PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Hierarchical organization of macaque and cat cortical sensory systems explored with a novel network processor

Claus C. Hilgetag, Marc A. O'Neill and Malcolm P. Young

Phil. Trans. R. Soc. Lond. B 2000 **355**, 71-89
doi: 10.1098/rstb.2000.0550

THE ROYAL SOCIETY

Hierarchical organization of macaque and cat cortical sensory systems explored with a novel network processor

Claus-C. Hilgetag*, Marc A. O'Neill and Malcolm P. Young

Neural Systems Group, Department of Psychology, University of Newcastle upon Tyne, Ridley Building, Newcastle upon Tyne NE1 7RU, UK

Neuroanatomists have described a large number of connections between the various structures of monkey and cat cortical sensory systems. Because of the complexity of the connection data, analysis is required to unravel what principles of organization they imply. To date, analysis of laminar origin and termination connection data to reveal hierarchical relationships between the cortical areas has been the most widely acknowledged approach. We programmed a network processor that searches for optimal hierarchical orderings of cortical areas given known hierarchical constraints and rules for their interpretation.

For all cortical systems and all cost functions, the processor found a multitude of equally low-cost hierarchies. Laminar hierarchical constraints that are presently available in the anatomical literature were therefore insufficient to constrain a unique ordering for any of the sensory systems we analysed. Hierarchical orderings of the monkey visual system that have been widely reported, but which were derived by hand, were not among the optimal orderings. All the cortical systems we studied displayed a significant degree of hierarchical organization, and the anatomical constraints from the monkey visual and somatomotor systems were satisfied with very few constraint violations in the optimal hierarchies. The visual and somato-motor systems in that animal were therefore surprisingly strictly hierarchical. Most inconsistencies between the constraints and the hierarchical relationships in the optimal structures for the visual system were related to connections of area FST (fundus of superior temporal sulcus). We found that the hierarchical solutions could be further improved by assuming that FST consists of two areas, which differ in the nature of their projections. Indeed, we found that perfect hierarchical arrangements of the primate visual system, without any violation of anatomical constraints, could be obtained under two reasonable conditions, namely the subdivision of FST into two distinct areas, whose connectivity we predict, and the abolition of at least one of the less reliable rule constraints.

Our analyses showed that the future collection of the same type of laminar constraints, or the inclusion of new hierarchical constraints from thalamocortical connections, will not resolve the problem of multiple optimal hierarchical representations for the primate visual system. Further data, however, may help to specify the relative ordering of some more areas. This indeterminacy of the visual hierarchy is in part due to the reported absence of some connections between cortical areas. These absences are consistent with limited cross-talk between differentiated processing streams in the system. Hence, hierarchical representation of the visual system is affected by, and must take into account, other organizational features, such as processing streams.

Keywords: visual hierarchy; cortical layers; anatomical connectivity; evolutionary optimization; area FST

1. INTRODUCTION

Structures in the central nervous system typically possess numerous connections with other structures. V1, for example, may exchange substantial numbers of projection fibres with at least 60 other cortical and subcortical regions (Young *et al.* 1995). Considering only the cerebral cortical part of the network, the average ipsilateral connectivity of cortical areas is about ten (Felleman & Van Essen 1991). The complexity of a network with this number of connections per node is too great to allow conclusions about its organization to be drawn by unaided intuition. Connectivity data, however, are widely believed to provide an important key to unravelling principles of brain organization. The complexity of the data, and the promise that they hold for providing insight into the organization of central nervous processing, have encouraged the development of several different data-analytical methods for treating connection data systematically.

A widely acknowledged approach is hierarchical analysis (Rockland & Pandya 1979; Maunsell & Van

*Author and address for correspondence: Boston University School of Medicine, Department of Anatomy and Neurobiology, 700 Albany Street W746, Boston, MA 02118, USA (claush@bu.edu).

Essen 1983; Felleman & Van Essen 1991; Coogan & Burkhalter 1993). This approach examines the patterns of laminar origin and termination of projections between cortical areas in the framework of some simple rules, which determine whether projections should be classified as ascending, descending or lateral. Application of these rules to empirical data yields a set of constraints which defines pairwise hierarchical relationships between areas. It is then possible to arrange the cortical areas into largely consistent hierarchies in which there are few violations of the pairwise hierarchical constraints. Hence, this analysis demonstrates that some sensory systems are hierarchically organized overall, and gives insight into the ordering of structures in each hierarchy (e.g. Felleman & Van Essen 1991).

Reanalysing the problem of hierarchical classification for data from the monkey visual system, we calculated that the total number, w, of possible hierarchies for 30 areas is greater than 10^{37} (Hilgetag *et al.* 1996), according to

$$\omega = \sum_{m=1}^{n} S_n^{(m)}(m!) = \sum_{m=1}^{n} \sum_{k=0}^{m} (-1)^{m-k} \binom{m}{k} k^n,$$

where $S_n^{(m)}$ is a Stirling number of the second kind describing all possible ways of partitioning a set of n elements (here $n = 30$) into m non-empty subsets (Abramowitz & Stegun 1972), where m ranges from 1 to 30 levels. k is an index variable and $(m!)$ stands for the factorial giving the number of all different orderings for the m levels. The primate somato-motor system (14 areas included in our analysis) and the cat visuo-limbic system (22 areas) carry smaller numbers of possible arrangements, but still yield in the order of 10^{14} and 10^{24} potential solutions, respectively. Considering the huge numbers of possible hierarchies, the incompleteness and partial inconsistency of the experimental data, and the use of a discrete cost function (number of violated rule constraints), we doubted that a uniquely optimal solution for any given data set could be obtained by hand.

Accordingly, we developed a computer program that could perform hierarchical analysis automatically. The large size of the solution space ruled out exhaustive analytical approaches, and the complexity and inconsistency of the anatomical data ruled out straightforward sorting methods. Instead, we employed an optimization approach, which was intended to find solutions with minimally few departures from a perfect hierarchy, in which all hierarchical constraints would be satisfied. Recently and increasingly, neural networks have been used in classification and optimization problems (e.g. Fausett 1994). The data available here, though, are not numbers but relational constraints, and the nature of the input data made conventional neural network methods difficult to apply. Neural networks typically take numerical vectors as input, compute over them by varying the weights of fixed connections, and offer a numerical vector as output, which can then be interpreted with reference to knowledge of the training history of the network. Considering the relational nature of the data, we decided to construct a novel system more appropriate to their characteristics. To do this, we caused a simulated network's structure itself, rather than just the activity of its elements, to reflect the relationships between the cortical areas. An algorithm then manipulated the structure of this network by simulated annealing until its structure optimally reflected the input rules. We constructed and tested the processor, and then applied it to the data from primate and cat cortical sensory systems. Preliminary results from the analysis of the primate cortical visual system with the simplest cost function were briefly reported previously (Hilgetag *et al.* 1996).

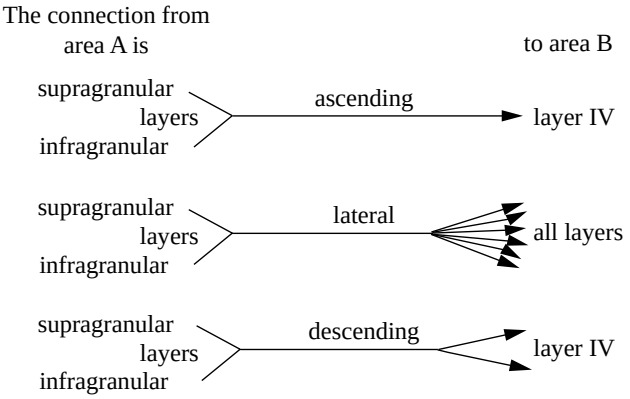

Figure 1. Hierarchical classification scheme for connections between areas in the monkey visual cortex, adapted from Felleman & Van Essen (1991).

2. METHODS

(a) *Data*

For the present analyses, we used the most comprehensive published collations of laminar origin and termination data for the primate cortical visual and somato-motor system (Felleman & Van Essen 1991) and the cat cortical visual system (Scannell *et al.* 1995). We used these published collations, rather than more recent and comprehensive databases we are developing, particularly because the data can be readily examined by other researchers to determine its reliability. We followed Felleman & Van Essen's (1991) classification of cortical connections on the basis of the laminar origin and termination of connections to derive hierarchical constraints on the relationships between linked areas (see figure 1). Connections proceeding from one area to another were classified as ascending if they originated in a unilaminar pattern from supragranular layers of the cortex, or bilaminarily from superficial and deeper layers, and terminated predominately in layer 4. Descending connections also arose bilaminarily from upper and deeper layers, and in a unilaminar pattern from infragranular layers, but tended to avoid layer 4 in their terminations. Connections considered as lateral originated in superficial and infragranular layers and terminated in a columnar pattern throughout the cortical layers (see figure 1). This classification scheme does not accommodate some 10% of the connections in the monkey visual system for which some laminar information is known, and these unclassified connections were excluded from the analysis. Webster *et al.* (1994), Cusick *et al.* (1995) and Barbas & Rempel-Clower (1997), among others, have questioned the extent to which this classification scheme can be applied universally to cortical regions, and have asked whether there should not be alternative schemes for certain parts of the cortex (e.g. frontal, prefrontal and polysensory regions). It should, moreover, be debated whether the same classification scheme as for connections in the macaque

brain can usefully be applied for categorizing connections in the cat brain (Scannell *et al.* 1999). Almost all data from the cat's visual system were derived by retrograde tracing methods, which do not provide clear indications of termination patterns, and so leave significant questions about whether cat connections are fitted by the classification scheme (Scannell *et al.* 1999). While we accepted the Felleman & Van Essen (1991) rules as the basis for study of cat connectivity here, we address these questions further in § 4.

We slightly modified the set of areas subject to hierarchical constraints (Felleman & Van Essen 1991, table 7) in that we excluded MIP and MDP, for which no hierarchical relationships were available. If a rule was given for the whole of an area that possesses subdivisions, we assumed that the constraint holds for all subareas, following previous authors (Felleman & Van Essen 1991). The Felleman & Van Essen (1991) compilation can be interpreted in a number of ways. We based our analyses mainly on their table 7, which lists Felleman & Van Essen's interpretation of the anatomical data in terms of hierarchical constraints for visual areas. In this interpretation, a bilateral, hierarchical relationship between two areas could arise from only a single connection between the areas. For instance, there is a clearly documented connection from VP to FST (Felleman & Van Essen 1991), based on information from Boussaoud *et al.* (1990) and Burkhalter & Van Essen (1983), and although the back projection from FST to VP appears to be missing (Boussaoud *et al.* 1990), two hierarchical constraints, VP < FST and FST > VP, were assigned to the pair of areas. (Other cases where there is a hierarchical constraint for two areas that appear to be unconnected anatomically are V1 to V4t, V4t to MSTd, MSTl to V1, MSTl to DP, MSTl to V4t, FST to VP and FST to V2; Felleman & Van Essen 1991.) Alternatively, one can interpret Felleman & Van Essen's collection of laminar data in their table 5 with the help of their classification scheme, which yields a slightly different set of fewer hierarchical constraints. The different data sets, however, led to similar results, as will be shown below.

The data for the primate somato-motor system were also taken from Felleman & Van Essen (1991), and analysed both for the paired hierarchical relationships suggested by Felleman & Van Essen (1991, table 8, column 11) and for individually classified laminar patterns. The data set from this second approach is displayed in figure 3. Hierarchical constraints for the cat visuolimbic system were obtained from the database provided in conjunction with Scannell *et al.* (1995). The latter hierarchical rules are listed in figure 4.

Additionally, we used a range of modified tables of constraints as input for the processor in further analyses of the primate visual system. The variations included 'merging' areas PITd and PITv, CITd and CITv, and STPp and STPa into combined areas PIT, CIT and STP, respectively. For some computer experiments, we also excluded constraints that were labelled with a question mark in figure 2, indicating that they might not be fully reliable due to varying degrees of inconsistency or interpretability in the neuroanatomical data (Felleman & Van Essen 1991). Table 1 gives an overview of the data sets employed in the additional studies.

(b) *The optimization algorithm*

We developed a computer program to order areas according to the given constraints into optimal hierarchical structures. This processor was programmed in ANSI C/C++ to run on UNIX workstations and possesses three key features. Data are stored in the form of objects to which relationships with other objects can be attached; there are routines to rearrange the network of objects; and routines for evaluating each rearrangement according to a cost function. This architecture makes it possible to use high-level rules directly as input to the processor, and to obtain output in an easily understandable form. A full technical account of the system will be presented elsewhere. The program was run on the most powerful UNIX machines available to us, but even so, computing large sets of solutions often required a number of months.

