, so sample sizes for each connection were small. There were typically three to five tracer injections per connection, with a range of two to ten. For the data from our laboratories we found between 1000 and 25 000 retrogradely labelled neurons in the thalamus or cortex per tracer injection. Because of the small number of injections per connection, no single connection provides enough data for a good estimate of the distribution of densities.

### (a) *Great variability in individual connection densities*

The high degree of variability in the densities of individual corticocortical and thalamocortical connections is illustrated in figure 1. Figure 1*a* shows cases where the density of the same thalamocortical connection varies

Table 1. *Distinguishing between within-individual and inter-individual variability*

The table shows how we treat data from studies that made two tracer injections within single individuals. Here, two animals (cat 69 and cat 89) each received two tracer injections (Bb and NY) and the proportions of labelled neurons in thalamic nuclei were recorded. (*a*) To estimate the inter-individual variability in connection density (containing within- plus between-animal factors), we did not average across all four injections. Rather, we produced two subsamples in which neither animal was represented more than once. Means and standard deviations were then computed for the subsamples. In practice, such subsamples contained data from two to seven animals. (*b*) To estimate the intra-variability in connection density (containing only within-animal factors), we did not average across all four injections. Rather, we produced subsamples, each of which contained data from only one animal. Means and standard deviations were then computed for the subsamples. The size of these subsamples was always two, because no more than two distinguishable tracers (NY and Bb) were ever injected into the same individual.)

| | cat | | | | | |
|---|---|---|---|---|---|---|
| | 69 | 69 | 89 | 89 | | |
| | injection site | | | | | |
| | CGa | CGa | CGa | CGa | | |
| | tracer | | | | | |
| te of label | Bb | NY | Bb | NY | mean | s.d. |
| *) Within- and between-animal (Musil & Olson 1992)* | | | | | | |
| VAm/VMP | 0.055 | — | 0.032 | — | 0.044 | 0.016 |
| MD | 0.029 | — | 0.068 | — | 0.049 | 0.028 |
| P | 0.269 | — | 0.260 | — | 0.265 | 0.006 |
| M | 0.292 | — | 0.372 | — | 0.332 | 0.057 |
| H | 0.035 | — | 0.065 | — | 0.050 | 0.021 |
| E | 0.136 | — | 0.066 | — | 0.101 | 0.049 |
| M | 0.010 | — | 0.039 | — | 0.025 | 0.021 |
| Ad | 0.046 | — | 0.039 | — | 0.043 | 0.005 |
| MB | 0.000 | — | 0.000 | — | 0.000 | 0.000 |
| VAm/VMP | — | 0.112 | — | 0.117 | 0.115 | 0.004 |
| MD | — | 0.040 | — | 0.118 | 0.079 | 0.055 |
| P | — | 0.139 | — | 0.291 | 0.215 | 0.107 |
| M | — | 0.220 | — | 0.152 | 0.186 | 0.048 |
| H | — | 0.036 | — | 0.031 | 0.034 | 0.004 |
| E | — | 0.045 | — | 0.040 | 0.043 | 0.004 |
| M | — | 0.202 | — | 0.040 | 0.121 | 0.115 |
| Ad | — | 0.112 | — | 0.125 | 0.119 | 0.009 |
| MB | — | 0.000 | — | 0.000 | 0.000 | 0.000 |
| *) Within-animal (Musil & Olson 1991)* | | | | | | |
| VAm/VMP | 0.055 | 0.112 | — | — | 0.084 | 0.040 |
| MD | 0.029 | 0.040 | — | — | 0.035 | 0.008 |
| P | 0.269 | 0.139 | — | — | 0.204 | 0.092 |
| M | 0.292 | 0.220 | — | — | 0.256 | 0.051 |
| H | 0.035 | 0.036 | — | — | 0.036 | 0.001 |
| E | 0.136 | 0.045 | — | — | 0.091 | 0.064 |
| M | 0.010 | 0.202 | — | — | 0.106 | 0.136 |
| Ad | 0.046 | 0.112 | — | — | 0.079 | 0.047 |
| MB | 0.000 | 0.000 | — | — | 0.000 | 0.000 |
| VAm/VMP | — | — | 0.032 | 0.117 | 0.075 | 0.060 |
| MD | — | — | 0.068 | 0.118 | 0.093 | 0.035 |
| P | — | — | 0.260 | 0.291 | 0.276 | 0.022 |
| M | — | — | 0.372 | 0.152 | 0.262 | 0.156 |
| H | — | — | 0.065 | 0.031 | 0.048 | 0.024 |
| E | — | — | 0.066 | 0.040 | 0.053 | 0.018 |
| M | — | — | 0.039 | 0.040 | 0.039 | 0.000 |
| Ad | — | — | 0.039 | 0.125 | 0.082 | 0.061 |
| MB | — | — | 0.000 | 0.000 | 0.000 | 0.000 |

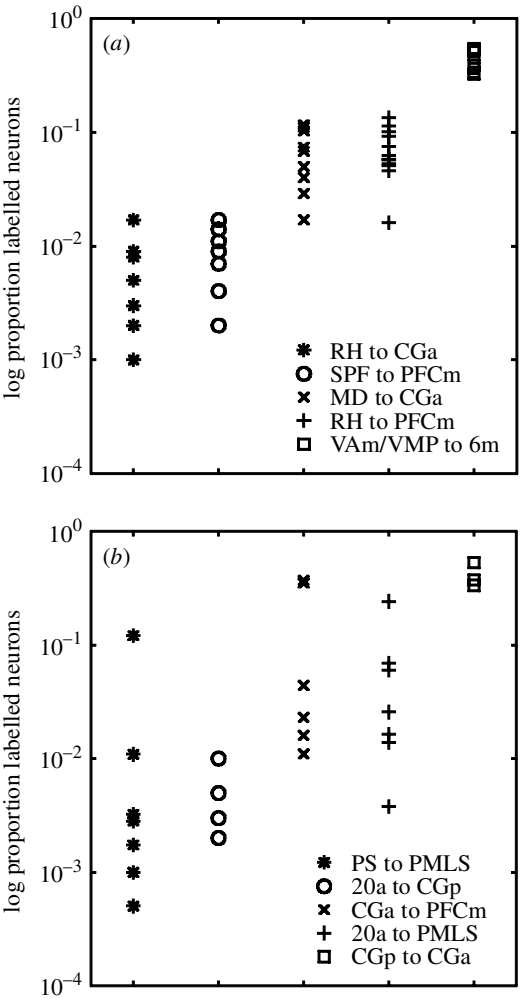
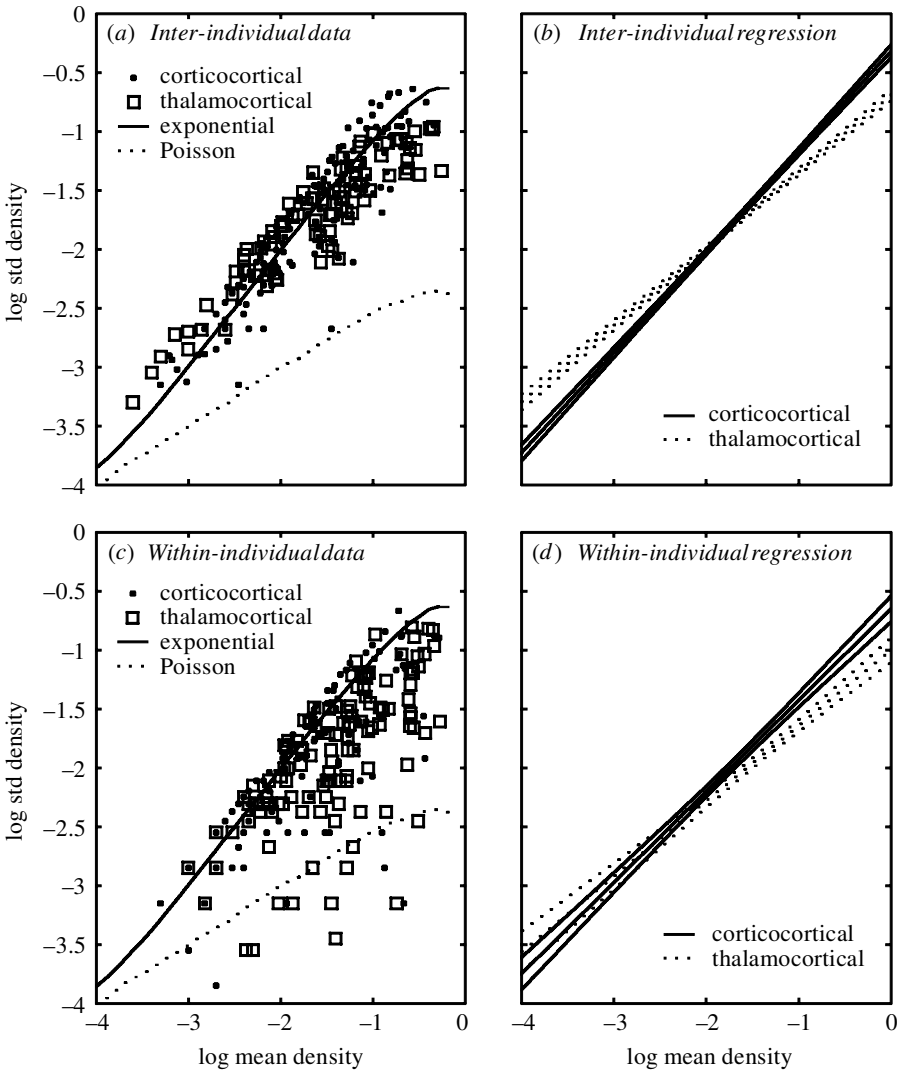



Figure 1. Typical examples of variability in extrinsic thalamocortical (*a*), and corticocortical (*b*), connection density. The vertical axis shows the log of the proportion of labelled neurons in an afferent area or nucleus following a retrograde tracer injection. Each point represents the proportion of total labelled cortical or thalamic neurons following a single retrograde tracer injection in a single individual. The mean density of the strong connections (e.g. CGp to CGa) may be over 1000 times that of the weak connections (e.g. PS to PMLS). There is also great variability in the proportion of labelled neurons for the same connection across different tracer injections. For example, the proportion of labelled cortical neurons in cortical area PS following injections in PMLS, and the proportion of labelled neurons in cortical area CGa following injections in PFCm, vary over 100-fold (*a*). Similarly, the proportion of labelled thalamic neurons in RH and MD following injections in CGa varies over tenfold (*b*).

over a factor of ten between different tracer injections. Figure 1*b* shows cases where the density of the same corticocortical connection varies over a factor of 100 between different tracer injections. In both the thalamocortical and corticocortical cases, the mean densities of the stronger connections may be over 1000 times that of the weaker connections.

### (b) *Relationship between mean and standard deviation of connection density*

To explore the relationship between the variability of connections and their mean density, we performed a



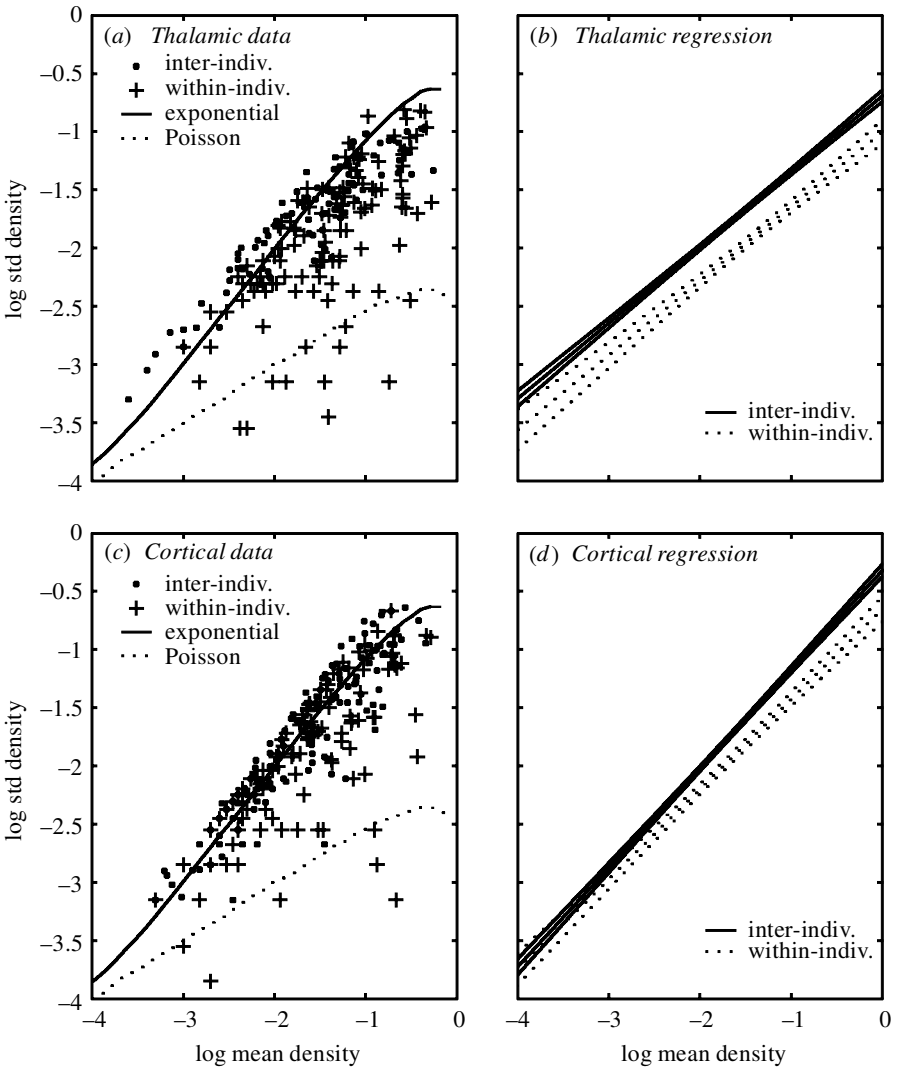

igure 2. Relationship between mean and standard deviation of thalamocortical and corticocortical connection density. The ertical axes show log standard deviation of the proportion of labelled neurons. The horizontal axes show log mean proportion of abelled neurons. (*a*) and (*c*) show raw data from inter- and within-individual cases, respectively. Each point was computed from wo or more tracer injections. The within-individual data (*c*) is more scattered because sample sizes were never greater than two. he solid and dotted lines show the relationship between mean and standard deviation for samples from an exponential and a oisson distribution, respectively. As connection data were scaled (see §2), we also scaled the data used to compute the lines for he exponential and Poisson. The scaling process reduces estimates of variability for strong connections. (*b*) and (*d*) show egression lines (±1 s.e.) for inter-individual and for within-individual data, respectively. The gradient of the regression for orticocortical data is higher than that for thalamocortical data (see also table 2). Strong corticocortical connections tend to be nore variable in density than strong thalamocortical connections, although the situation may be reversed for weak connections.

reliminary regression analysis of standard deviation ersus mean for thalamocortical and corticocortical onnections. Densities vary over several orders of magnitude, so we used the logs of the mean and standard deviation. We distinguished within-individual statistics (two istinguishable tracer injections in the same area in the ame animal) from inter-individual statistics (no more han one tracer deposit in any animal). Figure 2 ompares thalamic and cortical data. Figure 3 compares nter- and within-individual data.

Because of possible sampling bias and scaling bias in omputing standard deviation (see §3(e)), the uncorected regression analyses in figures 2 and 3 must be reated with caution. However, the figures show several obust features of the data that we consider briefly here nd return to later.

First, for stronger connections (more than 1% of labelled neurons, figure 2), thalamocortical variability is lower than corticocortical variability. The regressions suggest, however, that the relationship may be reversed for very weak connections where thalamic projections appear more variable. Second, for both corticocortical and thalamocortical connections, inter-individual standard deviation is moderately higher (roughly 0.3 log units for the uncorrected data) than within-individual standard deviation across a wide range of connection densities (figure 3).

Third, and importantly, log standard deviation connection density is roughly proportional to log mean connection density over several orders of magnitude (figures 2 and 3, table 2). The lines superimposed on the raw data points represent the relationships between mean and




igure 3. Relationship between mean and standard deviation of connection density for within-individual and inter-individual ases. The vertical axes show log standard deviation of the proportion of labelled neurons. The horizontal axes show log mean roportion of labelled neurons. (*a*) and (*c*) show raw thalamic and cortical data, respectively. Each point was computed from wo or more tracer injections within an individual study. The intra-individual data (*a*) and (*c*) are more scattered because sample zes were never greater than two. The solid and dotted lines show the relationship between mean and standard deviation for amples from an exponential and a Poisson distribution, respectively. As connection data were scaled (see § 2), we also scaled the ata used to compute the lines for the exponential and Poisson. The scaling process reduces estimates of variability for strong onnections. (*b*) and (*d*) show regression lines (± 1 s.e.) for inter- and for within-individual data, respectively. Within-individual ata show a similar slope to inter-individual data in both (*b*) and (*d*) but has a lower intercept (see also table 2), indicating roportionately lower variability.

tandard deviation that would be expected from a simple 'oisson process and from an exponential distribution figures 2*a*,*c* and 3*a*,*c*). Both of these theoretical lines were omputed taking account of scaling bias, and this xplains why their slopes decline for strong connections. t is clear that the Poisson model predicts unrealistically ow connection variability. In contrast, the line repre- enting samples drawn from an exponential distribution rovides a good description of the relationship between ariability and mean.

### (c) *Relationship between mean and median connection density*

The relationship between mean and standard deviation onnection density does not look like that predicted by a 'oisson process, but does resemble an exponential. However, figures 2 and 3 could, in principle, show that connection densities follow a normal distribution in which the standard deviation scales in proportion to the mean.

Normally distributed data are not skewed. A robust measure of skew, particularly when sample size is small, is the ratio of median to mean. This ratio will be centred on one for samples drawn from a normal distribution. It will be less than one for positively skewed data and more than one for negatively skewed data.

To explore the samples of connection data, we calculated the ratio of sample median to sample mean for all sample sizes greater than two (the minimum sample size necessary for mean and median to be different). This included the vast majority of inter-individual samples but, unfortunately, excluded all within-individual samples. Figure 4 shows that, with the exception of one data point (representing data from a single study of thalamocortical connections, see § 4), the median to mean ratio is

Table 2. *Coefficients for linear regression between log standard deviation (s) and log mean ($\bar{x}$) given by the equation* $\log s = k\log\bar{x} + c$, *where k is the slope and c is the intercept*

The table also shows the standard error in the slope and intercept estimates, and the goodness of fit, $R^2$, of the regression. The expected standard deviation for any mean density is given by $s = \log 10^{(c+k\bar{x})}$. (*a*) Results of regressions of raw data in which estimates of standard deviation have not been corrected for sampling bias or scaling bias; (*b*) results of regressions of corrected data where estimates of standard deviation have been corrected for sampling bias and scaling bias.)

| | *k* | s.e. | c | s.e. | $R^2$ |
|---|---|---|---|---|---|
| (*a*) *uncorrected* | | | | | |
| within thalamus | 0.64 | 0.06 | −1.00 | 0.10 | 0.45 |
| between + within thalamus | 0.65 | 0.03 | −0.69 | 0.05 | 0.87 |
| within cortex | 0.77 | 0.06 | −0.65 | 0.11 | 0.61 |
| between + within cortex | 0.85 | 0.03 | −0.32 | 0.05 | 0.87 |
| (*b*) *corrected* | | | | | |
| within thalamus | 0.68 | 0.10 | −0.63 | 0.10 | 0.53 |
| between + within thalamus | 0.73 | 0.02 | −0.46 | 0.05 | 0.90 |
| within cortex | 0.84 | 0.05 | −0.35 | 0.10 | 0.66 |
| between + within cortex | 0.90 | 0.03 | −0.09 | 0.05 | 0.89 |

significantly less than one, and is closer to 0.7. Therefore, the distribution of connection densities is highly non-normal and strongly positively skewed.

### (d) *The distribution of connection densities*

The results so far force us to reject two candidate distributions that could, in principle, describe the spread of corticocortical and thalamocortical connection densities. First, the ratio of mean to median shows that the distribution is highly non-normal (with the possible exception of thalamic projections to MS, see § 4). Second, variability is too high to be accounted for by a simple Poisson model.

One frequency distribution that resembles our data in terms of skew is the geometrical distribution. When the mean number of events (e.g. labelled neurons) is large, as in all but the weakest of anatomical connections, the geometric distribution approximates its continuous analogue, the exponential or Boltzmann distribution. The exponential distribution is defined in equation (4), where the probability of obtaining a score $x$ is $p(x)$ and $\mu$ is the mean of the distribution.

$$p(x) = \frac{1}{\mu}e^{-x/\mu}. \quad (4)$$

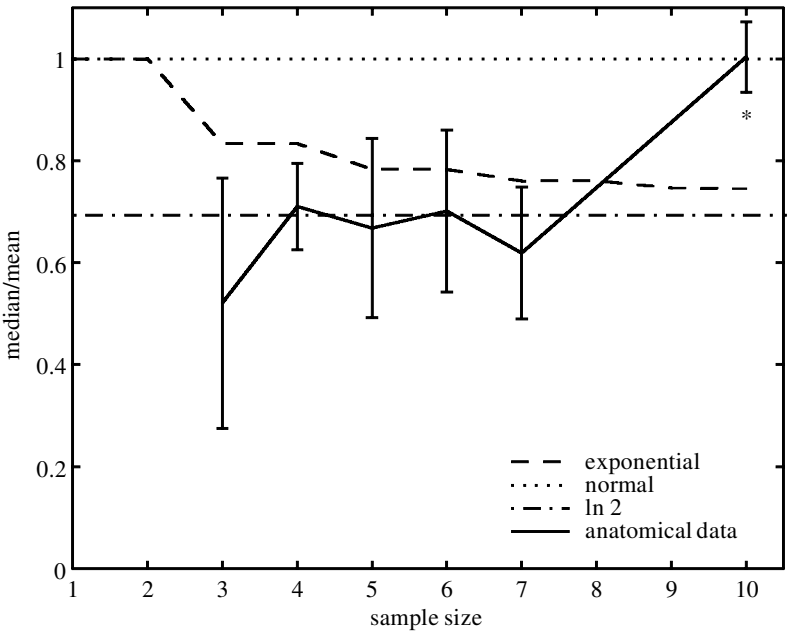
Figure 4 shows the ratio of median to mean for samples drawn from an exponential distribution for a range of sample sizes (dashed line in figure 4). Although the smaller experimental samples are more skewed than would be expected for an exponential, it is clear that the median to mean ratio of the connection data is much closer to an exponential than a normal distribution.

For the next section of this paper and in Appendix A, we adopt the exponential distribution as a simple 'working model' of the frequency distribution of connection densities. However, we qualify our use of the exponential for several reasons. First, the exponential approximation is not a perfect description of the variability that we find. This is shown by the fact that the slopes and intercepts of the regressions between standard deviation and mean are not one and zero, the values predicted for an exponential (table 2). As a result, the exponential underestimates variability for the very weak connections and overestimates variability for very strong connections. Second, figure 4 shows that the connection data appear even more skewed than predicted by an exponential distribution. This feature would be expected under a geometrical distribution (of which the exponential is the continuous analogue) for weak connections where $\mu^2$ is not much larger than $\mu$, the mean number of labelled neurons. These factors lead us to suggest that a more general class of frequency distributions known as the negative binomial (of which the geometrical distribution is a special case) probably provides a better description of the distribution of connection densities (see § 4). Negative binomials are attractive, as they arise under hierarchical generative models (Solomon 1983; Casella & Berger 1990), which might apply if a range of random processes are involved in the measurement of connection strengths with connection probabilities drawn from a continuous distribution. This would occur, for example, in a patchy cortical area containing domains that varied in their connection densities. While attractive, we do not develop a negative binomial model here for two reasons. First, it is not clear that the negative binomial distribution offers a sufficient improvement, or substantially alters our results or conclusions. Second, a large amount of quantitative data are required to distinguish between competing models. Such data should come from a purpose-designed study carried out with consistent methods in a single laboratory, and is beyond the quality of the data that are currently available.

The implications of an exponential distribution of connection densities are as follows. First, methods of inferential statistics based on the assumption of a normal distribution simply do not apply. Second, connection data are highly skewed and highly variable so rather large amounts of data are necessary for confident estimates of connection density and/or variability (see Appendix A, and figures 8 and 9). Third, and more deceptively, the standard measure of variability, sample standard deviation, systematically underestimates population standard deviation when given small samples drawn from highly skewed distributions such as the exponential (figure 8*b*). This bias depends on sample size, so estimates of variability based on small samples will be lower than estimates based on large samples. We consider sampling bias below and in greater detail in Appendix A (see also figure 8*b*).

### (e) *Sampling bias and scaling bias*

We performed a Monte Carlo simulation to assess the effect of sampling bias on the standard deviation of samples drawn from exponentially distributed data. We took 10 000 samples in a range of sample sizes (from two to ten) from an exponentially distributed population of random numbers with a population mean of one and standard deviation of one. Sampling bias in standard deviation was therefore equal to the mean value of sample




igure 4. The median proportion of labelled neurons is systematically less than the mean proportion of labelled neurons. The ertical axis shows the ratio of median to mean connection density; a robust measure of skew. The horizontal axis shows sample ze (the number of tracer injections; one per animal) from which the median to mean ratio for each connection was computed. rror bars show the 95% confidence intervals of the mean ratio of sample median to sample mean. The dotted line is the xpected median to mean ratio for samples drawn from a normal distribution. The dashed and dotted line is the expected nedian to mean ratio for samples drawn from an exponential distribution, which tends towards $\log_e 2$ (dashed line) when sample ze is large. In general, the empirically derived skew is closer to that expected by sampling from an exponential than from a ormal distribution. In fact, the data are even more skewed than an exponential distribution (see text). One data point, marked y an asterisk, is an obvious outlier. These data, which appear normally distributed, are from ten tracer injections in a single tudy of the thalamic projections to MS cortex (see § 4).

tandard deviation for each sample size (see Appendix A, olid line in figure 8*b*).

Sampling bias was substantial for small samples. For xample, with sample sizes of two, three and four, sample tandard deviation was only 0.70, 0.80 and 0.84 times the rue population standard deviation. We corrected the neasures of variability computed from anatomical data y dividing the sample standard deviation by sampling ias for the appropriate sample size.

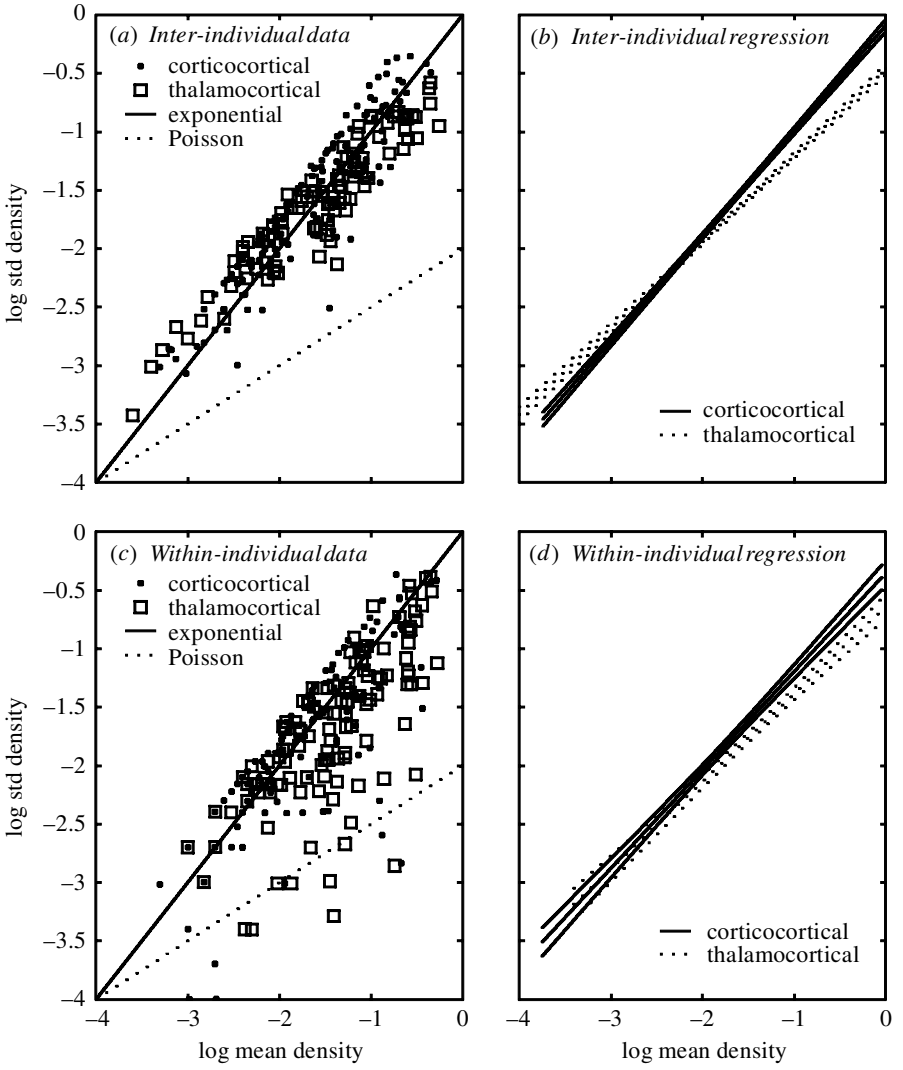
Scaling measures of connection density, by dividing cell ounts for individual connections by the total count, also educes estimates of the variability of connections. The caling bias is particularly severe for strong connections ecause these covary strongly with the total count. caling bias is intuitively understandable if one thinks of very strong connection that contains nearly all the abelled neurons. For such a connection, the proportion of otal labelled neurons will always be close to one, whatever the absolute variations in its density. Therefore, this onnection will appear to have a very low level of variaility across cases.

Scaling bias in estimates of variability occurs whenever ndividual measures are divided by a summed measure with which they covary, but the details of the bias vary with he nature of the covariance between individual measures nd the total. To obtain an accurate estimate of scaling bias n our data, we performed a Monte Carlo simulation that sed distributions that were very similar to those found in he experimental data. The model is outlined in equation 5). Here $S_c$ and $R_c$ are random variables that represent, respectively, the proportion and number of neurons labelled via connection $c$. In line with the anatomical data, the distribution of $R_c$ is exponential. $R_o$ represents the number of neurons labelled by all the other connections. As $R_o$ is the consequence of adding a large number of exponentially distributed random variables (i.e. the other connections), we assume that it is normally distributed.

$$S_c = \frac{R_c}{R_c + R_o}. \tag{5}$$

We ran our simulations with the mean total number of labelled neurons $T = R_c + R_o = 10\,000$. Therefore, the mean of $R_o = 10\,000 - R_c$. We found that scaling bias is only weakly dependent on total neuron number. The value of $T$ that we have chosen produces results that are representative for the range of values of $T$ (roughly 1000 to 20 000 labelled neurons) present in the data.

We took a large number of large samples in a range of connection strengths (mean $R_c$ was from 1 to 8000 labelled neurons, and mean $R_o$ was from 9999 to 2000 labelled neurons) to represent cell counts for connection $c$ and all the other connections. We then computed $S_c$ (equation (6)) and the standard deviation of $S_c$ for each mean connection strength. The results of the simulation (and of an equivalent simulation for Poisson data) are shown by the lines in figures 2*a*,*c* and 3*a*,*c*. In the absence of scaling bias, these lines would be straight, with slopes of 1 and 0.5, respectively. Equivalent lines for the Poisson and exponential, correcting for scaling bias, appear in figures 5*a*,*c* and 6*a*,*c*.



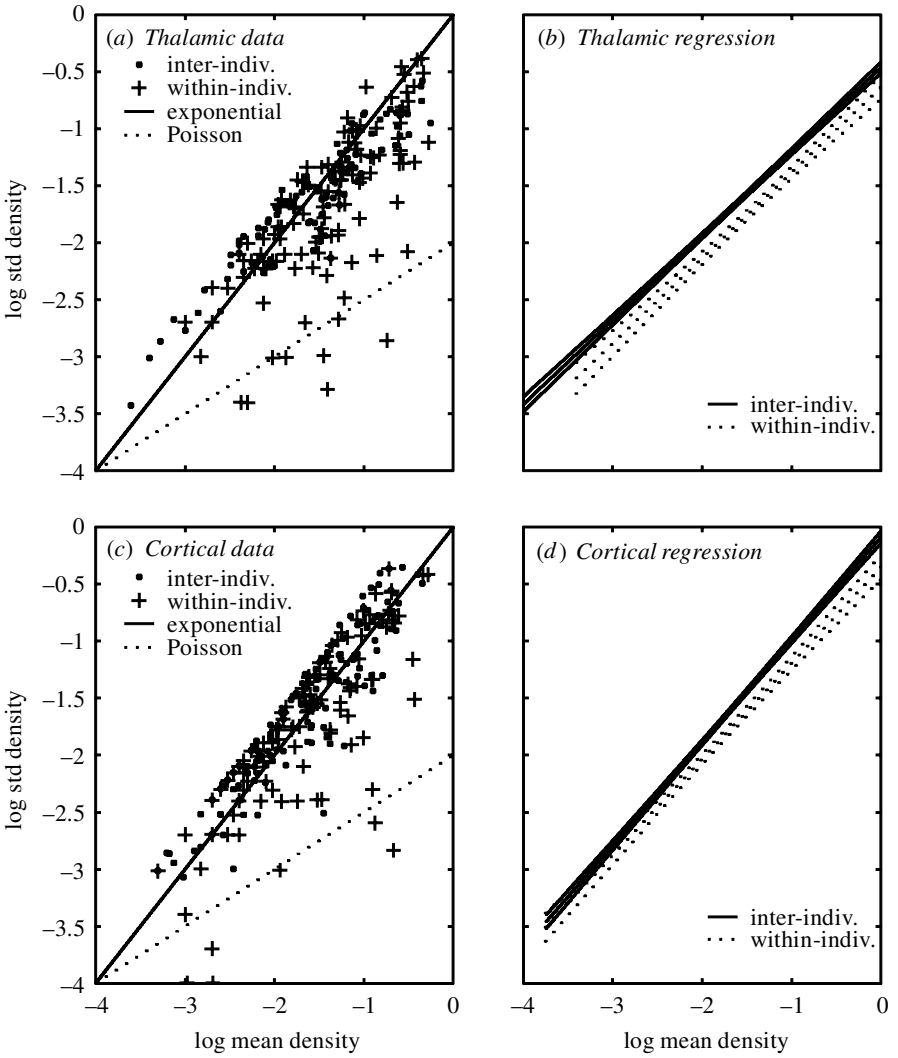

igure 5. Relationship between the mean and standard deviation of thalamocortical and corticocortical connection density when ampling bias and scaling bias in standard deviation are corrected. The vertical axes show the log of the standard deviation of he proportion of labelled neurons. The horizontal axes show the log of the mean proportion of labelled neurons. (*a*) and (*c*) show orrected data points for inter- and within-individual cases, respectively. Each point was computed from two or more tracer njections within an individual study. The solid and dotted lines show the relationship between mean and standard deviation for amples drawn from exponential and Poisson distributions, respectively. (*b*) and (*d*) show regression lines (± 1 s.e.) for inter- and or within-individual data, respectively. The gradient of the regression for corticocortical data appears higher than that for halamocortical data (see also table 2). Strong corticocortical connections tend to be more variable in density than strong halamocortical connections, although the situation may be reversed for weak connections.

A comparison of figures 2 and 3 with figures 5 and 6 hows that scaling bias is substantial for strong connections. or connections that account for 25% of the total number f labelled neurons, scaling bias will reduce standard devia- on estimates to roughly 0.66 of their true value. We omputed a correction factor corresponding to the differ- nce between the log standard deviation of scaling-biased xponential data (figures 2 and 3) and the log standard eviation of unbiased exponential data (figures 5 and 6). We hen added this factor to the standard deviation estimates of he connection data (figures 5 and 6).

### (f) *Comparison of corticocortical and thalamocortical variability, and of intra- and inter-individual variability: corrected data*

The data in figures 2 and 3 are likely to suffer from caling bias and sampling bias, which both tend to reduce the standard deviation. Both biases must be corrected to provide good estimates of the true relationship between mean and standard deviation connection density. This is necessary to reveal differences in corticocortical and thalamocortical connections, and to assess quantitative differences in within- and inter-individual variability. Figures 5 and 6 show the relationships between mean and standard deviation after appropriate correction. The results here are qualitatively similar, but quantitatively different, to figures 2 and 3.

First, strong corticocortical connections are more variable than strong thalamocortical connections (figure 5). For example, corticocortical connections with around 10% of labelled neurons tend to have a standard deviation 1.5–2 times greater than equivalent density thalamic connections. This is true for both inter- and within-individual cases. For connections that include




igure 6. Relationship between the mean and standard deviation of connection density for inter-individual and intra-individual ases when sampling bias and scaling bias in standard deviation are corrected. The vertical axes show the log of the standard eviation of the proportion of labelled neurons. The horizontal axes show the log of the mean proportion of labelled neurons. *a*) and (*c*) show corrected data points for thalamic and cortical data, respectively. Each point was computed from two or more racer injections within an individual study. The solid and dotted lines show the relationship between mean and standard eviation for samples drawn from exponential and Poisson distributions, respectively. (*b*) and (*d*) show regression lines (±1 s.e.) or thalamic and for cortical data, respectively. Within-individual data shows a similar slope to inter-individual data in both *b*) and (*d*) but has a lower intercept (see also table 2). The difference in intercept suggests that in both thalamocortical and orticocortical cases, the standard deviation of within-individual samples is roughly 1.5 times lower than the standard deviation f intra-individual samples with a comparable mean.

round 1% of labelled cells, thalamocortical and cortico-ortical connections have very similar variability. For veak connections (0.1% of labelled cells) inter-individual halamic connectivity may be more variable than cortical onnectivity, but within-individual variability is similar or thalamus and cortex.

Second, figure 6 confirms that inter-individual ariability is greater than within-individual variability. However, the differences with corrected data (figure 6) are maller than the estimates based on uncorrected data figure 3). For example, in the region of the corticocortical egressions (figure 6*d*) where the lines are significantly ifferent (roughly corresponding connection densities of .1% and upwards), the ratio of standard deviation of vithin- to inter-individual cases is only in the range of :1.3 to 1:1.5. Similarly, in the region of the thalamocortical egressions (figure 6*b*) where the lines are significantly different (roughly corresponding to connection densities of 0.1% and above), the ratio of standard deviation of within- to inter-individual ranges from 1:1.3 to 1:1.6.

Variance is the square of standard deviation, so standard deviation ratios can be converted into variance ratios. As the ratio of within- to inter-individual variance ranges from 1:1.7 to 1:2.6, the ratio of within-individual variance to true between-individual variance (i.e. inter-individual minus within-individual variance), ranges from 1:0.7 to 1:1.6. These values show that both true within-individual factors (local heterogeneity in tracer uptake, experimental error) and true between-individual factors (systematic differences between animals) contribute similar amounts of variance to the results of most connection tracing experiments.

Third, as with the uncorrected data, the exponential provides a good description of the data, while the Poisson

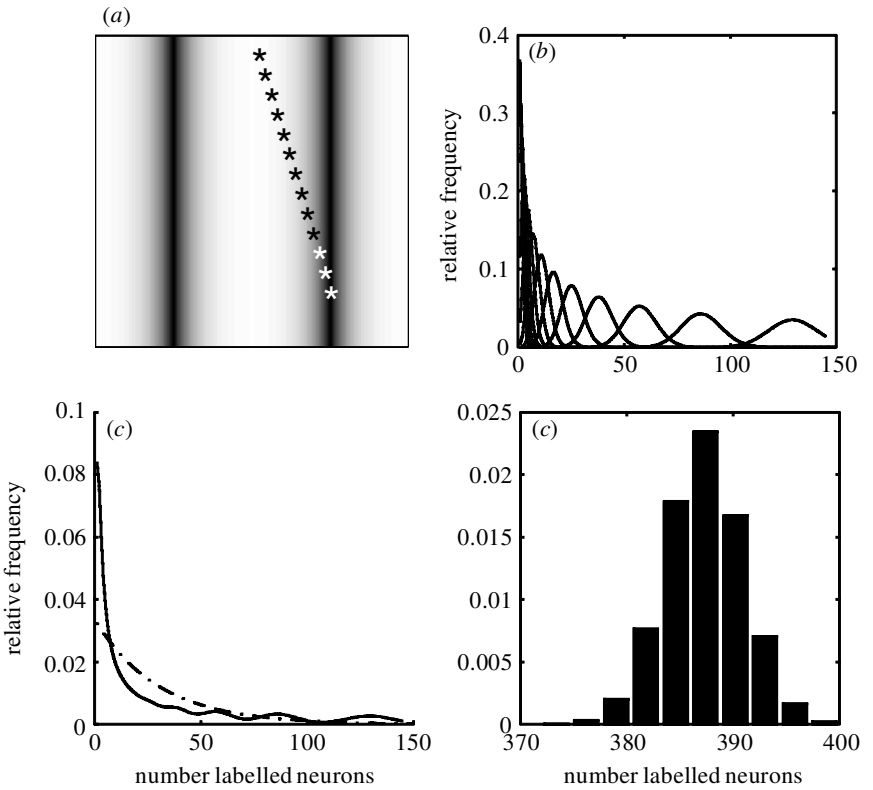



Figure 7. A model to account for the experimentally derived distribution of connection densities. (*a*) shows a surface view of a small region of a hypothetical cortical area. The area has striped variations in connection density with another structure. Light shades show weak connections and dark shades show strong connections. Possible locations of small tracer injections are marked by asterisks. (*b*) illustrates the frequency distribution of label density in a distant structure that would result from small injections made at the locations shown in (*a*). Each small curve represents a Poisson distribution of connection densities, that would occur if it were possible to repeatedly inject exactly the same location. (*c*) shows the frequency distribution that would result from a series of small injections at randomly selected asterisks (*a*). The solid line is the sum of the individual Poisson distributions shown in (*b*). The dotted line is an exponential with the same mean density as the solid line. It would be very difficult to distinguish between the exponential (dotted line) and the distribution based on random injections (solid line) on the basis of small samples. In fact, if the distribution of local mean densities within an area was gamma distributed, the combined distribution would be a negative binomial (see § 4). (*d*) shows the frequency distribution of label density in the distant structure from a large injection that covered all the asterisks. By averaging across all the domains, the distribution of densities from large injections will be very close to a normal.

substantially underestimates variability in the density of connections.

## 4. DISCUSSION

### (a) *Consequences of variability for neuroanatomical studies*

Our results have implications for interpreting the current neuroanatomical literature, for the design of connection tracing experiments and for the way connection data are reported. The first and most general implication is that caution is required when interpreting connection densities reported in studies with small sample sizes. We suggest that sample sizes of around ten are necessary for reasonable estimates of the density of connections. Our experience indicates that the majority of published studies on corticocortical and thalamocortical connections in cats and macaques use samples that are substantially smaller than ten. However, even with sample sizes of ten or more, only very large differences in mean connection density will prove 'significantly' different on a reliable basis (see Appendix A).

Second, quantitative connection tracing studies have tended to report scaled data, as such data make it easier to compare between cases. However, many statistics computed directly from scaled data will be misleading. Therefore, scaled data should be supplied with an indication of the total number of labelled cells, so that they may be easily 'unscaled'.

Third, as connection densities are highly variable, presenting any single individual's results as representative is difficult. This is because most individuals depart considerably from the average pattern and no single individual can be very similar to all the others. To reflect this genuine aspect of connection data, the distribution of connection densities should be quantified (Cherniak 1990), reported and represented in future attempts at collation and modelling (MacNeil *et al.* 1997).

### (b) *Thalamic projections to PMLS*

Both corticocortical and thalamocortical connection data typically show a skewed distribution of densities in which the median is substantially lower than the mean. The only obvious exceptions are the thalamocortical projections to MS from ten tracer deposits made in the laboratory of Dr Payne (asterisk in figure 4). The densities of these connections are roughly normally distributed, and show very low variability and very little skew.

everal factors could account for these 'exceptional' data. he first, and least interesting, is that they represent a tatistical quirk. It is possible that they differ from the ther data just by chance. This would not be altogether urprising, as we have computed the median to mean atios for a large number of samples.

Second, the difference could be due to the fact the tudy in Dr Payne's laboratory took several steps to educe random variability due to the spatial characteristics of tracer deposits (MacNeil *et al.* 1997). These are utlined in the § 2, but in addition to the usual precautions: i) tracer deposits were limited to a particular region of he visual field representation; (ii) tracer deposits covered reasonable area of the cortex, so avoiding differential abelling in subsets of afferent neurons with small or atchy terminal arborizations (Sherk & Ombrellaro 1988; 'ayne *et al*. 1991; Shipp & Grant 1991); (iii) all cortical ayers were exposed to tracer, so avoiding differential abelling in subsets of afferent neurons with different aminar terminations. These factors may be very important if the highly variable distribution of connection ensities is due to local heterogeneity in connectivity (see elow and figure 7). For patchy connection patterns, mall localized deposits could produce a near exponential istribution of densities, while larger deposits could yield more normal distribution (see figure 7 and below). This because the larger deposits could simultaneously span he different connectional subcompartments.

We note that the corticocortical data from Dr Payne's aboratory, which are based on the same set of tracer njections as the 'exceptional' thalamocortical data, bear a lose resemblance to the other corticocortical data. These acts could be explained if the corticocortical projections o MS cortex are more patchy (Montero 1981; Shipp & Grant 1991), or have coarser patches, than the thalamocortical projections (see Sherk 1986).

Third, the MS data raise the possibility of a difference n the thalamocortical connectivity of MS cortex and the ther areas for which we have data. MS is a relatively ow-order visual area, while the others (medial area 6, ingulate and prefrontal cortex) are all higher-order reas. The variability in the thalamic connections of these igher areas is more similar to most of the corticocortical onnections than to the thalamic data from MS. This bservation has several possible implications. First, halamic projections to the higher areas may be more atchy than thalamic projections to the lower areas. econd, epigenetic factors, which may contribute to a high egree of variability in corticocortical connections MacNeil *et al*. 1997), could play a greater role in shaping he thalamocortical connections of higher-order cortical reas.

### (c) *Interpreting the distribution*

The connections of MS, considered with the near exponential distributions that are observed in some other iological systems, may provide a clue to the generation f variability in the measurement of neuroanatomical rojections. Highly variable distributions of the kind we bserve in connection data are commonly found in the istribution of parasites in populations (Shaw *et al*. 1998; tear *et al*. 1998). In these cases, it can be assumed that for ny given individual, parasites follow a Poisson distribution. If all individuals were the same, this would result in a Poisson distribution of parasites across the population. However, the distribution of parasites appears much more clumped than the Poisson predicts. This pattern is obtained because parasites spread between nearby hosts, so that neighbouring animals have similar infection rates (a feature known as 'aggregation').

If instead of a simple Poisson process we have a distribution of Poisson processes with different mean rates, then this can lead to distributions very much like the one we observe in the connection data. This potentially unintuitive argument is shown much more simply in figure 7. Imagine that instead of host animals we have single tracer injections; instead of parasites we have labelled cells; and instead of 'aggregation' we have 'blobs' or 'stripes' in the cortex (Montero 1981; Symonds & Rosenquist 1984; Sherk 1986; Shipp & Grant 1991). Provided that cortical areas routinely contain patches with very different patterns of extrinsic connections, it is straightforward to understand how the distribution of densities that is observed with small injections (made into one stripe or patch), and assayed with small samples, resembles an exponential. If deposits were consistently made into the same stripe or patch in different experiments then we would observe a Poisson distribution. Alternatively, large tracer deposits that consistently cover an entire 'wavelength' of stripe or patch would generate a normal distribution of densities. These inferences have an obvious resonance with observed differences in variability in the corticocortical and thalamocortical projections to MS cortex.

### (d) *Measuring the distribution*

A challenging programme of empirical work is needed to put quantitative details on the simple model that we propose to account for variability. Until this is done, it will be very difficult to interpret within- and inter-individual differences, or differences in the distributions of connection densities obtained by injections in different areas. First, it is necessary to examine variations in connection density with the size and laminar distribution of tracer deposits, by making two or more deposits of distinguishable tracers nearby in the same cortical area of the same individual. Second, it will be necessary to determine the natures of within- and inter-individual variability. It requires injections of three or more distinguishable tracers into the same cortical area of a reasonable number individuals. A minimum of three tracers is required to provide information on the shape of the within-individual distribution. It is also necessary to repeat the procedures in several cortical areas to assess the variability in the nature of connectivity patterns from one region of cortex to another.

In conclusion, high variability appears to be a feature of all corticocortical and many thalamocortical connections. Variability presents a challenge for empirical neuroanatomy, for attempts to collate and analyse connection data, and for modelling studies. Connectional variability may be important for individual differences in behaviour, and can give us an insight into the local architecture of cortical areas. However, given the laborious nature of quantitative connection tracing, it is unlikely that variability will be properly addressed without




PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

istological and image-processing methods that allow abelled neurons to be counted automatically.

his work was supported by the Wellcome Trust (J.W.S.), US lational Institute of Neurological Diseases and Stroke and US lational Institute of Mental Health (B.P.). Thanks to Malcolm oung and Claus Hilgetag for helpful comments on the anuscript.

## APPENDIX A. PRACTICAL CONSEQUENCES OF A EAR EXPONENTIAL DISTRIBUTION OF CONNECTION DENSITIES FOR EMPIRICAL NEUROANATOMY

In this appendix, we consider some of the important onsequences of a near exponential distribution of onnection densities for experimental design in connec- onal neuroanatomy. As the exponential is the most vari- ble possible continuous distribution given non-negative alues, the estimates presented in this section are likely to epresent a relatively pessimistic, yet realistic, picture of he problems of sampling and statistical inference.

We provide three practical guides to data that are expo- entially distributed. First, it is easy to generate random xponentially distributed data from which to calculate tatistics (e.g. mean, median, standard deviation, etc.) and erform simulations. Numbers from an exponentially istributed population, $x_e$, with a mean and standard eviation of $m$ may be made by generating uniform andom numbers, $x$, between 0 and 1, and substituting hem into the following equation (equation (A1)).

$$_e = -\mu \ln (x). \qquad \text{(A1)}$$

econd, for simple guidance on the relationships between onfidence intervals, bias, and sample size, we provide wo graphs (figures 8 and 9). These were produced using quation (3) to generate 10 000 samples in each sample ze. The graphs show the 95% confidence intervals, 70% onfidence intervals (analogous to standard error), edian estimates, and bias for the mean (figure 8*a*) and tandard deviation (figure 8*b*), for sample sizes ranging rom two to 20. We note that the bias in standard devia- on (solid line in figure 8*b*) is worse for small samples. lso, confidence intervals are asymmetrical, unlike the ormally distributed case. Figure 8 illustrates the fact that rrors in estimates of connection strengths based on small amples can be very large. For example, the 95% confi- ence interval of the mean ranges over a factor of ten hen the sample size is five. Estimates of variability ased on small samples also show systematic bias.

Third, we provide an indication of the chance of orrectly identifying differences in the mean values of two opulations from which samples of connections are taken. igure 9 shows the probability that pairs of samples from opulations with different mean densities can be correctly dentified as 'significantly different' at the $p < 0.05$ level. he table was computed by sampling from exponentially istributed populations with known mean values. We udged pairs of samples as 'significantly different' if they ad completely non-overlapping 85% confidence intervals. e chose 85% confidence intervals, because pairs of amples drawn from exponential populations with the ame mean have non-overlapping 85% confidence inter- als just less than 5% of the time. Therefore, this

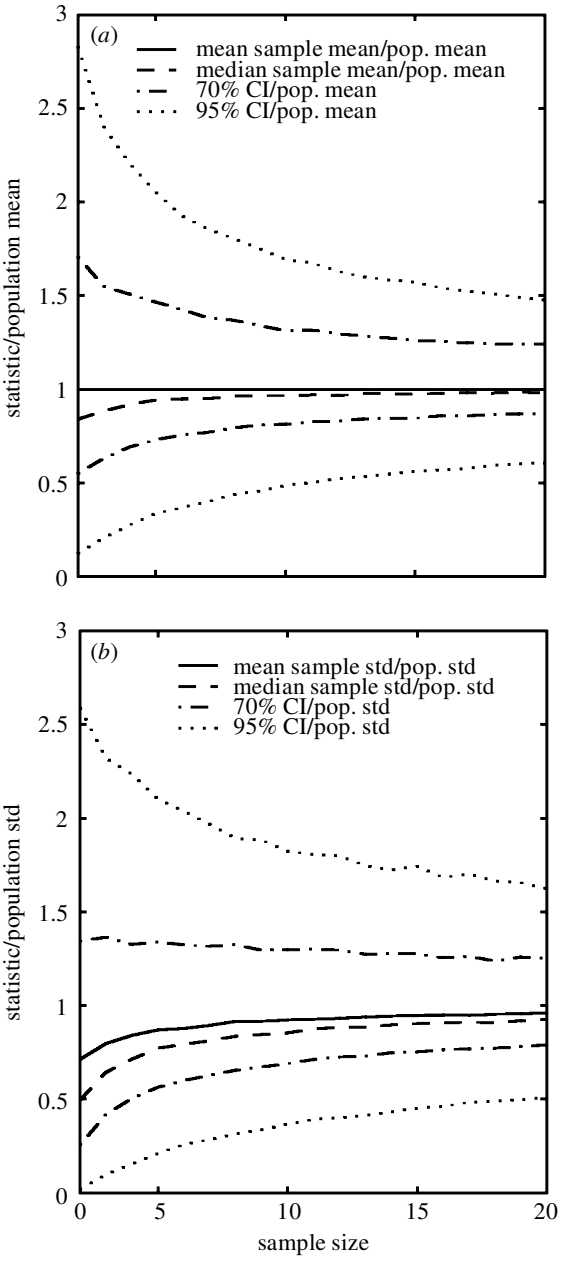
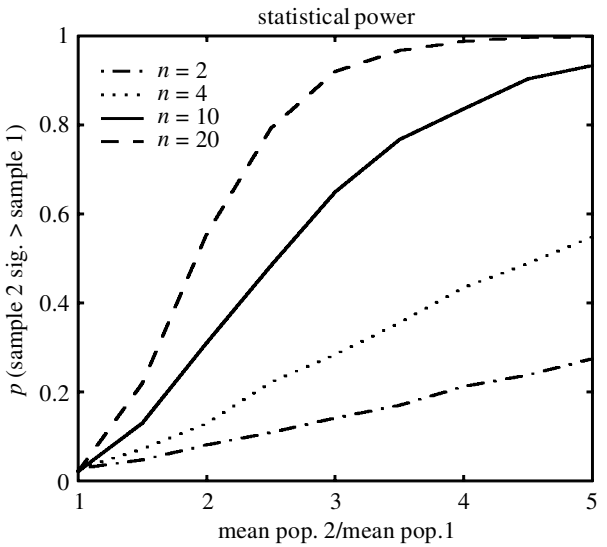



Figure 8. Sampling from an exponentially distributed population. The figure provides a guide to sampling bias, and the confidence of estimates of mean (*a*) and standard deviation (*b*) when sampling from an exponential distribution. The exponential distribution provides a better description of connection data than the normal distribution. The relationships are independent of mean connection density, so are expressed as a ratio of the relevant statistic (e.g. sample mean) to the true parameter (e.g. population mean). (*a*) shows the mean sample mean (solid line), median sample mean, 70% confidence intervals (analogous to standard error) and 95% confidence intervals for a range of sample sizes. Sample mean is an unbiased estimate of population mean ($y$-value of the solid line = 1), but the confidence intervals are wide. (*b*) shows mean sample standard deviation (solid line), the median sample standard deviation, the 70% confidence intervals for population standard deviation and the 95% confidence intervals. In contrast to sample mean, sample standard deviation is less than one, so systematically underestimates population standard deviation. This bias is particularly severe for small sample sizes.




igure 9. Statistical power of comparisons between samples rawn from exponential distributions. Samples 1 and 2 were rawn from populations 1 and 2. The vertical axis shows the roportion of pairs of samples in which sample 2 is identified s significantly greater than sample 1 (at $p < 0.05$). The orizontal axis shows the ratio of the means of population 2 nd population 1. A ratio of 1:1 indicates identical population eans, a ratio of 1:5 indicates a factor of five difference in opulation mean. The curves show the relationship between tatistical power and effect size for sample sizes of two (dots nd dashes), four (dots), ten (solid line) and 20 (dashes). We udged pairs of samples as 'significantly different' if they had ompletely non-overlapping 85% confidence intervals. We hose 85% confidence intervals, because pairs of samples rawn from exponential populations with the same mean have on-overlapping 85% confidence intervals just less than 5% of he time. Therefore, this corresponds to the conventionally ccepted significance level of a type I error rate of $p < 0.05$ for two-tailed test. The figure shows that for reasonable sample zes (e.g. ten individuals), connections have to have very ifferent connection densities to be reliably identified as gnificantly different.

orresponds to the conventionally accepted significance evel of a type I error rate of $p < 0.05$ for a two-tailed test. igure 9 shows that differences in population mean are ery difficult to reliably detect when sampling from an xponential distribution. In other words, type II error ates tend to be high. Given a sample size of four, which is ot unusual in connection tracing experiments, and a ratio n mean population connection density of 4:1, we would nly get a 'significant' difference around 40% of the time. his indicates that typical connection tracing experiments ay reliably distinguish only very large differences in the ean density of connections.

## REFERENCES

asella G. & Berger, R. L. 1990 *Statistical inference*. Belmont, CA: Wadworth.

herniak, C. C. 1990 The bounded brain: towards quantitative neuroanatomy. *J. Cogn. Neurosci.* **2**, 58–68.

owan, W. M., Gottlieb, D. I., Hendrickson, A. E., Price, J. L. & Woolsey, T. A. 1972 The autoradiographic demonstration of axonal connections in the central nervous system. *Brain Res.* **37**, 21–51.

De Yeo, E. A. & Van Essen, D. 1985 Segregation of efferent connections and receptive field properties in visual area V2 of the macaque. *Nature* **317**, 58–61.

Felleman, D. J. & Van Essen, D. C. 1991 Distributed hierarchical processing in the primate cerebral cortex. *Cerebr. Cortex* **1**, 1–47.

Gerfen, C. R. & Sawchenko, P. E. 1984 An anterograde neuroanatomical tracing method that shows the detailed morphology of neurons, their axons, and terminaimmunohistochemicallocalization of an axonally transported plant lectin, Phaseolus-vulgarisleukoagglutinin PHA-L. *Brain Res.* **290**, 219–238.

Grant, S. & Shipp, S. 1991 Visuotopic organization of the lateral suprasylvian area and of an adjacent area of the ectosylvian gyrus of cat cortex—a physiological and connectional study. *Vis. Neurosci.* **6**, 315–338.

Kristensson, K., Olsson, Y. & Sjostrand, J. 1971 Axonal uptake and retrograde transport of exogenous proteins in the hypoglossal nerve. *Brain Res.* **32**, 399–406.

Le Gros Clark, W. E. 1932 The structure and connections of the thalamus. *Brain* **55**, 406–470.

Le Gros Clark, W. E. 1942 The visual centres of the brain and their connections. *Phys. Rev.* **22**, 205–232.

MacNeil, M. A., Lomber, S. G. & Payne, B. R. 1997 Thalamic and cortical projections to the middle suprasylvian cortex of cats: constancy and variation. *Expl Brain. Res.* **114**, 24–32.

Marchi, V. & Algeri, G. 1895 Sulle degenerazioni discendenti consecutive a lesioni della corteccia cerebrale. *Nota Pre. Riv. Sper. Di Freniat.* **11**, 429.

Meynert, T. 1890 Über das Zusammenwirken der Gerhirntheile. *Verhandlungen des X. Internat. Mediz. Kongress. Berlin* **1**, 173–190.

Montero, V. M. 1981 Topography of cortico-cortical connections from the striate cortex in the cat. *Brain Behav. Evol.* **18**, 194–218.

Musil, S. Y. & Olson, C. R. 1988*a* Organization of the cortical and subcortical projections to the medial prefrontal cortex in the cat. *J. Comp. Neurol.* **272**, 219–241.

Musil, S. Y. & Olson, C. R. 1988*b* Organization of the cortical and subcortical projections to the anterior cingulate cortex in the cat. *J. Comp. Neurol.* **272**, 203–218.

Musil, S. Y. & Olson, C. R. 1991 Cortical areas in the medial frontal-lobe of the cat delineated by quantitative-analysis of thalamic afferents. *J. Comp. Neurol.* **308**, 457–466.

Nauta, W. J. H. & Gygax, P. A. 1954 Silver impregnation for degenerating axons in the central nervous system: a modified technique. *Stain Technol.* **29**, 91–93.

Olson, C. R. & Musil, S. Y. 1992 Topographic organization of cortical and subcortical projections to posterior cingulate cortex in the cat—evidence for somatic, ocular, and complex subregions. *J. Comp. Neurol.* **324**, 237–260.

Pandya, D. N. & Yeterian, E. H. 1985 Architecture and connections of cortical association areas. In *Cerebral cortex*. 4. *Association and auditory cortices* (ed. A. Peters & E. G. Jones), pp. 1–61. New York and London: Plenum Press.

Payne, B. R., Connors, C. & Cornwell, P. 1991 Survival and death of neurons in cortical area PMLS after removal or areas 17, 18, and 19 from cats and kittens. *Cerebr. Cortex* **1**, 469–491.

Polyak, S. 1927 An experimental study on the association, callosal, and projection fibres of the cerebral cortex of the cat. *J. Comp. Neurol.* **44**, 197–254.

Polyak, S. 1933 *The main afferent fiber systems in the cerebral cortex of primates*. Berkeley, CA: University of California Publ. Anat. vol. 2.

Raczkowski, D. & Rosenquist, A. C. 1983 Connections of the multiple visual cortical areas with the lateral posterior-pulvinar complex and adjacent thalamic nuclei in the cat. *J. Neurosci.* **3**, 1912–1942.

Rose, J. E. & Woolsey, C. N. 1948 Structure and relations of limbic cortex and anterior thalamic nuclei in rabbit and cat. *J. Comp. Neurol.* **89**, 279–438.

cannell, J. W. 1997 Determining cortical landscapes. *Nature* **386**, 452.

cannell, J. W., Blakemore, C. & Young, M. P 1995 Analysis of connectivity in the cat cerebral cortex. *J. Neurosci.* **15**, 1463–1483.

haw, D. J., Grenfell, B. T. & Dobson, A. P. 1998 Patterns of macroparasite aggregation in wildlife host populations. *Parasitology* **117**, 597–610.

herk, H. 1986 Coincidence of patchy inputs from the lateral geniculate complex and area 17 to the cat's Clare–Bishop area. *J. Comp. Neurol.* **253**, 105–120.

herk, H. & Ombrellaro, M. 1988 The retinogeniculate match between area 17 and its targets in visual suprasylvian cortex. *Expl Brain Res.* **72**, 225–236.

hipp, S. & Grant, S. 1991 Organization of reciprocal connections beteen area 17 and the lateral suprasylvian area of cat visual cortex. *Vis. Neurosci.* **6**, 339–355.

olomon, D. L. 1983 The spatial distribution of cabbage butterfly eggs. *Life science models*, vol. 4 (ed. H. Marcus-Roberts & M. Thompson), pp. 350–366. New York: Springer.

Stear, M. J., Bairden, K., Bishop S. C., Gettinby, G., McKellar, Q. A., Park, M., Strain, S. & Wallace, D. S. 1998 The processes influencing the distribution of parasitic nematodes among naturally infected lambs. *Parasitology* **117**, 165–171.

Symonds, L. L. & Rosenquist, A. C. 1984 Corticocortical connections among visual areas in the cat. *J. Comp. Neurol.* **229**, 1–38.

Young, M. P. 1993 The organization of neural systems in the primate cerebral cortex. *Proc. R. Soc. Lond.* B **252**, 13–18.

Zeki, S. & Shipp, S. 1988 The functional logic of cortical connections. *Nature* **335**, 311–317.

Zeki, S. & Shipp, S. 1989 Modular connections between areas V2 and V4 of macaque monkey visual cortex. *Eur. J. Neurosci.* **1**, 494–506.



PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES