## CONCLUSION

The development of the human nervous system is a complex and protracted phenomenon, involving many processes spanning prenatal as well as postnatal development.

At birth, while the generation of neurons is by no means finished, the number and locations of neurons have essentially been established, and their differentiation to adult form is complete or well under way. However, the development of important structures still has a long way to go. Establishment of the basic structure of major brain areas such as the neocortex, hippocampus, and cerebellum will continue for many months postnatally. Major axon bundles are still forming at birth, including the corpus callosum which will bridge the cerebral hemispheres, and the corticospinal tract with its motor projections to the spinal cord. The development of the myelin sheaths that speed the conduction of action potentials along axons will proceed throughout adolescence, and the major changes in connectivity allowed by synaptic remodeling through synapse elimination will take years to complete.

This development involves both progressive and regressive processes that are influenced by multiple interacting factors. The dynamic assembly and destruction of the nervous system during normal development ultimately requires both genetic as well as epigenetic sources of information, including the behavioral experience of the organism.

### Further Reading

Brown MC, Hopkins WG and Keynes RJ (1991) *Essentials of Neural Development*. Cambridge, UK: Cambridge University Press.

Cowan WM (1979) The development of the brain. *Scientific American* **241** (September): 112–133.

Jacobsen M (1991) *Developmental Neurobiology*. New York, NY: Plenum.

Purves D and Lichtman JW (1985) *Principles of Neural Development*. Sunderland, MA: Sinauer.

Sanes DH, Reh TA and Harris WA (2000) *Development of the Nervous System*. San Diego, CA: Academic Press.

# Neural Development, Models of

Introductory article

*Geoffrey J Goodhill*, Georgetown University Medical Center, Washington, DC, USA

**CONTENTS**



*Models of neural development attempt to produce precise mathematical and computational descriptions of the biological mechanisms that regulate the formation of the nervous system.*

## INTRODUCTION

Neural development is the process whereby the nervous system of a developing animal is formed. It involves several stages, each of which is regulated by different types of mechanisms. At each stage a complex interaction of genetic and environmental influences determines the final outcome. The effect of the environment becomes especially important in later stages of development, so that the mature nervous system is specifically adapted to the environment in which it must survive. Mathematical/computational models (hereafter referred to as 'models') have been formulated for several stages of development. As with all modeling, the goal is to take qualitative hypotheses about possible mechanisms and make them quantitatively precise. This rigorously defines the conditions under which such qualitative hypotheses are indeed plausible candidate mechanisms, and makes explicit the implicit assumptions that often underlie them. A desirable result of such modeling is that it makes novel and testable predictions about the outcome of experiments.

Neural development can be conceptually divided into four stages, although they overlap temporally to some degree. These stages are neural

induction and neural-tube formation, pattern formation and regionalization, establishment of connectivity, and activity-dependent refinement of connections. The majority of modeling research has concentrated on the last of these stages, attempting to understand the principles by which connection strengths between neurons are adjusted in response to specific patterns of neural activity from the sense organs. This article will consider these stages sequentially.

## NEURAL INDUCTION AND NEURAL-TUBE FORMATION

The first stage of neural development is the specification of neural ectoderm by the notochord, which occurs within the first few weeks of gestation in humans. This region, which is called the neural plate, then folds in on itself to form the neural tube. The neural tube gives rise to the peripheral and central nervous systems. As the neural folds come together to form the neural tube, neural-crest cells migrate from the neural folds away from the neural tube. The neural crest forms the peripheral nervous system, while the rest of the neural tube forms the central nervous system. The molecular interactions that lead to the specification of one region of the gastrula as the neural plate have been extensively investigated experimentally. Diffusion of certain molecules from the organizer region is crucial, and mathematical modeling of diffusion processes in general is very well understood. However, application in this area is complicated by the fact that these molecules may be transported by mechanisms other than free diffusion. Furthermore, as yet there are virtually no models of the mechanical forces that drive the formation of the neural tube.

## PATTERN FORMATION

In the second stage of neural development, the neural tube expands and changes shape as new cells are generated and migrate to new locations. In conjunction with this, certain regions start to become specialized for particular fates. Moving from anterior to posterior, boundaries begin to form between the prosencephalon (embryonic forebrain), mesencephalon (embryonic midbrain) and rhombencephalon (embryonic hindbrain). Further subdivisions then occur as the prosencephalon divides into the telencephalon and diencephalon, and the rhombencephalon divides into the metencephalon and myelencephalon. Within these general regions discrete structures are constructed, such as the hippocampus and cerebrum from the telencephalon, the retina and thalamus from the diencephalon, and the cerebellum and pons from the metencephalon. The main question that modeling can attempt to address at this stage of development is how such a spatially non-uniform structure can emerge from an apparently spatially uniform tube.

### Reaction–Diffusion Systems

One way to approach this question is to treat it as a very general one. Pattern formation occurs in many other biological systems, such as the spatially regular pattern in which leaves are added to a plant stem, and the division of a fly embryo into individual body segments. In fact, evolutionary conservation of genetic programs has led to some of the same genes controlling regionalization of the nervous system as control fly body segment formation. Pattern formation has also been studied in many physical systems, such as convection patterns in heated liquids and certain types of chemical reactions. The most influential class of mathematical models in this area has been that of reaction–diffusion systems, first proposed by Turing in 1952 and subsequently developed in specific biological contexts by Meinhardt. The principal assumptions are that cells in an initially spatially uniform array produce several chemicals or 'morphogens' that diffuse at different rates from cell to cell through the array. In its simplest form there is one 'activator' chemical and one 'inhibitor', often called an activator–inhibitor system. The activator diffuses slowly and stimulates its own production, while the inhibitor diffuses more rapidly and suppresses production of the activator. In any real system some noise will be present, which means that each cell in the array will not initially have exactly the same concentrations of morphogen. The activator–inhibitor system amplifies such small fluctuations dramatically, leading to a stable state of regularly spaced regions dominated by either the activator or the inhibitor. This elegant mathematical idea has been developed in numerous ways to model biological pattern formation. One aspect of neural development where such algorithms may operate is the selection of just one cell in a proneural cluster to become a neuroblast. The ligand Delta, operating through its receptor Notch, acts as a lateral inhibitor, so that if one cell is specified as neural, then neighboring cells are inhibited from assuming that fate.

## Morphogen Gradients

However, there is no evidence that reaction–diffusion mechanisms underlie the formation of overall regionalization in the embryonic nervous system. Rather than asking how symmetry is broken in a spatially uniform array, the key question seems to be how patterns arise from an initial gradient of some morphogen over the array. This initial polarization is created by the positioning of the egg inside the mother. Wolpert first framed this question in 1969 as the 'French flag' problem, and subsequently discussed how various types of chemical reaction kinetics could transform an initially smoothly varying morphogen concentration into sharply defined regions of different concentrations. The fundamental process by which the nervous system becomes initially regionalized seems to be a biochemical cascade, whereby initial gradients of maternal gene products turn on genes in a concentration-dependent manner, creating gradients of new proteins and so on, to form structure at ever finer levels of detail. One of the first signs of patterning along the anterior–posterior axis of the embryo is the expression of homeobox genes of the Pax family in different regions. These gene products then control the regional expression of other genes. As yet there has been no theoretical research addressing the structure and outcome of these complex networks of interacting proteins in the context of nervous system regionalization. However, some models have recently been proposed for fly segment formation, which may turn out to be relevant to the nervous system of vertebrates as well. In such models, certain gene products are assumed to regulate the expression of other genes in a network, somewhat akin to an artificial neural network with weighted connections. Gene products also diffuse within the embryo. Starting initially with a network of random weights (strength of regulatory influences), the outcome of the developmental process for this network is calculated. Using methods such as gradient descent and simulated annealing, weights are then incrementally updated until the developmental outcome matches that which is seen biologically. These optimal weights represent quantitative predictions of the strengths of the regulatory influences that exist in the real biological system, and predictions can be made about the effect of specific gene deletions.

# ESTABLISHMENT OF CONNECTIVITY

The task that is achieved in the third stage of neural development is the establishment of initial patterns of connectivity between neurons. Structures such as the retina send out axons which navigate through a complex environment to find appropriate targets (e.g. the superior colliculus and lateral geniculate nucleus). On their journey, they often have to make many guidance decisions, where for instance some axons in a group are routed in one direction and other axons are routed in another direction (e.g. the split between axons originating from the nasal and temporal halves of the retina that occurs at the optic chiasm). On reaching the correct target region, both the tip of the axon and the target have to change their structure in order to form a specialized synapse. The principal way in which the growing axon senses its environment is via the growth cone (a specialized structure at its tip). The growth cone continuously extends and retracts thin filopodia which sample the local region. The types of molecular guidance cues that can be detected include pathways of growth-permissive factors (e.g. laminin), barriers of growth-inhibitory factors (e.g. myelin-associated glycoprotein) and concentration gradients of molecules (e.g. netrins and semaphorins) which can be attractive or repulsive. These cues can be bound to or expressed by the substrate, or they can diffuse more or less freely. This type of diffusion process (e.g. from a target region) is one way in which a concentration gradient can be established.

## Models of Growth-Cone Guidance

Models have been proposed that address several different types of questions at this stage of development. One class considers the types of gradient shapes that axons might encounter *in vivo* or *in vitro*, and how this shape constrains the maximum distance over which it is possible for an axon to be guided. Such models assume that gradient detection can only occur when two principal constraints are satisfied. First, the absolute concentration of the guidance factor at the growth cone must exceed a certain minimum, and secondly, the fractional change in concentration across the growth cone must exceed a certain steepness. Such models predict that the maximum distance over which a growth cone could be guided *in vivo* is about 1 mm for a target-derived diffusible factor, and about 1 cm for a substrate-bound gradient with shape optimized for maximum guidance length. Both of these are reasonably consistent with experimentally measured values. Another class of models attempts to understand how growth-cone gradient detection is limited by noise. The external ligand gradient causes different numbers of receptors to

be bound on one side of the growth cone than the other. However, since receptor binding is a stochastic process, at each instant a small difference from the mean level of binding will be measured. One influential hypothesis has been that the threshold for reliable gradient detection occurs when the average size of these small differences becomes comparable with the true difference in concentration between the two sides of the growth cone. Such models can produce numerical values for growth cones of about 1%, which is similar to experimentally measured values. Another class of models focuses on the role of filopodia in growth-cone guidance. Such models propose rules (usually stochastic) with regard to the way in which filopodia are generated in response to environmental cues. Complete axonal trajectories can then be simulated, and both short-term and filopodial dynamics and overall trajectories can then be compared with reality. The mechanisms whereby a difference in receptor binding is internally amplified to give a turning response toward or away from the direction of higher concentration have not yet been considered for growth cones. This is partly because the signal transduction pathways in the growth cone that cause this response are only how beginning to be revealed experimentally. However, several models which may be applicable have been proposed for analogous signaling pathways in bacteria, white blood cells and slime molds. One hypothesis is that the polarization of a cell in response to the gradient may occur by a similar type of reaction–diffusion mechanism previously described in the context of spatial regionalization.

### Retinotectal Map Formation

Perhaps the most popular area for theoretical modeling at this stage of development is the formation of topographic maps. There are numerous examples in the brain, such as the retina and superior colliculus (optic tectum), where the connections between two structures form a continuous map, so that neighboring neurons in one structure connect to neighboring neurons in the target structure. Although, as will be discussed later, neural activity plays an important role in fine-tuning the structure of such maps, their initial formation is probably governed by gradient-directed axon-guidance mechanisms. The basic idea, first proposed by Sperry several decades ago, is that gradients of some molecules in the input structure give each neuron in that structure a unique regional identity. Analogous gradients are postulated to exist in the target structure, and the map forms because each axon follows the gradients in the target structure to a topographically appropriate location. During the past few years, direct evidence for this type of process has been discovered in the form of gradients of Eph receptors in structures such as the retina, and their ligands (the ephrins) in target structures such as the optic tectum. Starting in the 1970s, theoreticians proposed models that made the qualitative ideas of Sperry and others more precise. An early model suggested that, for retinal and tectal gradients of 'labels' (receptors and ligands) that run in the same direction, all retinal axons seek to connect to tectal regions with the highest level of tectal label. Competition for tectal space ensures that only those retinal axons with the highest available level of retinal label can connect to any particular region. A map forms because the 'best' retinal axons occupy the 'best' regions of tectum, and so on down the line. In more abstract multiple constraint models, systems matching arises as the optimal balance between the opposing forces of competition for tectal space and matching by stable gradients. Other models address retinal axon navigation over a continuous substrate in response to gradients running in opposite directions. In the future it will be important for such models to take into account the rapidly expanding experimental literature on Eph/ephrin gradients and their interactions.

## ACTIVITY-DEPENDENT REFINEMENT OF CONNECTIVITY

In the fourth stage of neural development, the initially crude patterns of connections that were established by molecularly based axon-guidance mechanisms are refined by neural activity. Patterns of firing among groups of neurons, generated either internally or by sensory transduction of signals from the environment, help to mold the architecture of the nervous system for optimum matching with the structure of the external world. This process of continued refinement continues throughout life, and it is generally believed that the mechanisms of synaptic plasticity that operate during development are similar to those that govern adult learning. Recently it has also become clear that as well as refinement there is also growth (i.e. the generation of new neurons and new connections), and that this continues throughout life.

### Hebbian Learning

The most basic principle of activity-dependent synaptic plasticity is the rule proposed by Hebb in

1949, which can be paraphrased as 'cells that fire together wire together'. That is, the connection between a presynaptic neuron and a postsynaptic neuron is strengthened when their activity is correlated. This is directly illustrated experimentally by the phenomenon of long-term potentiation (LTP). This was first discovered in the hippocampus, but has now also been observed elsewhere (e.g. in the visual cortex). Here synaptic strengths can be increased when an input pathway is strongly stimulated, causing a correlated depolarization of the postsynaptic cell, or when weaker stimulation of the input pathway is paired with direct depolarization of the postsynaptic neuron. A related phenomenon is long-term depression (LTD), where under certain circumstances synaptic strengths can decrease due to a lack of correlation between pre- and postsynaptic activity. Further experimental evidence for Hebbian learning mechanisms comes from the development of the neuromuscular junction and the visual system, which will be discussed below.

Hebb's qualitative statement of his rule can be made mathematically precise in a number of different ways. Each of these versions leads to different outcomes when the same series of input patterns is presented. Since postsynaptic activity is crucial, one important source of variability in the mathematical statements of Hebb is how this activity is calculated as a function of the activity of presynaptic neurons and the strengths of connections ('weights'). One common assumption is that the output is just a linear weighted sum of activities multiplied by weights. The change in the weight vector at each time can often then be expressed as the covariance matrix of the presynaptic input patterns multiplied by the current weight vector. The final weight pattern that then emerges is the principal eigenvector of the covariance matrix. This is computationally useful, since this weight vector is the principal component of the input patterns. This is the direction of maximum variance in the input data, which in one sense is the most 'interesting' projection of the input data.

An alternative common assumption with regard to how the output of postsynpatic neurons is calculated is that some form of nonlinear competition exists between these neurons. This is considered (either explicitly or implicitly) to arise via lateral connections between postsynaptic neurons. In the simplest case this is implemented computationally by a 'winner-takes-all' mechanism, whereby the postsynaptic neuron in the group that has the largest initial activity in response to an input pattern is the only one that has its weights updated. What tends to emerge from this type of rule is weight vectors that point to the centers of clusters of input points. This is computationally useful since such cluster analysis can help to divide the input patterns into useful categories based on the degree of overlap of different patterns.

In both of these linear and nonlinear cases an important component is often competition in the form of normalization. Since Hebb's rule specifies only increases in synaptic strengths, without some such mechanism for decreasing weights they would all increase without limit. Therefore the total weight impinging on a postsynaptic neuron is often constrained so that when some weights are increased, others are decreased.

## The Neuromuscular Junction

One particular process in neural development that has been modeled using Hebbian learning is synaptic elimination at the neuromuscular junction. When growth cones first encounter their targets, molecular signals are passed back and forth which induce the development of the presynaptic and postsynaptic specializations that form a synapse. This has been most extensively studied in the development of synapses of motor neurons on muscle fibers. Initially, several motor neurons contact each muscle fiber (polyneuronal innervation). However, by a process of activity-dependent competition all but one of these connections is eliminated. Several Hebbian-based models of this process have been proposed, and again an important component is normalization. In fixed-resource models, terminals (e.g. measured by synaptic strength or arbor size) compete for a fixed amount of a resource (e.g. total synaptic strength or space). In fixed-rate models, terminals compete for a resource that is generated at a fixed rate (e.g. neurotrophins). In these models a critical variable is the function which describes the dependence of receptor production rate on the amount of neurotrophin that is bound. Such models make testable predictions with regard to experimental perturbations such as a reduction in the supply of neurotrophin or a change in the level of activity.

## Visual System Development

Processes of activity-dependent development have been most extensively modeled in the visual system, due to the large amount of experimental data that is available to constrain and inspire theories. As mentioned previously, the map from the retina to the optic tectum is topographic. If neural

activity in the retina is blocked, a topographic map forms that is cruder than normal, indicating that activity is important for map refinement. This refinement can be modeled using Hebb-type rules, with two important assumptions. The first assumption is that 'neighborliness' between neurons in the retina is encoded by spatially correlated activity, so that nearby neurons are more highly correlated than more distant neurons. This is true of natural images – the correlation between pixels tends to drop off smoothly with distance. Similar correlations can also be generated earlier in development, before eye opening, by spontaneous activity of retinal neurons in the form of waves or blobs of activity that sweep periodically across the retina. The second assumption is that neighborliness between neurons in the tectum is encoded by lateral connections, so that nearby neurons have stronger connections and are thus more highly correlated than more distant neurons. These connections are often assumed to be in the form of short-range excitation and longer-range inhibition. Although there is some evidence that such connections exist in reality, there is also evidence for longer-range excitatory connections which are patchy rather than spatially uniform.

Another important area of theoretical modeling of activity-dependent processes in the visual system is the development of ocular dominance and orientation columns. In the adult visual cortex of mammals such as cats and monkeys, a beautiful mosaic pattern of varying response properties is seen. Each neuron responds better to an input in one eye than to one in the other eye, and to an edge or bar of light at a specific orientation angle within its receptive field. The response properties of nearby neurons mostly vary smoothly as one moves across the surface of the cortex, so that a patch of neurons that prefers one orientation blends into a patch that prefers a similar but slightly different orientation, and there is a smooth variation in preference for one eye compared with the other eye (Figure 1). However, these patterns of stimulus preference are not present in very young animals, and they may emerge by a process of activity-dependent Hebbian learning during development. The evidence for this is particularly strong in the case of ocular dominance columns, since their overall structure can be altered by changing the correlations in the visual stimuli that are seen during the critical period for ocular dominance development. For instance, if one eye is sutured shut early in development, the size of the regions representing that eye in V1 shrink relative to the regions representing the other eye.

The development of ocular dominance and orientation columns has been modeled by a variety of different versions of Hebbian learning. Three of the most important classes of models are high-dimensional, low-dimensional and filtering models. In high-dimensional models there is an explicit representation of each neuron (or each small group of neurons) in both the cortex and input layers (retina and/or lateral geniculate nucleus), and of all of the connections between layers. Patterns of stimulation are presented to the retina representing images either from the environment or generated by spontaneous activity. Activity is propagated to the cortex, and the lateral spread of activity through intracortical connections is taken into account. Weights are then updated by a Hebbian rule. Under appropriate conditions, orientation and ocular dominance columns can emerge in the cortical layer. Some high-dimensional models include hypotheses about the role of neurotrophins in mediating competition between cortical neurons. Low-dimensional models are more abstract, and instead of presenting explicit patterns of pixel intensity they consider that input images can be characterized by a small number of features, such as position in the visual field, orientation and ocular dominance. Each of these features is represented by an orthogonal dimension, and the weights of a cortical neuron now represent the selectivity of that neuron in the low-dimensional space of features. The learning rules can be seen as trading off 'matching' and 'stretching', so that all feature combinations in the input space are represented, while at the same time maximizing the degree to which neighboring cells in the cortex represent similar features. Although such models are more difficult to interpret biologically than high-dimensional models, they can often be more robust, and they actually tend to produce a better match with the fine structure of real orientation and ocular dominance maps (Figure 1). More abstract still are the filtering models, which essentially repeatedly convolve the cortical layer with a filter, such as a difference of gaussians, as an abstract representation of the effect of intracortical processing on the cortical input.

## SUMMARY

The way in which the immense complexity of the nervous system arises during development is still somewhat nuclear. Theoretical modeling has been applied to several different stages of neural development in the hope of shedding light on some of these mysteries. The best developed areas are axon

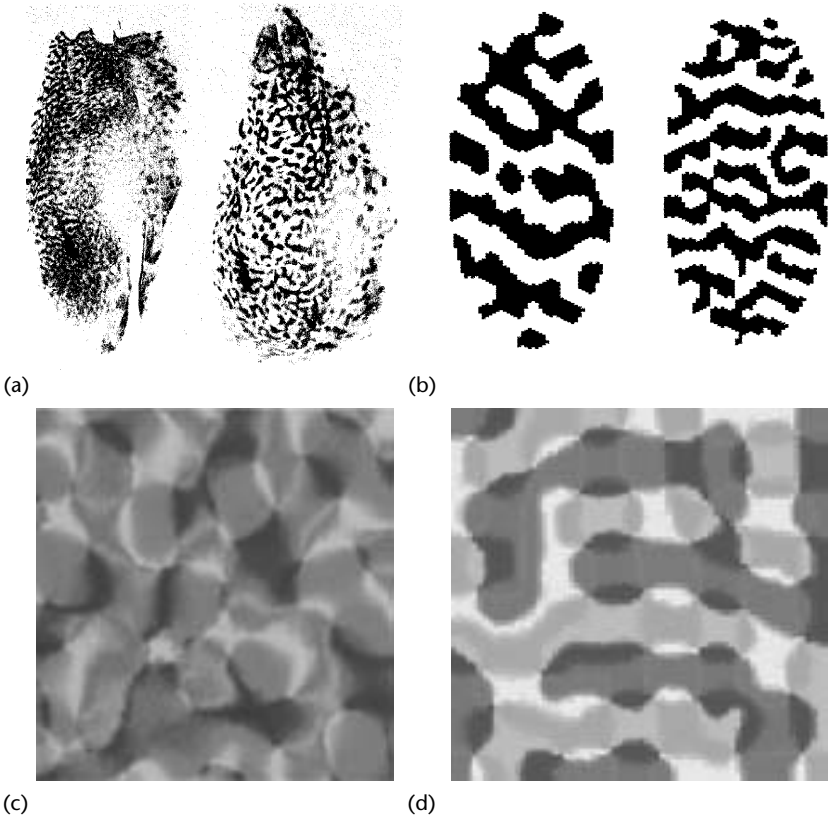


**Figure 1.** [*Figure is also reproduced in color section*.] (a) The pattern of ocular dominance columns in cat primary visual cortex. Dark patches are those regions of cortex dominated by one eye, and light patches are regions dominated by the other eye. Left: normal cat; right: strabismic cat. Note that columns are wider in the strabismic cat than in the normal cat. (From Löwel (1994) *Journal of Neuroscience* **14**: 7451–7468. Copyright Society for Neuroscience.) (b) Results of the author's analogous simulations using the elastic net algorithm. Left: 'normal' case; right: 'strabismic' case. (c) The orientation map in primary visual cortex of the tree shrew. The different colors represent patches that have different orientation preferences. (From Bosking *et al*. (1997) *Journal of Neuroscience* **17**: 2112–2127. Copyright Society for Neuroscience.) (d) Results of the author's analogous simulation using the elastic net algorithm.

growth, branching and gradient detection, and activity-dependent development in the visual system and at the neuromuscular junction. In each case, specific quantitative predictions can be derived. The comparison of these predictions with experimental data allows the assumptions underlying the models to be developed and refined.

## Further Reading

Brown MC, Keynes RJ and Lumsden A (2001) *The Developing Brain*. Oxford: Oxford University Press.

Churchland PS and Sejnowski TJ (1992) *The Computational Brain*. Cambridge, MA: MIT Press.

Goodhill GJ and Richards LJ (1999) Retinotectal maps: molecules, models, and misplaced data. *Trends in Neurosciences* **22**: 529–534.

Katz LC and Shatz CJ (1996) Synaptic activity and the construction of cortical circuits. *Science* **274**: 1133–1138.

Mueller BK (1999) Growth-cone guidance: first steps towards a deeper understanding. *Annual Review of Neuroscience* **22**: 351–388.

Murray JD (1993) *Mathematical Biology*, 2nd edn. New York: Springer.

Price DJ and Willshaw DJ (2000) *Mechanisms of Cortical Development*. Oxford: Oxford University Press.

Swindale NV (1996) The development of topography in the visual cortex: a review of models. *Network* **7**: 161–247.

Wolpert L *et al*. (1998) *Principles of Development*. Oxford: Oxford University Press.

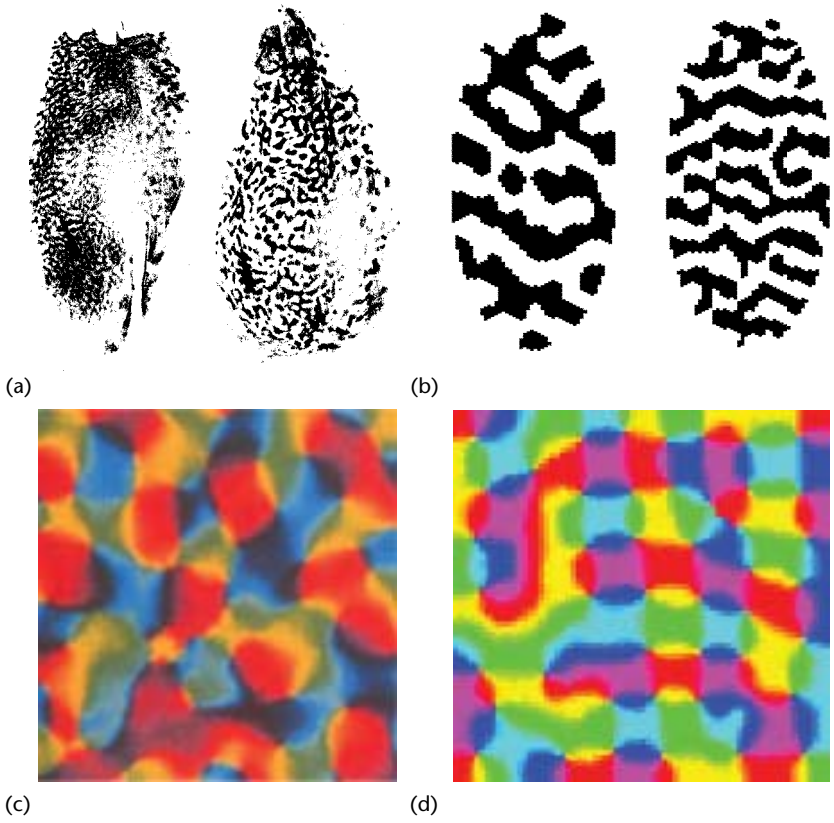


**Plate 2 [Neural Development, Models of]** (a) The pattern of ocular dominance columns in cat primary visual cortex. Dark patches are those regions of cortex dominated by one eye, and light patches are regions dominated by the other eye. Left: normal cat; right: strabismic cat. Note that columns are wider in the strabismic cat than in the normal cat. (From Löwel (1994) *Journal of Neuroscience* **14**: 7451–7468. Copyright Society for Neuroscience.) (b) Results of the author's analogous simulations using the elastic net algorithm. Left: 'normal' case; right: 'strabismic' case. (c) The orientation map in primary visual cortex of the tree shrew. The different colors represent patches that have different orientation preferences. (From Bosking *et al*. (1997) *Journal of Neuroscience* **17**: 2112–2127. Copyright Society for Neuroscience.) (d) Results of the author's analogous simulation using the elastic net algorithm.