For lack of a suitable definition for a continuous cost, we used discrete cost functions, like those employed in previous work (e.g. Felleman & Van Essen 1991). The simplest cost function calculated the number of violated rules for a given hierarchy, and the goal of the optimization was then to find solutions that minimized this function. We also used a more complex cost function that additionally considered the numbers of hierarchical levels crossed by a particular rule violation, so that violations between areas separated by multiple hierarchical levels were more costly than those between hierarchical neighbours. The procedures for rearranging the areas within a hierarchy are based on an algorithm resembling cumulative evolution and simulated annealing (Laarhoven & Aarts 1987). We now turn to a detailed description of this method.

(i) *Initialization*

A random assignment of all areas to certain levels of an arbitrary hierarchy simulated a 'heating' of the initial structure. Each heated structure served as a starting point for an epoch, in which up to 2500 (a value related to the capacity of available computer systems) different optimal structures could be found. For performing comprehensive searches, numerous epochs (up to 15 000) were computed to search for possible optimal solutions sufficiently thoroughly. The number of hierarchical levels was arbitrarily chosen and modified by the program, although computations could be restricted to optimal hierarchies with a given number of levels, if required.

(ii) *Evolution*

Mutation

A structure existing at some point of the computation was modified to a descendant structure by one of the following mutations. An area could be assigned to another hierarchical level, which could already exist or was thus newly created, or two areas on different levels could be swapped. These simple mutations ensured that descendant structures varied only minimally from their parents so that the search space was examined at very high resolution.

Evaluation

The value of a hierarchical structure produced during an epoch was measured in units determined by the prevailing cost function. As a default, a weight of equal importance was assigned to each of the different constraints regardless of its hierarchical relationship or reliability. The processor could, however, accommodate different weights for reliable and for possibly less reliable constraints (see § 3(b)).

Selection

If a descendant structure possessed a cost within a range of up to 125% of the parental cost, it was considered for further optimization, given that in turn a descendant structure within the same range could be found for it. This optimal upper bound value was

areas	V1	V2	VP	V3	PIP	V3A	V4	PO	MT	V4t	DP	VOT	LIP	VIP	MSTd	MSTl	PITd	PITv	7a	STPp	CITd	CITv	STPa	AITv	FEF	TF	46	FST	TH	AITd
V1		<	Ø	<	<	<	<	<	<	<	Ø	Ø	Ø	Ø	Ø	<?	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø
V2	>		<	<		<	<	<	<	<	Ø		Ø		<	<	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	<	Ø	Ø	<	Ø	Ø
VP	Ø	>			<	≤	<	<	<	Ø		<	<	<	<	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	<	<	Ø	<	Ø	Ø
V3	>	>			<	<	<	<	<	<		Ø	<	<	<	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	<	<	Ø	<	Ø	Ø
PIP	>		>	>			<	<			<						Ø	Ø		Ø	Ø	Ø	Ø	Ø						Ø
V3A	>	>	≥	>			<		≤		<		<		<	<	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	<	Ø	Ø	<	Ø	Ø
V4	>	>	>	>	>	>		Ø	≤	≤	<	<	<	Ø	Ø	Ø	<	<	Ø	Ø	<	<	Ø	<	<?	<	<	<	<	Ø
PO	>	>	>	>	>		Ø		≤		≤				<	<	Ø	Ø	<	Ø	Ø	Ø	Ø	Ø	<	Ø		Ø	Ø	Ø
MT	>	>	>	>		≥	≥	≥		=	Ø		<	<	<	<	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	<	Ø	<	<	Ø	Ø
V4t	>	>	Ø	>			≥		=						<?	<?									<			<		Ø
DP	Ø	Ø	Ø	Ø	>	>	>	≥	Ø				≤		≤	<?	Ø	Ø	<	Ø	Ø	Ø	Ø	Ø		Ø	<		Ø	Ø
VOT	Ø		>	Ø			>										<	<												
LIP	Ø		>	>	Ø	>	>		>		≥				≤		Ø	=	<		Ø	Ø		Ø	<	≤	<	≤	Ø	Ø
VIP	Ø		>	>					>						=	≤	Ø	Ø			Ø	Ø		Ø	<			≤		Ø
MSTd	Ø	>	>	>		>	Ø	>	>	>?	≥		≥	=			<?	<?	≤	<	Ø	Ø		Ø	<	<		≥		Ø
MSTl	>?	>	Ø	Ø	Ø	>	Ø	>	>	>?	>?	Ø		≥			Ø	Ø	Ø	<	Ø	Ø		Ø	<			≤		Ø
PITd	Ø	Ø	Ø	Ø	Ø	Ø	>	Ø	Ø			>	Ø	Ø	>?	Ø			Ø			<?		<?	<		<	≤		<?
PITv	Ø	Ø	Ø	Ø	Ø	Ø	>	Ø	Ø			>	=	Ø	>?	Ø			Ø		<?	<?		<?	<	<?	<	≤	<	<?
7a	Ø	Ø	Ø	Ø		Ø	Ø	>	Ø		>		>		≥		Ø	Ø			Ø	Ø			≤	<	≤?	Ø	<	<
STPp	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø				Ø	Ø	>	>					=	=	<			<		>	<	Ø
CITd	Ø	Ø	Ø	Ø	Ø	Ø	>	Ø	Ø		Ø		Ø	Ø	Ø	Ø		>	Ø	=				<?					<	<?
CITv	Ø	Ø	Ø	Ø	Ø	Ø	>	Ø	Ø		Ø		Ø	Ø	Ø	Ø	>	>	Ø	=				<?		<			<	<?
STPa	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø				Ø	Ø						>				Ø		<	<		<	
AITv	Ø	Ø	Ø	Ø	Ø	Ø	>	Ø	Ø				Ø	Ø	Ø	Ø	>?	>?			>?	>?	Ø			<	Ø	Ø	<	
FEF	Ø	>	>	>		>	>?	>	>	>			>	>	>	>			≥								≤			<
TF	Ø	Ø	>	>	Ø	Ø	>	Ø	Ø		Ø		≥	Ø	>			>	>	>		>	>	>			≤	≤		
46	Ø	Ø	Ø	Ø		Ø	>		>		>		>		Ø		>	>	≥?				>		≥	≥		Ø	≤	<
FST	Ø	>	>	>	Ø	>	>	Ø	>	>			≥	≥	≤	≥	≥	≥		<	Ø	Ø		Ø		≥	Ø			Ø
TH	Ø	Ø	Ø	Ø	Ø	Ø	>	Ø	Ø				Ø	Ø	Ø	Ø		>	>	>	>	>	>	>			≥			
AITd	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø				Ø	Ø	Ø	Ø	>?	>?	>		>?	>?			>		>	Ø		

Figure 2. Hierarchical constraints for the monkey visual system, slightly modified from Felleman & Van Essen (1991, table 7). Only clearly categorized constraints are shown. Thus, the absence of an entry does not necessarily mean that there is no anatomical projection between the respective areas. Nor does the existence of a hierarchical constraint always imply the existence of a connection, see §2(a). Different symbols stand for different classes of hierarchical constraints: ' < ' and ' ⩽ ' symbols stand for feed-forward and mixed feed-forward–lateral constraints, respectively; ' > ' and ' ⩾ ' stand for feedback and mixed feedback–lateral constraints; ' = ' represent lateral constraints; and '*Ø*' symbolize connections that have been looked for, but have been found absent. Question marks denote data that might not be fully reliable, according to Felleman & Van Essen (1991). We used a colour scheme to make the constraints more easily recognizable. Simple and mixed feed-forward constraints are shaded in blue, (mixed) feedback constraints in yellow and lateral constraints in green. Lighter hues have been used for the not fully reliable constraints, and grey entries stand for established absent connections. The figure is meant to be read as 'area in the column on the left is lower than/ lateral to/ higher than area in the top row'. The areas have been sorted as to arrange all feed-forward constraints above and all feedback constraints below the leading diagonal; however, some exceptions will be found for any order of the areas. This reflects the fact that there are inconsistencies within the hierarchical rules.

established experimentally. We found that, for values in a range above 135% for the necessary descendant cost, convergence to optimal solutions was very slow. If the requirements for the descendant cost could not be met, the algorithm tracked back to the start of step (ii) and introduced a different mutation.

(iii) *Validation*

A computed structure was accepted as new, and kept in the set of solutions, if (i) it possessed a cost equal to or lower than that of any solution found up to that point, and (ii) was different from all other stored solutions. Condition (ii) required that each potential solution structure was checked against all existing optimal solutions, and this set the computational limit for calculating large sets of solutions, which for practical reasons was around 150 000.

3. RESULTS

(a) *Primate visual system: solutions derived with the simplest cost function*

The processor found more than 150 000 different solutions using the standard data set and the simplest cost function. All the hierarchies had the same optimally small number of rule violations. The processor was able to find a full set of different solution structures in each of the computed epochs, giving reason to believe that the actual number of equally good optimal solutions is very much greater, probably by the order of several millions, than the 150 000 optimal hierarchies obtained.

areas	3a	3b	1	2	5	Ri	SII	7b	Ig	4	6	Id	SMA	35
3a				<			<							
3b			<	<			<							
1	>	>		<			<	<						
2	>?	>	>?		<		<	<		<			<	
5			>	<				<		<			<	
Ri					>		<	<	<					
SII	>	>	>	>		>		≤	<	<?	<?	<		
7b				<?	>	>			<	>?	<			
Ig						>	>	>						
4					<?						≤?		≤?	
6								>		≥?			=?	
Id							>		≥					<?
SMA					>?					≥?	=?			
35												>		

Figure 3. Hierarchical constraints for cortical somatosensory and motor areas of the macaque monkey. The constraints were created on the basis of table 8 from Felleman & Van Essen (1991) using the classification approach described in the text. The format of the entries is identical to the one used in figure 2.

Table 1. *The different data sets used in the hierarchical analysis of the primate visual system*

'standard'	set as used by Felleman & Van Essen (1991), excluding MIP and MDP; 30 areas, 318 constraints
'reliable'	standard excluding 'question mark constraints', excluding MIP and MDP; 30 areas, 282 constraints
'merged'	standard areas PITd, PITv, CITd, CITv, STPp and STPa combined into merged areas PIT, CIT and STP, excluding MIP and MDP; 27 areas, 286 constraints
'reliable merged'	standard with merged areas PIT, CIT and STP, excluding 'question mark constraints', excluding MIP and MDP; 27 areas, 262 constraints

All the computed hierarchies possessed a cost of six violated rules. In all cases, this involved two symmetrical pairs of three anatomical connection constraints that could not be satisfied for any particular hierarchy. This is a smaller number of violations than for any previous manually obtained solution. The familiar Felleman & Van Essen scheme of the visual cortical hierarchy (Felleman & Van Essen 1991), for instance, possesses eight constraint violations, when violations are counted in the same way as we have done throughout our analyses (counting subdivisions of areas independently). This cost is impressive given the informal methods used, but it is unlikely that the Felleman & Van Essen (1991) hierarchy is found in the top million hierarchies. The number of rule violations in the computed solutions was very small when compared to the cost for the random structures that give the starting point for the optimization, for which costs were around 170 violations, and very small by comparison to optimal solutions derived from randomly shuffled tables (see § 3(b)).

areas	18	17	PS	VLS	19	PMLS	SVA	21a	PLLS	ALLS	AMLS	20a	DLS	21b	20b	7	CGp	CGa	AES	ER	35	36
18		>		<	<	<			<		<	<		<		<						
17	>				<	<			>?		<	≤?		>								
PS									=?													<
VLS																						
19	>	>		>?		<?	<?	>?	<?	<?		<	<?	<	<	<?						
PMLS	>?	>	>?		=?			<	<?	<?	<	<							<		<	<
SVA			>?		=?										<?							<
21a	>?	>?			>?							<?				<?						
PLLS	>	>			<?	>?				<?		>?							<		<	<
ALLS	>?		>?		>?														<		<	<
AMLS	>	>	>?		<?	=	>			>?									<		<	<
20a		>?	>?	>?	>?	=?	>?	>?	>?					<?	<?	>?	<?	<?			<	<
DLS									>?										<			
21b	>?					=?						>?	=?									
20b			>?		>?		>?				>?	>					<?	<?			<	<
7	>?		>		>?			>?	>?	<?	>?	>?		>?	>							
CGp																				<	<	<
CGa			>?								>?				>?	>?					<	<
AES				>		>			>	>	>		>									<
ER																						
35																						
36																						

Figure 4. Hierarchical constraints for cortical visual and 'limbic' (cingulate and perirhinal) areas of the cat. The constraints are displayed in the same format as in figures 2 and 3. This figure shows hierarchically classified connections, rather than reciprocal hierarchical constraints, that were taken from a connectivity database published in conjunction with Scannell *et al.* (1995). The database is available online at http://www.psychology.ncl.ac.uk/jack/nature/scann95.txt.

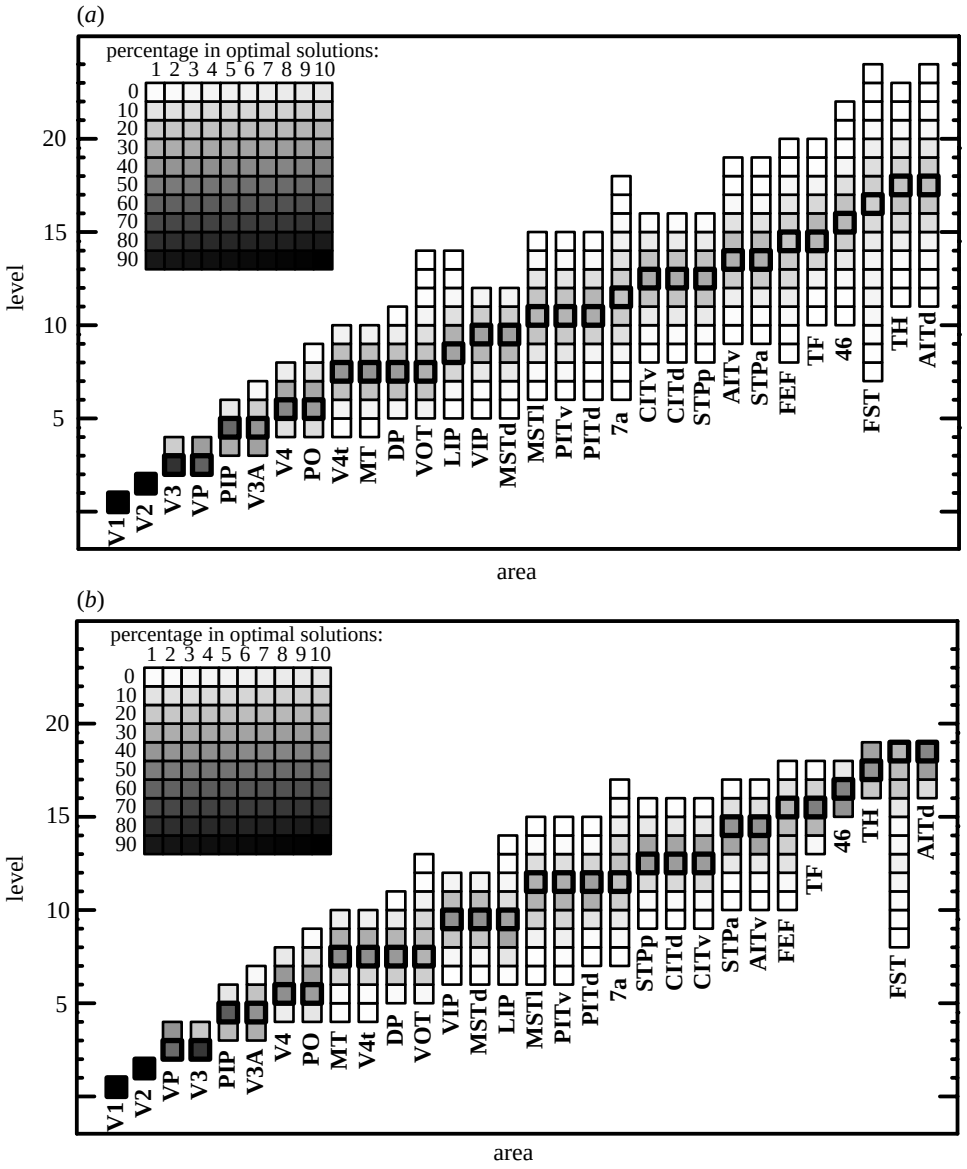

Figure 5. (*a*) Area frequency distribution of optimal visual hierarchies for the simple rule violation cost function. The boxes are shaded according to the relative occurrence of an area at a particular level in all 152 803 computed solutions. The main peaks of the area frequencies are denoted by frames in thicker lines, and the ordering of the peak solution also represents an optimal hierarchy. Violations of the hierarchical constraints are FST $\leqslant$ MSTd (relative occurrence in all solutions *ca.* 17%), FST < STPp (14%), LIP = PITv (12%), LIP $\leqslant$ MSTd (2%), FST $\geqslant$ TF (2%), MSTd < PITv (2%), FST $\geqslant$ PITd ($\ll$1%) and MSTd < PITd ($\ll$1%) (together with their corresponding symmetrical rules, e.g. MSTd $\geqslant$ FST). The first three violations, together with their counterparts, are the violations for the peak solution. (*b*) Area frequency distribution for all optimal visual hierarchies with 19 levels (40 131 cases), using the simple cost function.

The number of levels in the optimal hierarchies ranged between 13 (for 21 solutions) and 24 (for three solutions), with the peak of the distribution of levels being found around 18 levels (39 636 solutions) and 19 levels (40 131 solutions). Hence, the majority of optimal solutions possessed markedly more hierarchical stages than was apparent hitherto.

It is misleading to single out particular hierarchies from the large solution set, and impractical to show all of the optimal hierarchies. A statistical summary of the many solutions may therefore be the most appropriate format in which to represent the hierarchical structure of the primate visual system revealed by these data and rules. Figure 5*a* reflects the frequency with which one area appears on a particular hierarchical level, taking into account all 152 803 different solutions. The shading of the boxes in a column shows the proportion of solutions in which the area occupied the particular level. The sum of the frequency distribution for each column is 100%.

Figure 5*b* uses the same format as explained for figure 5*a*, and shows the distribution of areas in all optimal solutions that possessed 19 levels. Note that, despite the fact that all these optimal solutions shared the same total number of levels, there was still considerable variability in the placement of areas, especially for those that occupied intermediate levels of the optimal hierarchies.

The hierarchical constraints were sufficient to fix only V1 and V2 uniquely. These areas were always found on levels 1 and 2, respectively. The placement of all other areas depended on the structure of a particular solution. The positions of some areas were strongly associated with those of others, since these areas had an identical frequency distribution, indicating a fixed position relative to each other in all the solutions. These relationships concerned, besides V1 and V2, areas V4t and MT (same level), MSTd and VIP (same level), and CITv, CITd and STPp (same level). These areas could be collected into larger units, or building blocks, without restricting the variability of the solutions, see figure 6.

For each of the areas, the level with the highest occurrence among the solutions is marked in figure 5*a* (main peak level), and from the ordering of these levels a peak solution can be specified. This peak solution is shown in

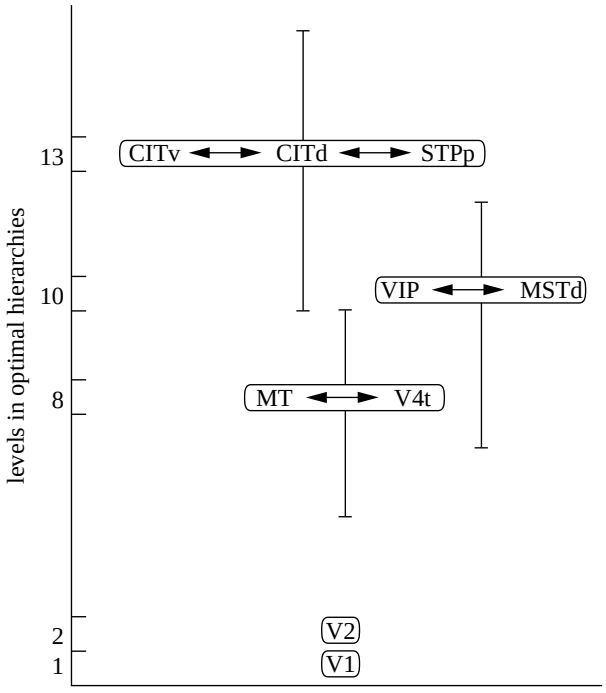

Figure 6. Fixed area relationships within optimal hierarchies. The figure represents the relationships between areas with identical level frequencies in the optimal arrangements. Three pairs of areas displayed fixed relative positions on their levels, while areas V1 and V2 had fixed absolute positions on levels 1 and 2, respectively.

figure 7, as are the connections of the areas. The ordering of areas in the peak solution is itself an optimal hierarchy. It should be noted, however, that even the peak hierarchy is just one optimal hierarchy from at least 150 000 that we found, and that it is noteworthy only because many other solutions have partial overlap with it. The ordering of the cortical areas in a particular optimal solution can, for this reason, be considerably different from the statistical main peak solution. This also explains the gaps apparent in the peak ordering in figure 5*a*, which are present even though no single solution possessed any gaps. A further example of the peculiarity of the single solutions is given by the low frequency of some of the rule violations. MSTd < PITd and FST ⩽ PITd, for example, occur in only 21 100 and 32 500 solutions, respectively.

The main peak level was the only peak in the distribution of almost all areas. An exception was apparent for area FST (fundus of superior temporal sulcus), which, beside the main peak at level 17 (for 29 550 solutions), had a second, lower, peak at level 12 (for 6779 solutions). We will examine the interesting behaviour of this area in a later section. The highest and lowest possible level for the areas typically possessed only a very low frequency among the optimal hierarchies.

(b) *Primate visual system: alternative input tables*

As well as the 'standard' rule table, we employed a number of different rule input tables (see table 1), while still using the simple cost function. We present only the resulting number and composition of rule violations of solutions arising from these tables.

In the 'reliable' table, all the data in the 'standard' table that were marked as unreliable by Felleman & Van Essen (1991) were omitted. The majority of rule violations for solutions from the 'standard' table concerned 'reliable' rule constraints, except for the infrequently occurring violations MSTd < PITv and MSTd < PITd. The cost for solutions arising from the 'reliable' table, however, was lower, having only two violations. These arose in every solution from the rule FST ⩽ TF and its counterpart. This means that removing the less 'reliable' rules allowed solutions better to satisfy the 'reliable' rules.

We investigated the violations arising if areas PIT, CIT and STP were assumed to be homogeneous, undivided areas by creating the 'merged' and 'reliable merged' tables. The 'standard' results described above indicated, however, that while this approach might be justified for CIT, whose two components appeared at the same level in all solutions, it might not be realistic to merge PITd and PITv, and STPp and STPa, as these areas had independent distributions in the solutions. For completeness, however, every rule involving either of these areas' subdivisions was reassigned to the relevant combined area, and for the 'reliable merged' table, all less 'reliable' rules were removed. Solutions computed from the 'merged' table possessed a cost of four rule violations, which in all cases were FST ⩽ TF, MSTd < PIT, and their complementary rules. The solutions arising from the 'reliable merged' table had a similar structure as those of the 'reliable' table. Again, FST ⩽ TF and its counterpart were the only rule violations, which yielded a minimal cost of two.

To compare the costs of the optimal solutions with non-hierarchical data, we examined a randomly composed rule table. Each hierarchical constraint listed in figure 2 was replaced by a constraint chosen at random from elsewhere in the set. The shuffled rule table was then optimized. The rule violation cost for solutions from rule tables processed in this way ranged between 117 and 134: much more than the six violations obtained for the optimized real anatomical data from the standard table. An interesting characteristic of the solutions from the shuffled rule tables was that only a small number of solutions with the minimal energy could be found for each shuffled rule table optimized.

We also computed hierarchies for data from table 5 in Felleman & Van Essen (1991), interpreted with the familiar classification scheme of figure 1. This yielded a slightly different set of 251 (including 82 less reliable) hierarchical constraints from the above table. This set also contained non-reciprocal relationships and constraints that were directly contradicting their complementary constraint. We used the simple cost function to compute 6828 optimal hierarchies for this table (within 1000 epochs, in each of which up to 30 solutions could be collected). All these solutions had 11 constraint violations, and possessed between 12 and 18 levels (with a peak of 4040 solutions at 15 levels). Violations in the optimal hierarchies from this rule table were LIP ⩽ PO (relative frequency in all optimal solutions 9%), MSTd = DP (9%), MSTd < PITv (9%), MSTd = 7a (9%), CITd < FEF (9%), CITv < FEF (9%), FEF > STPp (9%), TF < FST (9%), FST = TF (9%), MSTd < PITd (8%), 46 = TH (5%), TH > 46 (4%), FST ⩾ PITd (1%).

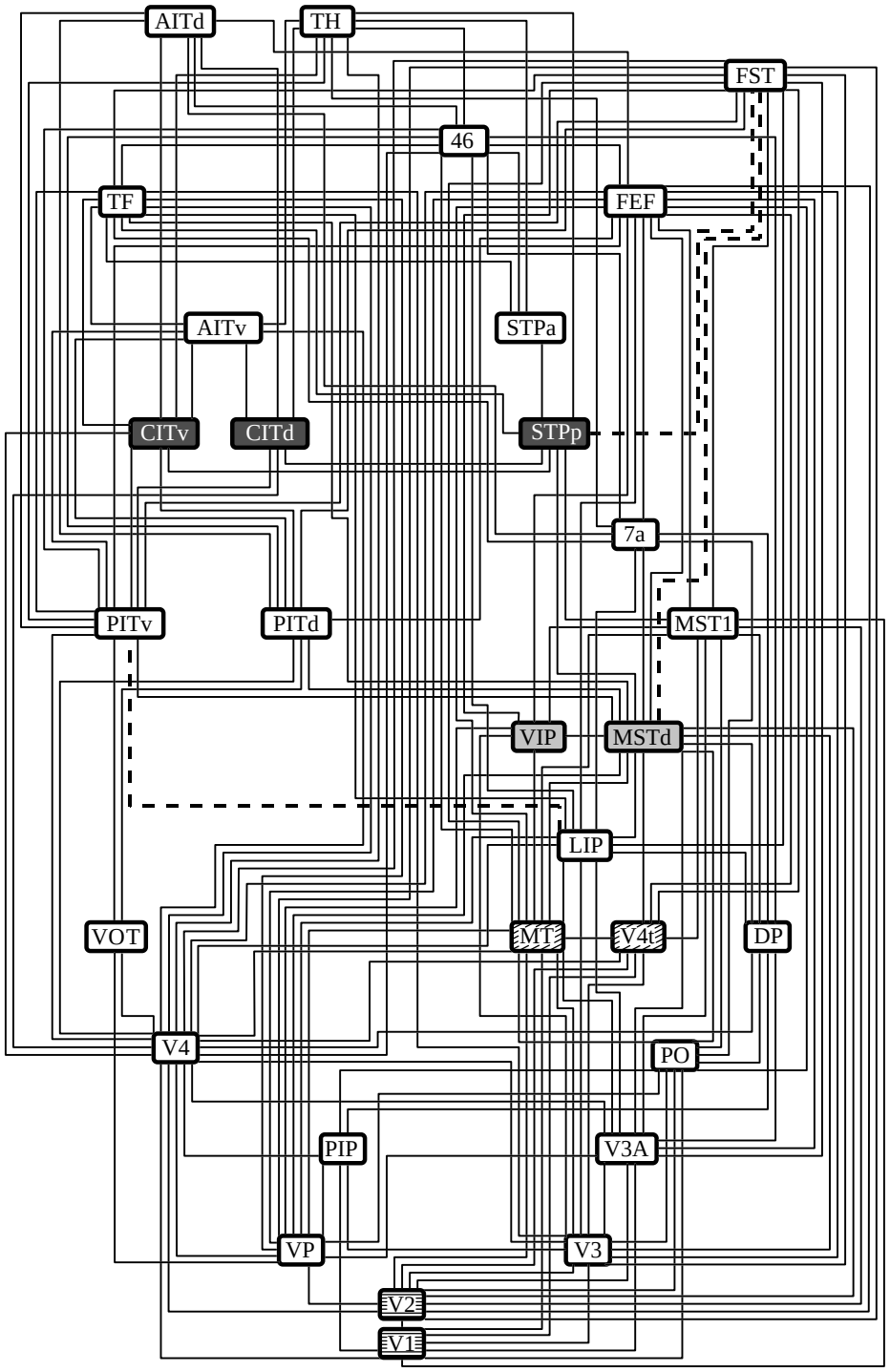

Figure 7. The figure shows the hierarchy derived from the peak positions of areas in all optimal solutions, as seen in figure 5*a*. Lines represent hierarchical relationships between areas. All lines between areas stand for reciprocal hierarchical constraints; and the three pairs of constraints that could not be fitted for this arrangement of areas, notably FST $\leqslant$ MSTd, LIP = PITv, FST < STPp, and their respective counterparts, are given in bold, dashed lines. Identically shaded areas have an identical distribution over all solutions, and so belong to the same building block (see § 3(a)). The figure also takes into account the organization of the visual system into at least two processing streams, guided by the multidimensional scaling representation from Young (1992). Thus, the left side of the diagram represents mainly occipito-temporal and the right side occipito-parietal components of the cortical visual system.

(c) *Primate visual system: a more complex cost function*

We examined optimal hierarchies for the visual system derived by optimizing with a more elaborate cost function, which we called the level-crossing cost function. This cost function evaluated each hierarchical structure according to the sum of two equally weighted parts: (i) the number of violations (the simple cost used above); and (ii) for each rule violation occurring, the number of levels separating the areas giving rise to the violation. To illustrate evaluation by the level-crossing cost function, we consider again the peak solution in figure 5*a* and figure 7.

The peak position in figure 5*a* for FST is on level 17 and that for MSTd on level 9. There are, however, a pair of hierarchical constraints requiring that the position of FST be on the same, or a lower, level than that of MSTd. Therefore, a level-crossing cost of eight has to be assigned to each of these violated rules. The total cost of the peak hierarchy in figure 5*a* according to the level-crossing cost function would thus be evaluated as (i) three violated rules and their counterparts, as before, giving a cost of six; plus (ii) 2×8 (FST $\leqslant$ MSTd), 2×2 (LIP = PITv) and 2×4 (FST < STPp). The peak hierarchy therefore yields a total cost of 34 according to the level-crossing function. The rationale for this modified cost function is

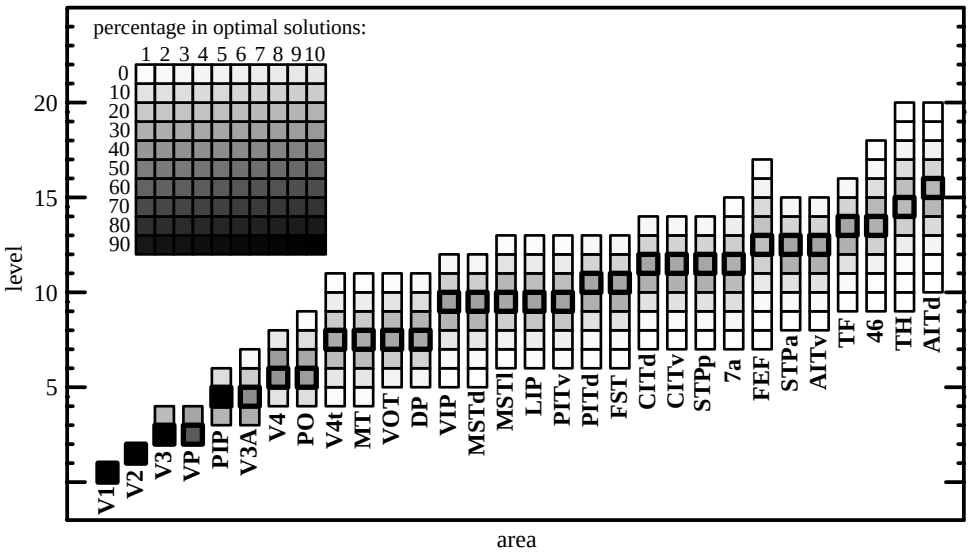

Figure 8. Area frequency distribution for the level-crossing rule violation cost. Shading of the boxes represents the relative occurrence of an area at a particular level in the 22 507 computed solutions. The main peaks of the area frequencies are denoted by frames in thicker lines, and the ordering of the peak solution also represents an optimal hierarchy under the level-crossing cost function. Violations of the hierarchical constraints are FST $\leqslant$ TF (relative occurrence in all solutions 17%), FST $\leqslant$ MSTd (15%), MSTd $<$ PITv (11%), MSTd $\leqslant$ LIP (6%) and MSTd $<$ PITd (2%) (together with their corresponding symmetrical rules, e.g. TF $\geqslant$ FST). The first three violations, together with their counterparts, are the violations for the peak solution, the misplacements of areas in these violations are three levels, one level and one level, respectively.

that if a rule has to be violated, then the departure from a perfect hierarchy should involve as few inappropriate levels as possible.

We computed 22 507 different solutions, which all possessed an optimal cost of 16 as evaluated by the level-crossing function. Again, computational limits rather than the availability of further optimal structures restricted the number of solutions we collected. All solutions had three rule violations and their correspondents, and they may be considered as a subset of the solution set for the simple cost function. The rule violations occurring for the level-crossing cost were also a subset of the ones found for the simple cost, although the violations appear with a different frequency among the solutions, and FST $<$ STPp, LIP $=$ PITv, FST $\leqslant$ PITd, and their counterparts were not contained in the set of rule violations derived using the level-crossing function.

The additional constraint imposed by costing the number of levels over which violations occur restricted the total number of levels in the hierarchies, which now ranged from 11 (for four solutions) to 20 levels (for 25 solutions) with the peak of the distribution around 16 levels (6670 solutions). A statistical representation of the solutions for level-crossing cost function is shown in figure 8.

The ordering of the peak solution again constituted an optimal hierarchical solution under this cost function. A larger number of areas lay in fixed positions relative to each other for this cost function, due to the restrictions the level-crossing cost placed on the variety of solution structures. V1 and V2, however, remained the only areas uniquely placed in all solutions at fixed levels in the hierarchies. Areas that possess the same relative position in all solutions were, beside V1 and V2, V4t and MT (on the same level), VIP and MSTd (same level), and LIP and PITv (same level). More complex fixed relationships were also apparent. These involved PITd and FST being fixed at the same relative level, while being locked with CITd, CITv and STPp at the next higher level; and STPa and AITv being fixed two levels higher than PITd, while TF was fixed three levels higher than PITd and one above STPa and AITv.

(d) *Primate somatosensory-motor system*

We investigated the hierarchical organization of the primate somatosensory-motor system using two different sets of hierarchical constraints. The first constraint set was identical to the 42 (2×21) pairwise hierarchical relationships suggested by Felleman & Van Essen (1991, table 8, column 11). The second set of data was derived from individually classifying the laminar patterns presented in table 8 of Felleman & Van Essen (1991), using the categorization shown in figure 1. This second approach yielded a total of 59 hierarchical constraints, including non-reciprocal rules and constraints that directly contradicted their reciprocal counterparts. Cases where this classification approach failed, however, were connections that originated bilaminarly from the supra- and infra-granular layers and whose termination pattern was unknown, or connections with contradictory termination patterns.

Figure 3 shows the hierarchical constraints obtained through the individual classification approach, and figure 9 summarizes all 144 optimal arrangements that were computed with the simple cost function on the basis of these data. There were diminishing returns in deriving further optimal hierarchies with further epochs of analysis, suggesting that there is a much smaller number of optimal hierarchies for this system by comparison to the macaque visual system. The hierarchies had between

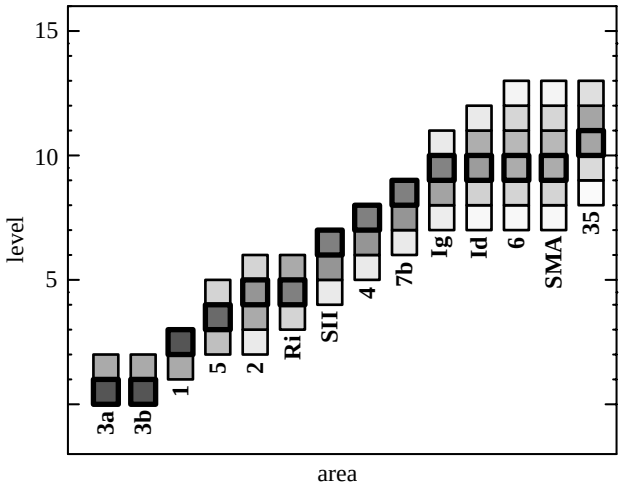

Figure 9. Statistical summary of 144 optimal hierarchies for the primate somatosensory-motor system. Fixed relative positions in these hierarchies are 3a and 3b on the same level, SII, 4 and 7b on three consecutive levels, and 6 and SMA on the same level.

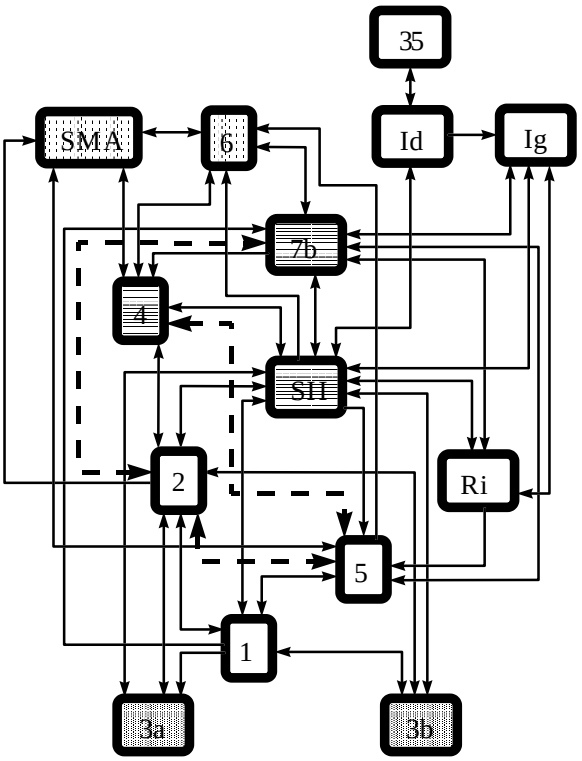

Figure 10. Optimal peak hierarchy for the primate somatosensory-motor system. Constraint violations for this optimal scheme are 7b < 2, 4 < 5 and 2 < 5. Areas marked by identical hatching have fixed relative positions in all optimal solutions, see figure 9.

nine (for two solutions) and 13 (for 24 solutions) levels, with the peak at 12 levels (56 hierarchies). All optimal arrangements had three violations of the hierarchical constraints. The constraint violations were 7b < 2 (relative frequency 33%), 4 < 5 (33%), 2 < 5 (25%), and 5 < 2 (8%), with the first three of these also being the constraint violations for the optimal peak hierarchy shown in figure 10.

In contrast, the original Felleman & Van Essen (1991) interpretation of the data yielded a much larger number of solutions, all without constraint violations. We computed 8974 such solutions, but the computational ease with which they could be obtained suggests that there may be many more of them. These solutions had between six (for 681 solutions) and ten (for 120 solutions) levels, with the peak at seven levels (4074 solutions). Consequently, the individual classification allowed a more specific determination of the ordering of the somatosensory-motor system, but at the cost of a larger number of constraint violations. The structure of the constraint violations for arrangements from individually classified constraints, however, suggests that all of them could be easily resolved. All of the violated rules, apart from the constraint 2 < 5, are based on less compelling, retrograde tracing data. The rules 7 > 2 and 4 < 5 should, moreover, be considered unreliable (Felleman & Van Essen 1991). Discarding all violated rules, except the constraint 2 < 5, would therefore abolish the constraint violations for the individually classified data set.

(e) *Cat visuo-limbic system*

Hierarchical arrangements of visual areas have also been suggested for the cat: two recently published schemes can be found in Felleman & Van Essen (1991) and Scannell *et al*. (1995). Whereas the former scheme arranged 16 cortical visual areas on eight hierarchical levels, the latter one included 22 visual and limbic areas, stretched over 14 levels, and gave rise to 15 violations of the known hierarchical constraints. Both schemes placed areas 17 and 18 on the first and second hierarchical level, respectively.

We reanalysed the data set shown in figure 4, which is identical to the one used by Scannell *et al*. (1995), with the network processor using the simple cost function. As for the primate visual system, there was a large number of optimal hierarchies, of which we computed more than 25 000 (25 631), a limit imposed solely by practicality. All optimal solutions created 15 constraint violations and possessed between ten and 20 hierarchical levels. The distribution of areas over hierarchical levels in all optimal solutions is shown in figure 11.

None of the areas possessed a unique level position in all optimal hierarchies, however, areas DLS and 21b had fixed relative positions throughout all solutions, on the same level. Area ALLS had a two-peak distribution, in about 18% of the solutions being on level 7, and in another 16% on level 12. As we argue in a similar case for primate area FST (see § 3(f)), this raises the possibility that area ALLS may be composed of two distinct areas, which differ in their anatomical connectivity. However, as the state of knowledge of cat connectivity is more uncertain than that of the primate visual system (e.g. the classification scheme fits poorly and there are a relatively large number of constraint violations), we did not pursue this peculiarity through further analyses.

Despite the higher number of constraint violations in the cat visual system, the system displays a significant degree of hierarchical organization. Optimized solutions from randomly reshuffled constraint tables yielded much higher numbers of constraint violations than the optimal

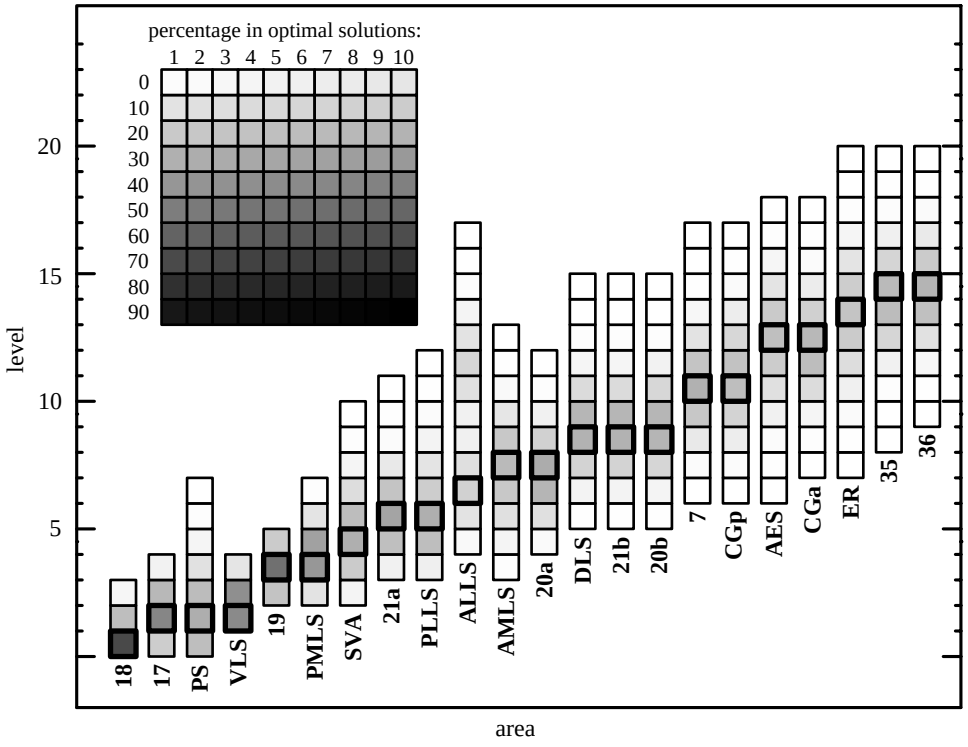

Figure 11. Distribution of cat cortical visual and limbic areas in 25 100 optimal solutions computed with the simple cost function. Constraint violations are 19 > 21a (with a relative frequency of about 7% in all optimal solutions), 17 > PLLS (7%), 17 > 21b (7%), PLLS < 19 (7%), AMLS < 19 (7%), 21b = PMLS (7%), 20a = PMLS (7%), 20a > 7 (7%), AMLS = PMLS (7%), PS = PLLS (6%), PLLS > 20a (6%), 187 > 17 (5%), SVA = 19 (5%), 19 < PMLS (4%), AMLS > ALLS (3%), 7 < ALLS (3%), PMLS = 19 (3%), 19 < SVA (2%), 17 > 18 (1%), 20a > PLLS (1%), PMLS > PS (< 1%), PMLS < AMLS (< 1%).

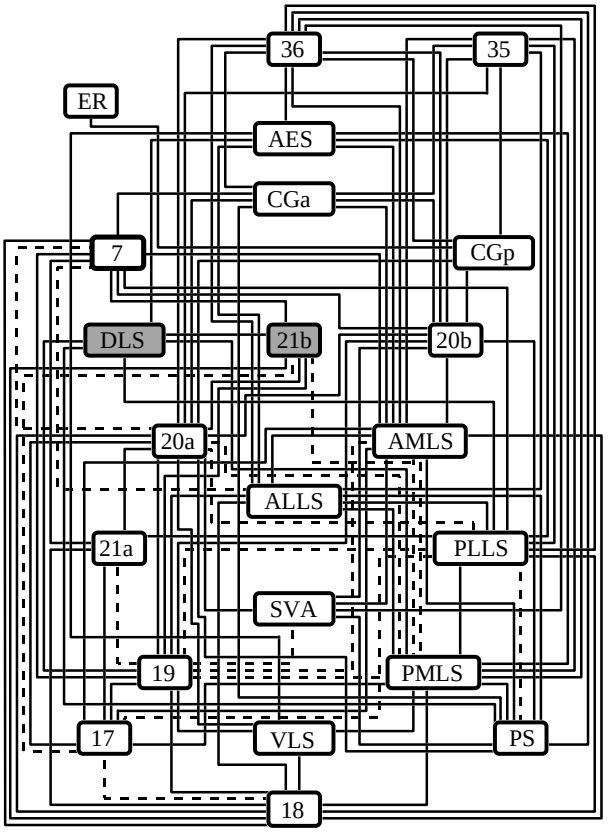

Figure 12. Optimal peak hierarchy for cat visuo-limbic areas. Identical hatching of areas DLS and 21b indicates fixed relative positions of these areas in all optimal hierarchies, see figure 11.

arrangements of the anatomical data (between 65 and 80 violations for ten randomized tables).

The peak hierarchy (figure 12) differs in a number of features from the unique hierarchy presented by Scannell *et al*. (1995). While the number of constraint violations of both hierarchies is 15, and the numbers of levels are similar (13 levels in the peak hierarchy versus 14 levels of Scannell *et al*.'s scheme), the hierarchical order of some of the stages varies. This concerns most prominently the order of areas 17 and 18, which is reversed in the two schemes. The placement of 18 at the very bottom of the optimal peak hierarchy follows from the fact that area 18 was placed on the lowest possible level in 71% of all optimal hierarchies, and on the second and third level only 25% and 3%, respectively. Interestingly, the order of early visual stages 18, 17, 19 and PMLS in the optimal peak hierarchy closely matches the sequence of area activations found in a latency study by Dinse & Krüger (1994).

(f) *Computer experiments on primate visual area FST*

As mentioned earlier, area FST showed a number of characteristics that clearly distinguished it from the other areas in the solutions. FST was involved in the most commonly occurring rule violations under both cost functions and all the different input tables. The majority of rule violations in the simple and level-crossing cost solutions were related to FST, and the only remaining rule violation for solutions obtained with the 'reliable' and 'merged reliable' input tables also concerned FST. Other areas were obviously also involved in these rule violations, but FST was the one with the highest frequency, and under level-crossing costing, the rule violation with the

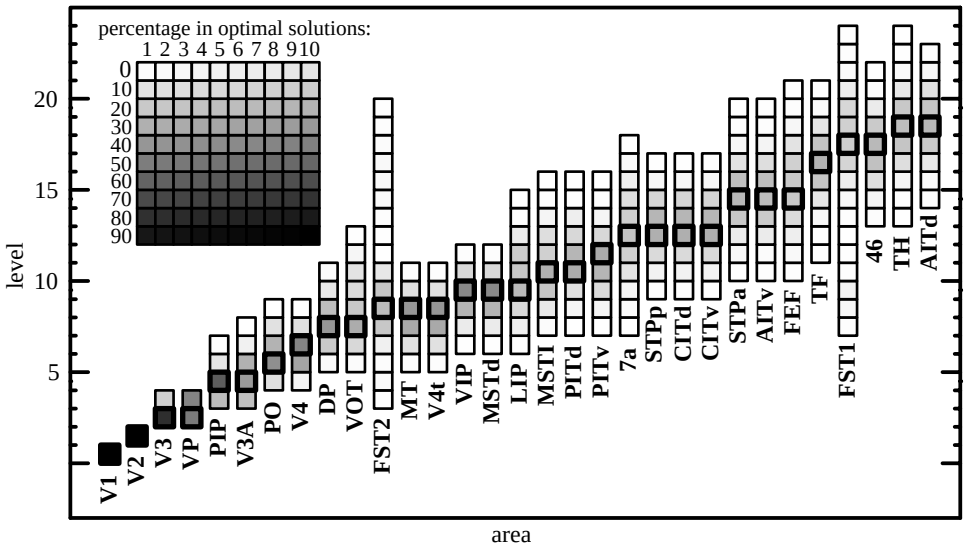

Figure 13. Area frequency distribution for FST split input tables and the simple cost function. Note the widely separated main peak levels of FST_1 and FST_2. The multipeak distribution of the FST subdivisions is due to the fact that their area distributions are created from solutions of different input tables, of which only some specify multipeak distributions for the FST subcomponents. Rule violations for these optimal solutions were LIP $\leqslant$ MSTd or LIP = PITv or MSTd < PITv, and their respective corresponding rules, of which MSTd < PITv and its correspondent is an unreliable constraint, see figure 2.

most costly misplacement (FST $\leqslant$ TF, three levels departure from fitting the constraint for the main peak solution) also involved FST.

FST, moreover, was the only area that possessed a two-peak distribution of levels in the statistics for the simple cost solutions (see figure 5*a*). The lower peak at level 12 had a relative frequency among the solutions of 4%, compared to a frequency of about 19% for the main peak at level 17, indicating a split of the solutions into two subsets, in which FST appeared at considerably different levels in the hierarchy. The solutions in the two subsets had the same optimal cost of six rule violations, but differed from each other in respect of the placement of area FST. Solutions that had FST on a lower hierarchical level were favoured by the level-crossing cost function, so that a single peak for FST appeared at level 11 for this set of solutions (see figure 8). This caused the main peak position of FST to change considerably from the simple cost to the level-crossing cost solutions, unlike any other area.

We investigated the causes for FST's behaviour by conducting further computer experiments. Two principal types of errors, or a combination thereof, could exist in the experimental data used for the previous analyses. The first error could be that anatomical connections to and from FST could have been assigned to the wrong hierarchical category. For example, ascending or lateral connections could have been misclassified as descending ones. However, none of the 15 connections FST makes with other areas has been marked as unreliable in the database, and the primary data are notably clear on the laminar patterns of origin and termination for these projections (Boussaoud *et al.* 1990). We are not, therefore, convinced that misclassification of laminar direction is a likely explanation.

The second possible error could be mispartitioning of area FST. FST as presently partitioned could include or exclude territory and connections belonging in reality to other, structurally and functionally independent areas. Alternatively, FST could itself contain more than one area, and the subdivisions might possess different connectivity. The latter possibility, a subdivision of a homologue of area FST into a dorsal and a ventral part, has been reported to hold in the owl monkey (Kaas & Morel 1993). For this reason, and because of the indications given above, we explored the assumption that FST is actually composed of two areas, which differ in the nature of their connections. Appropriate modelling of this difference should then produce hierarchical structures with fewer rule violations.

We considered FST as two areas, FST_1 and FST_2, which both obey the existing hierarchical constraints for FST, but which possess distinct subsets of the connections to other areas. That is, either FST_1 or FST_2 makes the connection to any of the 15 possible target areas, to which FST as a whole is connected. The number of combinations of the 15 connections of FST among FST_1 and FST_2 when overlapping sets of connections are allowed can be calculated as $\frac{1}{2}(3^n - 1)$, which for $n = 15$ yields 7 174 453 combinations. This is an impracticably large set of possibilities to compute over. Hence, we first restricted our modelling to non-overlapping cases, in which the sets of connections ascribed to the two supposed subcomponents of FST were exclusive. This constituted a simplification, since we would certainly expect some overlap in the connections made by the two areas (see below). Even without accounting for such possible overlap, the number of combinations, which here can be calculated according to $\frac{1}{2}2^n - 1$, was 16 383. We wrote a computer program to prepare the constraint tables for all these possible distributions of the 15 connections between FST_1 and FST_2. The resulting input data table exceeded 16 Mb. Only combinations, however, that involved one or more of the constraints in rule violations (see § 3(a); figure 5) were analysed, a restriction that reduced the input table by a further 1023 cases.

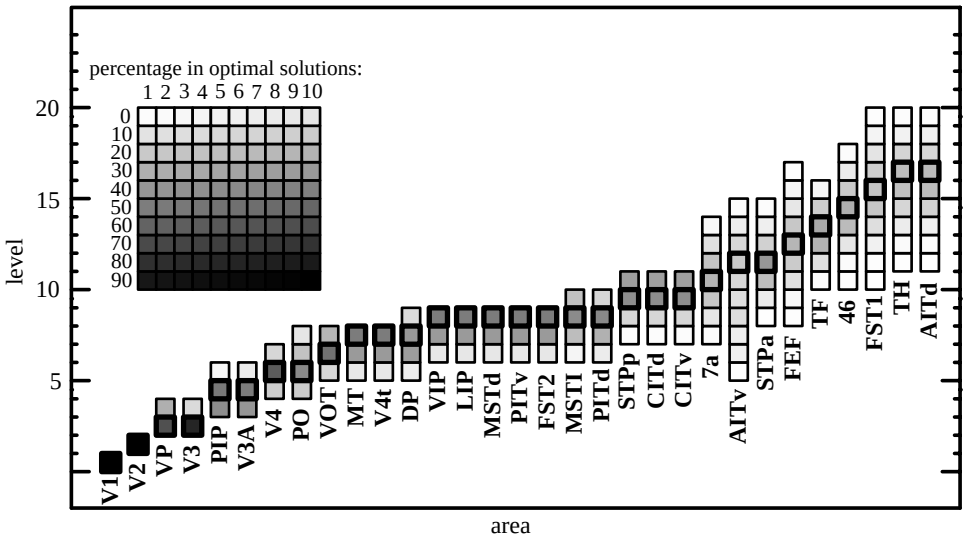

Figure 14. Area frequency distribution for 2500 perfect hierarchical solutions, without any constraint violations. These solutions have been obtained under the assumption that (i) FST is split into two distinctive subcomponents FST_1 and FST_2, and (ii) the 36 less reliable hierarchical constraints be excluded. Note that the constraint set from which these solutions arose is only one out of a large number of such sets which all allow perfect hierarchies, and that the assumed subdivisions of FST and their connections to other cortical visual areas will have to be established experimentally. The peak levels of FST_1 and FST_2 were on levels 16 and 9, respectively, whereas the peak levels for these areas across all the perfect hierarchies from different constraint sets were on levels 18 and 9, respectively.

The data sets from this input table were optimized by a modified version of the network processor, which was able to cross-compare solutions arising from different input tables. This version of the program also possessed the ability to store and process data in a compressed format, and to adapt runs automatically to the available computer, and could thereby handle the large amounts of data involved in the experiments. The results were that 137 constraint sets led to hierarchies with a new optimal cost of only two constraint violations. Neither of the two violations involved FST. Hence, this manipulation abolished FST violations completely. All the constraint sets leading to only two rule violations had in common that area PITd was exclusively connected to one subdivision of FST, whereas areas MSTd and STPp were exclusively connected to the other. All other connections could be made by either or both FST_1 and FST_2, accommodating the probable overlap in their connectivity.

The distribution of area levels in the optimal solutions is given in figure 13, and shows that the hypothetical subcomponents of FST, FST_1 and FST_2, have their main distribution peaks on hierarchical levels 9 and 18, respectively: that is, they are strikingly separated in the orderings.

To verify that this separation was not due to an artefact of our approach, we performed exactly the same experiment for a different area, PITd. This area also appeared in rule violations and shared a 'building block' with FST (that is, it had an identical distribution of the hierarchical levels as FST) in the level-crossing solutions. In the same fashion as described for FST, we redistributed the nine connections PITd makes with other areas between two hypothetical subcomponents of PITd in all possible combinations. In the solutions derived from all these rule input tables, however, the PITd divisions remained adjacent to each other, with the distribution peak on the same level.

We noticed that one of the pairs of violations that remained for optimal hierarchies with FST subdivisions, namely MSTd < PITv and its correspondent, concerned a hierarchical constraint whose reliability may be doubtful (Felleman & Van Essen 1991). We conducted further computer experiments and found that splitting FST into FST_1 and FST_2 with different connections, coupled with the exclusion of the 36 potentially less reliable constraints (or just MSTd < PITv and its counterpart) permitted solutions that violated none of the constraints. These two assumptions allowed hierarchies that fitted the pairwise relationships perfectly (see figure 14 for a summary of 2500 perfect hierarchies).

We found that, for a special case, we could also model all combinations of rule distributions between FST_1 and FST_2, including possible overlap in their connections to other areas. Under the assumption that the connectivity structure of FST in macaques is similar to that found in owl monkeys (Kaas & Morel 1993), it was possible always to assign analogous connections to FST_1, FST_2 or both, so that only redistributions of the connections for which there was no guidance from the owl money data had to be considered. FST_1 was given fixed connections to assumed macaque homologues MT, MSTd, MSTl, V3 and V3A (assuming that the latter two areas are equivalent to area DM of the owl monkey; see Kaas & Morel 1993), as well as V2. FST_2 was firmly assigned connections to MT and V2, while not having the other connections of FST_1. This reduced the number of combinations for the other nine FST connections at most to 9841. Optimal hierarchies derived from these input tables had a minimal energy of four rule violations. This cost was higher than for the non-overlapping approach, and this suggests that the connectivity structure of FST subdivisions in macaque monkeys (if such a split exists) is likely to be somewhat different to that in owl monkeys.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

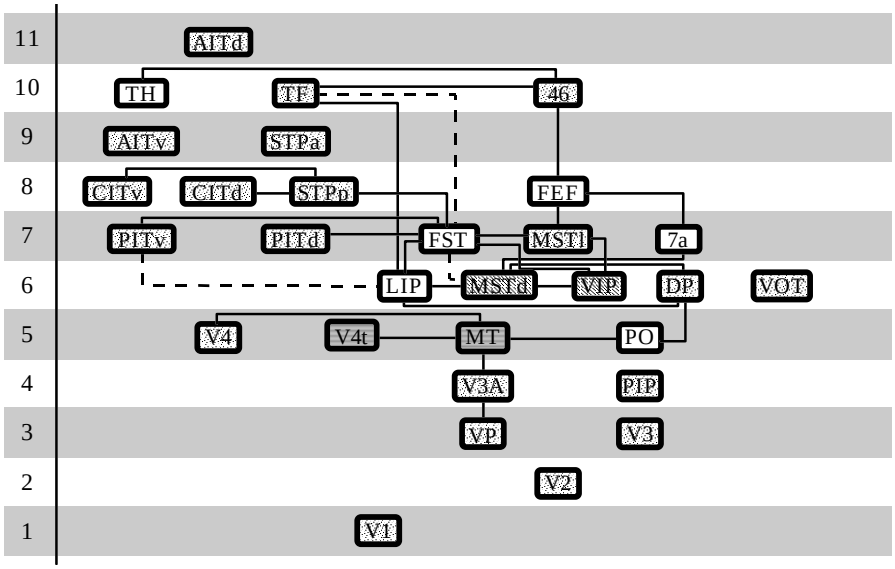

Figure 15. Optimal hierarchical enumeration of primate visual areas, in which each area is placed at the lowest level possible (adapting a constraint suggested by Van Essen & Felleman (1996) and Crick & Koch (1998)). In this optimal 11-level scheme, every area is placed on the lowest level of the area's distribution in all optimal 11-level hierarchies, see figure 16. To verify that no optimal scheme with lower positions for the areas existed, we attempted to compute optimal hierarchies with ten levels, but found none in extensive computations. Most connections between the areas are omitted for simplicity's sake; however, connections with lateral contributions in their connectivity patterns are shown. This points out a network that may implement a fast signalling system within the primate visual system, see §4(e). The constraint violations for this optimal arrangement are LIP = PITv, FST ⩽ MSTd, FST ⩾ TF, and their respective counterparts. Dashed lines denote these violations in the scheme. Areas marked with identical hatching have fixed relative, or absolute, positions in all optimal 11-level hierarchies computed, see figure 16. The scheme is arranged so as to represent the functional subdivisions of the primate visual system, with the left side representing the ventral and the right side the dorsal stream.

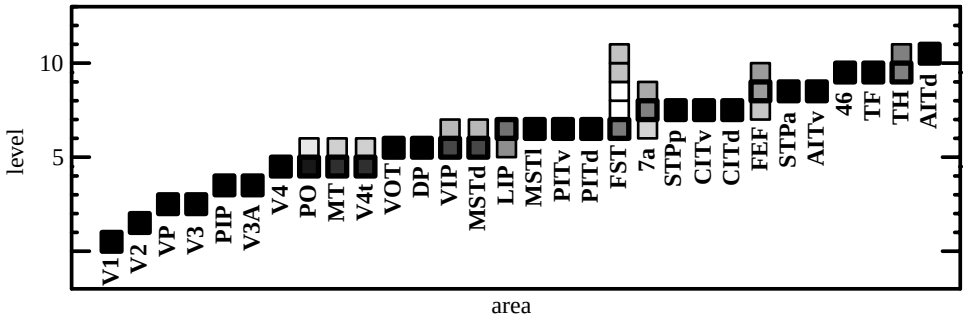

Figure 16. Distribution of primate visual areas in 108 optimal hierarchies with 11 levels. Constraint violations are: FST ⩽ MSTd (relative frequency in all solutions: 12%), FST ⩾ TF (9%), FST STPp (8%), LIP = PITv (7%), LIP ⩽ MSTd (5%), MSTd < PITd (5%), MSTd < PITv (5%), and their complementary rules. MT and V4t, as well as VIP and MSTd, have identical distributions in all optimal 11-level hierarchies, although their positions are not unique.

4. DISCUSSION

(a) *Hierarchical organization of sensory cortices*

Our results support, through computational analysis, the view that areas in the sensory cortices of cat and macaque monkey are organized hierarchically. The degree to which this principle is reflected in the structure of corticocortical connections, however, differs from species to species and between brain regions. The results show that the organization of the cortical visual system of the macaque monkey is remarkably strictly hierarchical. This is apparent in the marked difference between the small number of rule violations the optimal arrangements possessed and the number of violations for non-hierarchical, randomly created, rule sets. The finding of a strongly hierarchical organization of cortical systems should not be confused, however, with a unique hierarchical arrangement (see §4(b)). Neither should it be confused with a strictly serial hierarchy, in which only one linear pathway of information flow exists. There are very many opportunities for level-skipping and 'backward' interactions, and many different interacting paths could be traced through the hierarchy (see De Yoe & Van Essen 1995; Felleman & Van Essen 1991; Van Essen *et al.* 1992). Neither do the results imply that the physiological properties of the cortical areas should be strictly hierarchical. Both the 'backward' projections and the many interactions that the cortical areas have with subcortical structures serve to complicate the relationship

between the structural aspects of the system drawn out here and its functional properties (Young 1995). Moreover, this study ignored the sophisticated intrinsic, microanatomical connectivity, which may vary between different cortical areas.

(b) *Resolving hierarchical indeterminacy*

Our results show that no unique solution can be given for the primate cortical visual system with existing rules and laminar data. Instead, a very large number (more than 150 000) of hierarchical structures exist that all have the same low cost. No single hierarchical solution can therefore accurately represent the hierarchical structure of the system, and conclusions drawn from any single hierarchy will not be reliable.

One might expect that the indeterminacy of the visual hierarchy could be resolved through a number of advances, such as the addition of more experimental data on cortical connections or the inclusion of other hierarchical constraints, for instance from subcortical–cortical connections. There might also be alternative ways for defining and applying hierarchical constraints that do not give rise to a multitude of equally good arrangements. We now examine some methods that have been suggested for resolving hierarchical indeterminacy.

We addressed the possibility that adding further experimental data for the primate visual system could resolve hierarchical indeterminacy. We considered all possible ways of filling in the 235 potential hierarchical constraints (in figure 2) whose direction is presently unknown. We made the assumptions that (i) all unknown connections exist and can be hierarchically classified on the basis of their laminar patterns; and (ii) these artificial data should not create additional constraint violations in the optimal hierarchical arrangements. We found multiple solutions in all cases, even for constraint sets without any undetermined connections. The main reason for this result is that about half the presently known constraints have been looked for and found absent (317 out of 635 suggested hierarchical or established absent constraints in figure 2). The sparsity of connectivity, particularly between parietal and inferotemporal areas, leaves many degrees of freedom for arrangements of the areas that fit the constraints equally well.

Van Essen & Felleman (1996) and Crick & Koch (1998) have suggested that the multitude of equally plausible hierarchical schemes we find could be restricted to just one by placing each area on the lowest level possible, given the anatomical constraints for this area. The method can produce a unique hierarchical enumeration for a set of data, but may not do so when the system possesses hierarchical inconsistencies. Only exhaustive search would determine that the method produces a unique ordering for a particular data set. More importantly, it invokes a constraint that does not come from the neuroanatomy and is difficult to justify and to interpret biologically. Any attempt to limit hierarchical schemes to the smallest possible number of levels arbitrarily but indirectly gives greater weight to the lateral components of those connections that have been classified as lateral–feed-forward or lateral–feedback. We know of no suggested justification for this at present. Notwithstanding this reservation, other researchers may find this additional constraint justified, and so we show in figure 15 an optimal hierarchy for the primate visual system, in which each area is placed at the lowest level possible. However, restricting the total number of levels does not itself lead to an unambiguous hierarchy: we computed hierarchies of the primate visual system restricted to the lowest number of levels that was obtained for optimal hierarchies computed with any of the cost functions (i.e. 11 levels). With this approach, we gained another 108 optimal hierarchies (i.e. with six constraint violations each) which each had 11 levels. A statistical summary of these solutions in the familiar format is given in figure 16.

Recently, Crick & Koch (1998) envisaged that including thalamocortical connections as hierarchical constraints might allow the construction of an unambiguous visual hierarchy. They pursued a classification of corticothalamic connections into two main classes, E- (extended) and R- (round) axons (from Rockland 1996), and suggested that the R-axons would form 'driving' connections, similar to feed-forward connections within the cortex. With this assumption, and their hypothesis that strong feed-forward loops should be avoided (Crick & Koch 1998), they predicted 'that if a certain thalamic region, x, projects strongly to the middle layers of a certain cortical area, Z, then any projection x receives from R axons must come from a cortical area Y, which is lower in the cortical hierarchy than Z'. Scannell *et al.* (1999) argued, on the basis of cat connectivity, that such additional constraints are very unlikely to occur, as the same thalamic region generally does not link cortical areas that are not themselves connected at the cortical level. Even were such cases to arise, they still might not disambiguate cortical visual hierarchies. If the two unconnected areas Z and Y were in different cortical visual streams, they could possess corticothalamic projections of the kind described by Crick & Koch (1998) without creating closed strong loops, and the relative hierarchical position of these areas would thus remain unspecified. However, the additional constraints suggested by Crick & Koch (1998) may help to disambiguate relationships of the mixed feed-forward–lateral or feedback–lateral type.

A further reason for the large number of equally valid solutions is that hierarchical analysis employs an integer cost function (number of violations), and the anatomical rules are relational (e.g. higher than) and contain no information on the hierarchical 'distance' between areas (e.g. much-higher-than versus not-much-higher-than). Advances on this aspect of the problem could involve incorporating into the hierarchical cost the density or laminar spread of projections, or correlating hierarchical position with other anatomical and morphological parameters. Amir *et al.* (1993) and Elston & Rosa (1997, 1998), for instance, related morphological parameters such as dendritic arborization or the spread of intrinsic connections to the 'height' of an area in the cortical hierarchy. There is a tendency for neighbouring areas in the cortex to exchange denser connections and connections that involve more layers (Rockland 1997). It might be possible to use this information to restrict the hierarchical positions of some areas, by introducing additional terms to the cost function which calculate the cost related to the level separation of areas. The goal of deriving rational

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

cost functions to test concepts of cortical organization is thus intimately linked with the need to make progress in the experimental quantitation of anatomical parameters, such as connectivity strength.

(c) *Constraint violations*

The optimal solutions the processor obtained all had equal or lower costs than those previously published by other authors. For the hierarchical constraints of the cat visuo-limbic system provided by Scannell *et al.* (1995), our optimal arrangements possessed the same number of 15 constraint violations as Scannell *et al.*'s scheme, and fewer violations than the hierarchy suggested by Felleman & Van Essen (1991), which created 23 constraint violations. Our hierarchies of primate somatosensory-motor areas created no constraint violations for the set of pairwise hierarchical relationships (Felleman & Van Essen 1991) and three constraint violations if the constraints were derived from individually classified connections. Felleman & Van Essen's (1991) arrangement of these areas had no violations either for the pairwise rules, but created four violations for the individual constraints.

The unique primate visual hierarchy presented by Felleman & Van Essen (1991) and Van Essen *et al.* (1992) possessed a cost of eight under the simple cost function, whereas the 150 000 solutions computed here had a simple rule violation cost of six. For the level-crossing cost function, the Felleman & Van Essen (1991) solution possesses a cost of 22, compared to 16 for the 22 500 optimal solutions computed here. The structure of solutions (in particular that of the optimal peak solutions) and their rule violations also differed from those of the hierarchies proposed hitherto. We observed that the number of levels in the majority of solutions for the primate visual system (peak of distribution around 18 and 19 levels) was significantly larger than that of previous models.

Despite the large number of optimal solutions we obtained computationally, only a limited set of rule violations occurred. This concerned the rules (and their counterparts) FST $\leqslant$ MSTd, LIP = PITv, FST < STPp, LIP $\geqslant$ MSTd and FST $\leqslant$ PITd, as well as the rules FST $\leqslant$ TF, MSTd < PITv and MSTd < PITd. The latter three violations were also found by Felleman & Van Essen (1991). The rule violations involved seven cortical areas, namely FST, MSTd, LIP, PITv, PITd, STPp and TF. If hierarchy is an important anatomical principle, a thorough anatomical analysis might revisit these areas to determine the causes for the occurrence of rule violations in the optimal solutions. On the other hand, non-hierarchical interactions could be an essential part of the organization of cortical visual areas. Felleman & Van Essen (1991) also obtained one further rule violation, 46 $\leqslant$ AITd, which did not appear for any of our optimal solutions, in which AITd was always and unequivocally higher than area 46.

The level distribution of areas in the optimal solutions suggested that of the areas where there is some explicit doubt about internal differences (CITd–CITv, PITd–PITv, STPp–STPa, AITd–AITv), only CIT could be considered as a homogeneous unit. All the other pairs have two different level distributions; that is, the areas in these pairs meet partly different rule constraints in the optimal solutions. Interestingly, the combined area CIT could also possess a special relationship to STPp, which has the same level distribution as both CITd and CITv.

The results of our computer experiments in which all FST connections were reassigned to two hypothetical subdivisions, FST_1 and FST_2, showed the possible existence of such a split in macaques. We created rule tables that led to solution structures with very low cost, which did not involve connections of FST_1 or FST_2. For these solutions, the main peaks of FST_1 and FST_2 in the level distribution appeared on greatly different levels. Anatomical experiments will be required to show whether such subdivisions of FST, as already reported for owl monkeys (Kaas & Morel 1993), actually exist in macaques. Results from our experiments that were based on the assumption of a close relationship between owl and macaque monkeys suggested, however, that the pattern of connections of FST_1 and FST_2 in macaques may be somewhat different from that for FSTv and FSTd in owl monkeys. Moreover, our results did not indicate a split of area MT into subdivisions as found in the owl monkey (Kaas & Morel 1993)—or such a potential split concerns only the connections between areas MT and FST.

(d) *Hierarchy and other aspects of cortical organization*

We have suggested that the indeterminacy of the visual hierarchy is mainly due to the limited cross-talk between two distinguishable streams of the visual system. An alternative to outlining a global cortical hierarchy, which incorporates all visual cortical regions, would therefore be to segregate the system into separate hierarchical arrangements for the identified connectional subdivisions (Hilgetag *et al.* 1998; Young 1992). Figures 7 and 15 suggest such an arrangement. However, these separate hierarchies might not be unambiguous due to the nature of their constituent hierarchical constraints, a proportion of which are of the mixed (lateral–feed-forward or lateral–feedback) type. This aspect is exemplified by the occipito-parietal stream of the primate visual system, which comprises a particularly large number of such constraints: 14% (17 out of 122) of the connections within the dorsal stream are of the mixed type, compared to only 2% (one out of 45) within the ventral stream, following the dorsal–ventral categorization from figure 15. Mixed constraints allow more degrees of freedom for the relative placement of areas. The unequal distribution of mixed constraints thus indicates that the hierarchical organization of the dorsal stream is more flexible or uncertain than that of the ventral stream. We speculate below that this difference may have important consequences for information processing in the visual system. Additionally, the greater hierarchical flexibility of connections between dorsal stream structures may be mirrored by a more uniform intrinsic organization of dorsal stream areas, compared to a more hierarchically differentiated intrinsic anatomy of ventral stream areas. Elston & Rosa (1997), for example, found that morphological parameters thought to be correlated with the hierarchical position (e.g. size and complexity of basal dentritic fields, number of dentritic spines) are very similar for dorsal stream areas MT, LIP and 7a, in contrast to ventral stream areas V4 and TEO, which are

clearly separated according to their intrinsic parameters (Elston & Rosa 1998).

A further functionally relevant aspect of hierarchical structure was pointed out by Rockland (1997). She noted that, in a small but potentially important number of cases, projections to two different cortical targets can arise from the same projection neurons, through collaterals. This implies that the same functional information is relayed to the two targets. The phenomenon occurs for terminations both within and between areas. In the first case, axons that terminate in one layer might have collaterals terminating in another. In the second situation, secondary terminations of a projection may connect in a completely different area. An example of such across-area collaterals would be connections from V2 to V4 that sprout collateral projections to area TEO, thereby creating a direct V2–TEO route (Nakamura *et al*. 1993; more examples are reviewed in Rockland (1997)). Although the question of cortical hierarchy itself remains untouched by this aspect of connectivity, a better understanding of the extent and functional significance of the phenomenon is desirable and will depend on the systematic collation and analysis of data from further multilabelling studies.

We presently lack a clear-cut hierarchical classification scheme for all connections. When confronted by connections that are poorly fitted by present classification schemes, we wonder whether a better, quantitatively substantiated scheme for hierarchically classifying cortical connections could be derived. The problem may be particularly relevant for defining hierarchical categories of corticocortical connections in the cat brain. As Scannell *et al*. (1999) pointed out, the direct transfer of the macaque classification scheme is inappropriate given the consistently different laminar structure of connections in the cat. The problem of classification of projections can ultimately be approached best by systematic collation of connectivity data including laminar information, which should be specified in quantitative terms (i.e. counts of retrogradely labelled cells or quantitative measures for anterograde label distributions) wherever possible. Statistical procedures such as cluster analyses could then be used to group the different connection patterns, to establish whether there are distinct classes of laminar projections.

On the other hand, quantitative analyses of frontal and prefrontal connectivity in the macaque carried out by Barbas (1986) and Barbas & Rempel-Clower (1997) suggest that laminar connectivity patterns in the cortex may not in fact belong to distinct classes, but rather follow a steady gradient of change. These researchers found that the structure of the cortical layers themselves (i.e. the definition and number of layers) was a good predictor for the laminar patterns of originating and terminating connections, and consequently for a hierarchical arrangement of the areas. Connections would preferentially originate in the lower layers (5–6) of 'lower' areas (with a poor definition of layers) and terminate in the upper layers (1–3) of 'higher' areas (showing well-defined layer structure), and vice versa for connections originating from 'higher' areas terminating in 'lower' ones. A structural gradient for the layer definition could also be claimed along occipitotemporal and occipitoparietal axes in the primate visual cortex (Barbas & Rempel-Clower 1997), with V1 showing the clearest definition and largest number of identifiable layers. The classification scheme depicted in figure 1 would then simply be a consequence of this structural gradient. Some of the laminar patterns do indeed seem to follow the predictions from such a structural model. For instance, projections from V1 (the 'highest level' area according to its layer structure) originate from the upper layers and terminate in the lower layers (if counted as layers 4–6) of its cortical targets. A more systematic experimental analysis of layer structures and corresponding laminar patterns could clarify a structural model for primate sensory cortices.

Some indication about regional variations of laminar patterns could, however, already be gained from an analysis using a modified version of our network processor. The goal of this analysis would be to delineate cortical regions on the basis of the laminar patterns of connections within these regions. The modified integer cost function for this approach would be composed of two components. The first component would measure the number of connections (or any other defined neighbourhood relationship) between the potentially separated regions, each composed of any number of individual cortical areas. The second cost component would measure the number of differences in the connectivity patterns within the regions. Minimizing this combined cost with the optimization algorithm would yield groups of areas that are strongly interconnected by similar types of connections. If, however, a structural model applies to the primate visual cortex, and laminar patterns vary along a smooth gradient, no distinct regions should obtain.

(e) *Anatomical hierarchy versus neurophysiological anti-hierarchy*

Recent reviews (e.g. Nowak & Bullier 1997) and systematic studies of processing latencies in several visual areas (e.g. Schmolesky *et al*. 1998; Schroeder *et al*. 1998) have presented results that appear to contradict a hierarchical organization of the visual system. The responses to externally presented visual stimuli were much more variable than seems consistent with a strictly hierarchical organization of pathways within the system. In particular, the temporal overlap in the activation of dorsal stream neurons in response to flashed stimuli was so great that there were no significant differences in latency among any of the dorsal stream structures tested (Schmolesky *et al*. 1998). In contrast, there were significant latency differences between occipitotemporal areas V1, V2 and V4. Moreover, all dorsal activations occurred with much shorter latencies than ventral activations (Nowak & Bullier 1997; Schmolesky *et al*. 1998).

How can these results be reconciled with a hierarchical organization of cortical anatomy? We speculate that an important part of the answer lies in the unequal distribution in the different streams of connections involving 'lateral' laminar patterns. Connections with lateral components terminate in the same layers from which onward projections originate. Connections with lateral components could thus 'bridge' areas with fast, monosynaptic contacts, whereas strictly feed-forward connections have to cross at least one other intrinsic synapse before contacting cells in the output layers. We noticed that the

number of lateral or mixed lateral-feed-forward–feedback, connections is much greater in the dorsal stream than it is in the ventral: 23 out of 122 (19%) connections in the dorsal, but only three out of 45 (7%) connections in the ventral stream. This organization would give rise to very rapid onset times in dorsal areas compared to those in the ventral structures, and to great overlap in activation in the dorsal pathways. Hence, this anatomical organization could meet the observation in Nowak & Bullier (1997) that 'the major factor that influences the speed of activation of cortical areas appears to be whether or not a cortical area belongs to the dorsal stream'.

We further suggest that the dorsal stream may have both fast and slow signalling pathways. The fast pathways would be implemented by monosynaptic contacts in lateral or mixed inter-areal projections, and would allow rapid but simple computations without lateral inhibition. The slow pathways, implemented by feed-forward projections, would engage the intrinsic processing machinery in each area, including inhibitory interactions, and could implement more complex but slower computations. The ventral stream, by contrast, may possess only or mainly a feed-forward sequence. This would involve extra delays as the signal passes through the intrinsic connectivity of the areas and lead to the observed separation and relatively delayed onset times in the ventral areas. Some connections between the streams would also allow monosynaptic contacts on output neurons in the target (potentially 11 out of 70 connections = 16%) and could play a role in temporal coordination of the streams. In addition, we note that callosal connections terminate prominently in output layers (Clarke 1994) and that these connections may be biased toward compartments with richer magnocellular input (Clarke 1994; Olavarria & Abel 1996). These interhemispheric connections might therefore preferentially contact the fast signalling pathways in the two dorsal streams, so promoting temporal integration across the mid-line. Hence, hierarchical anatomical organization and anti-hierarchical neurophysiological processing need not be contradictory, but rather, may suggest previously unsuspected modes of functioning.

Supported by The Royal Society, the Medical Research Council, the Wellcome Trust and the University of Newcastle upon Tyne. We are very grateful to the many colleagues who provided feedback at the various stages of this project.

REFERENCES

Abramowitz, M. & Stegun, I. A. (eds) 1972 *Handbook of mathematical functions*. New York: Dover.

Amir, Y., Harel, M. & Malach, R. 1993 Cortical hierarchy reflected in the organization of intrinsic connections in macaque monkey visual cortex. *J. Comp. Neurol.* **334**, 19–46.

Barbas, H. 1986 Pattern in the laminar origin of corticocortical connections. *J. Comp. Neurol.* **252**, 415–422.

Barbas, H. & Rempel-Clower, N. 1997 Cortical structure predicts the pattern of corticocortical connections. *Cerebr. Cortex* **7**, 635–646.

Boussaoud, D., Ungerleider, L. G. & Desimone, R. 1990 Pathways for motion analysis—cortical connections of the medial superior temporal and fundus of the superior temporal visual areas in the macaque. *J. Comp. Neurol.* **296**, 462–495.

Burkhalter, A. & Van Essen, D. 1983 The connections of the ventral posterior posterior area (VP) in the macaque. *Soc. Neurosci. Abstr.* **9**, 46.5.

Clarke, S. 1994 Modular organization of human extrastriate visual cortex: evidence from cytochrome oxidase pattern in normal and macular degeneration cases. *Eur. J. Neurosci.* **6**, 725–736.

Coogan, T. A. & Burkhalter, A. 1993 Hierarchical organization of areas in rat visual cortex. *J. Neurosci.* **13**, 3749–3772.

Crick, F. & Koch, C. 1998 Constraints on cortical and thalamic projections: the no-strong-loops hypothesis. *Nature* **391**, 28–36.

Cusick, C. G., Seltzer, B., Cola, M. & Griggs, E. 1995 Chemoarchitectonics and corticocortical terminations within the superior temporal sulcus of the rhesus monkey: evidence for subdivisions of superior polysensory cortex. *J. Comp. Neurol.* **360**, 515–535.

De Yoe, E. A. & Van Essen, D. C. 1995 Concurrent processing in the primate visual cortex. In *The cognitive neurosciences* (ed. M. S. Gazzaniga), pp. 383–400. Cambridge, MA: MIT Press.

Dinse, H. R. & Krüger, K. 1994 The timing of processing along the visual pathway in the cat. *NeuroReport* **5**, 893–897.

Elston, G. N. & Rosa, M. G. 1997 The occipitoparietal pathway of the macaque monkey: comparison of pyramidal cell morphology in layer III of functionally related cortical visual areas. *Cerebr. Cortex* **7**, 432–452.

Elston, G. N. & Rosa, M. G. 1998 Morphological variation of layer III pyramidal neurones in the occipitotemporal pathway of the macaque monkey visual cortex. *Cerebr. Cortex* **8**, 278–294.

Fausett, L. V. 1994 *Fundamentals of neural networks: architecture, algorithms, and applications*. Englewood Cliffs, NJ: Prentice Hall.

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

Hilgetag, C.-C., O'Neill, M. A. & Young, M. P. 1996 Indeterminate organization of the visual system. *Science* **271**, 776–777.

Hilgetag, C.-C., Burns, G. A. P. C., O'Neill, M. A. & Young, M. P. 1998 Cluster structure of cortical systems in mammalian brains. In *Computational neuroscience: trends in research* (ed. J. M. Bower), pp. 41–46. New York: Plenum Press.

Kaas, J. H. & Morel, A. 1993 Connections of visual areas of the upper temporal lobe of owl monkeys—the MT crescent and dorsal and ventral subdivisions of FST. *J. Neurosci.* **13**, 534–546.

Laarhoven, P. J. M. V. & Aarts, E. H. L. 1987 *Simulated annealing: theory and applications*. Dordrecht, The Netherlands: Kluwer.

Maunsell, J. H. & Van Essen, D. C. 1983 The connections of the middle temporal visual area (MT) and their relationship to a cortical hierarchy in the macaque monkey. *J. Neurosci.* **3**, 2563–2586.

Nakamura, H., Gattass, R., Desimone, R. & Ungerleider, L. G. 1993 The modular organization of projections from areas V1 and V2 to areas V4 and TO in macaques. *J. Neurosci.* **13**, 3681–3691.

Nowak, L. G. & Bullier, J. 1997 The timing of information transfer in the visual system. In *Extrastriate cortex in primates*, vol. 12 (ed. K. S. Rockland, J. H. Kaas & A. Peters), pp. 205–241. New York: Plenum Press.

Olavarria, J. F. & Abel, P. L. 1996 The distribution of callosal connections correlates with the pattern of cytochrome oxidase stripes in visual area V2 of macaque monkeys. *Cerebr. Cortex* **6**, 631–639.

Rockland, K. 1996 Two types of corticopulvinar terminations: round (type 2) and elongate (type 1). *J. Comp. Neurol.* **368**, 57–87.

Rockland, K. 1997 Elements of cortical architecture: hierarchy revisited. In *Extrastriate cortex in primates*, vol. 12 (ed. K. S. Rockland, J. H. Kaas & A. Peters), pp. 243–293. New York: Plenum Press.

Rockland, K. S. & Pandya, D. N. 1979 Laminar origins and terminations of cortical connections of the occipital lobe in the rhesus monkey. *Brain Res.* **179**, 3–20.

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

Scannell, J. W., Burns, G. A. P. C., Hilgetag, C.-C., O'Neill, M. A. & Young, M. P. 1999 The connectional organization of the cortico-thalamic system of the cat. *Cerebr. Cortex*. **9**, 277–299.

Schmolesky, M. T., Wang, Y., Hanes, D. P., Thompson, K. G., Leutgeb, S., Schall, J. D. & Leventhal, A. G. 1998 Signal timing across the macaque visual system. *J. Neurophysiol*. **79**, 3272–3278.

Schroeder, C. E., Mehta, A. D. & Givre, S. J. 1998 A spatiotemporal profile of visual system activation revealed by current source density analysis in the awake macaque. *Cerebr. Cortex* **8**, 575–592.

Van Essen, D. C. & Felleman, D. J. 1996 Indeterminate organization of the visual system—response. *Science* **271**, 777.

Van Essen, D. C., Anderson, C. H. & Felleman, D. J. 1992 Information processing in the primate visual system: an integrated systems perspective. *Science* **255**, 419–423.

Webster, M. J., Bachevalier, J. & Ungerleider, L. G. 1994 Connections of inferior temporal areas TEO and TE with parietal and frontal cortex in macaque monkeys. *Cerebr. Cortex* **4**, 470–483.

Young, M. P. 1992 Objective analysis of the topological organization of the primate cortical visual system. *Nature* **358**, 152–155.

Young, M. P. 1995 Open questions about the neural mechanisms of visual pattern recognition. In *The cognitive neurosciences* (ed. M. S. Gazzaniga), pp. 463–474. Cambridge, MA: MIT Press.

Young, M. P., Scannell, J. W., O'Neill, M. A., Hilgetag, C.-C., Burns, G. & Blakemore, C. 1995 Non-metric multidimensional scaling in the analysis of neuroanatomical connection data and the organization of the primate cortical visual system. *Phil. Trans. R. Soc. Lond.* B **348**, 281–308.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES