

Neural Darwinism: Selection and Reentrant Signaling in Higher Brain Function

Review

Gerald M. Edelman
The Neurosciences Institute
New York, New York 10021

Summary

Variation and selection within neural populations play key roles in the development and function of the brain. In this article, I review a population theory of the nervous system aimed at understanding the significance of these processes. Since its original formulation in 1978, considerable evidence has accumulated to support this theory of neuronal group selection. Extensive neural modeling based on the theory has provided useful insights into several outstanding neurobiological problems including those concerned with integration of cortical function, sensorimotor control, and perceptually based behavior.

Introduction

Over the last several decades, remarkable advances have occurred in the molecular and cellular biology of the nervous system. A vigorous resurgence has also taken place in cognitive psychology. Nevertheless, a large gap remains in our understanding of the biological bases of psychological phenomena. It is unlikely that this gap will be filled simply by more extensive experimental observations in each subject domain, important as these are. What is needed is a theoretical framework sufficiently broad to connect biology and psychology in a fashion consistent with developmental and evolutionary mechanisms. The theory of neuronal group selection was proposed to provide such a framework (Edelman, 1978). Its formulation was prompted by the need to reconcile two sets of observations that seemed inconsistent with the then prevalent notions of brain function:

First, individual nervous systems (particularly those of vertebrate species) show enormous structural and functional variability. This striking variability was explicitly noted by Lashley (1947), who at the time could offer no explanation for it. Variability occurs in both space and time at many levels: molecular, cellular, anatomical, physiological, and behavioral (Edelman, 1987). Although there is an obvious commonality of neural structure within a species, the degree of individual variation far exceeds that which could be tolerated for reliable performance in any machine constructed according to current engineering principles. Notwithstanding this variation, adaptive behavior characteristic of each species emerges in the course of individual development.

Second, understanding the development of such adaptive behavior faces an additional challenge. The world of stimuli encountered by a newborn animal cannot be described adequately as preexisting, unam-

biguous information ready to be manipulated according to a set of rules similar to those followed by a computer executing a program. Although the real stimulus world certainly obeys the laws of physics, it is not uniquely partitioned into "objects" and "events" (Smith and Medin, 1981). To survive in its econiche, an organism must either inherit or create criteria that enable it to partition the world into perceptual categories according to its adaptive needs. Even after that partition occurs as a result of experience, the world remains to some extent an unlabeled place full of novelty (Edelman, 1987).

Contrary to these two sets of observations, information processing or functionalist views of higher brain function (see, for example, Underwood, 1978; Pylyshyn, 1987) assume that both neuroanatomy and the significance of signals are more or less fixed and unambiguous (like instructions) and that the brain is a kind of computer. The theory of neuronal group selection disputes this view. It argues that the ability of organisms to categorize an unlabeled world and behave in an adaptive fashion arises not from instruction or information transfer, but from processes of selection upon variation. In this view, the otherwise puzzling variability of individual brains is a feature that is central to their function. Like the theories of natural selection and of clonal selection in immunity, the theory of neuronal group selection is a population theory. According to the theory of neuronal group selection, the world becomes "labeled" or perceptually categorized as a consequence of two interactive processes of selection upon variation. The first occurs largely in embryonic and postnatal development, during which adjacent neurons tend to be strongly interconnected in collectives of variable size and structure called neuronal groups. The second process consists of alterations in synaptic strengths during an animal's activity, selecting the correlated responses of those neuronal groups that yield adaptive behavior.

Challenges to the Instructionist View

The burden of this theory is to show in detail how both perceptual and conceptual categorization can arise as a result of selection upon preexisting variance in structure and function of the nervous system. Before taking up this burden, however, it will be useful to mention briefly a number of observations consistent with a role for continuing variation and selection in the nervous system. These observations present insurmountable difficulties to the functionalist view that the world is like a computer tape (full of "instructions") and that the brain is a computer following instructions.

At the level of their finest connectivity, rich nervous systems such as those of vertebrates do not have precise, prespecified, point to point wiring (see Cowan,

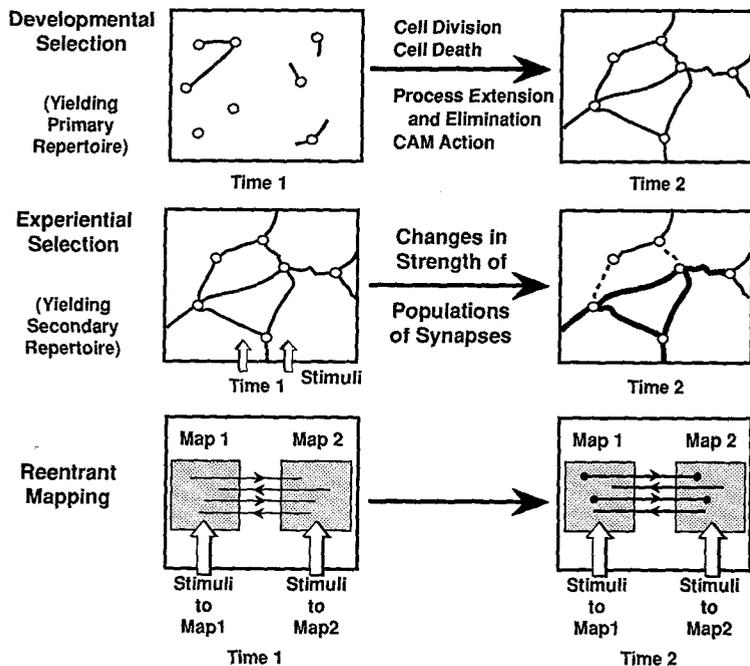


Figure 1. Basic Tenets of the Theory of Neuronal Group Selection, a Global Theory of Brain Function

Top: Developmental selection. This occurs as a result of the molecular effects of CAM and SAM regulation, growth factor signaling, and selective cell death to yield variant anatomical networks in each individual. These networks make up the primary repertoire. Center: Experiential selection. Selective strengthening or weakening of populations of synapses as a result of behavior leads to the formation of various circuits constituting a secondary repertoire of neuronal groups. The consequences of synaptic strengthening are indicated by bold paths; those of weakening, by dashed paths. Bottom: Reentry. Binding of functionally segregated maps occurs in time through parallel selection and the correlation of the various maps' neuronal groups. This process provides a fundamental basis for perceptual categorization. Dots at the ends of some of the active reciprocal connections indicate parallel and more or less simultaneous strengthening of synapses facilitating certain reentrant paths. Synaptic strengthening (or weakening) can occur in both the intrinsic and extrinsic reentrant connections of each map.

1978, and Edelman, 1988, for reviews). Not only are the majority of anatomical connections not functionally expressed at any one time, but there are major fluctuations in the physiologically detected boundaries of the neural territories and maps to which these connections contribute (Merzenich et al., 1983). Moreover, such maps are unique to each individual.

If we shift from structure to the level of psychological function, two other remarkable observations deserve notice. First, even in species without language, animals are capable of a remarkable range of generalization in their perceptual categorizations. When presented with a few examples of particular shapes belonging to a common class (say, photographs of fish), pigeons can respond effectively to large numbers of different novel shapes in the same class (Herrnstein, 1985). Second, objects and their properties are perceived to be unitary, despite the fact that a given perception results from parallel activity in the brain of many different maps, each with different degrees of functional specialization. A striking case is the extrastriate visual cortex (Zeki, 1981), in which different areas specialized for color, motion, and form act together to yield a coherent response to an object. There appears, however, to be no single superordinate executive map that links these variously mapped attributes or that binds them into a scene consisting of different objects. An adequate theory of higher brain function must account for this "binding problem" without invoking preexisting instructions from a computer executive or a homunculus.

All these observations challenge the instructionist

view (see Postscript in Edelman, 1992) that the actions of the nervous system can be accounted for by information processing operating upon codes. A different view is required. Because it is a population theory, the theory of neuronal group selection provides such an alternative—instead of ignoring the observed variance and fluctuations in neuroanatomy and neural dynamics, it treats them as key features that are essential to the function of the nervous system. By invoking a higher order correlative process called reentry (which I shall define below), it also provides a solution to the binding problem.

Basic Selectional Mechanisms

The theory of neuronal group selection proposes three mechanisms to account for the production of adaptive behavior by rich nervous systems (Edelman, 1987): developmental selection, experiential selection, and reentrant signaling (Figure 1). Each mechanism acts within and among collectives consisting of hundreds to thousands of strongly interconnected neurons, called neuronal groups. Neurons within a group are highly interconnected, and changes in their synaptic strengths tend differentially to enhance the adaptive responses of the group as a whole. While the structures underlying neuronal groups arise from local anatomical connections, the groups themselves are dynamic entities whose borders and characteristics are affected by such synaptic changes and by the nature of the signals the groups receive.

Before considering evidence for the proposal that

neuronal groups are basic units of selection, I shall give a brief account of each of the selectional mechanisms of the theory illustrated in Figure 1:

Developmental Variation and Selection

The structural diversity of the nervous system is not strictly programmed by a molecular code. Instead, it arises during development from dynamic epigenetic regulation of cell division, adhesion, migration, death, and neurite extension and retraction (Changeux and Danchin, 1973; Cowan, 1978; Edelman, 1988; Rakic, 1988). Neuronal adhesion and migration are governed by a series of morphoregulatory molecules called cell adhesion molecules, or CAMs, and substrate adhesion molecules, or SAMs (Edelman, 1988; see Takeichi, 1990, Edelman and Crossin, 1991, for reviews). These molecules interact with neuronal surfaces and affect the dynamics of cellular interactions as they occur at particular neural sites. Recent evidence suggests that such morphoregulatory molecules are under control of homeobox-containing genes that are known to govern place-dependent morphology (Jones et al., 1992a, 1992b). The temporal patterns and levels of expression of morphoregulatory molecules, while characteristic of a given anatomical area, are nevertheless dynamically regulated and are subject to epigenetic influences. A key process in this regulation, particularly at later stages of development, is the selection of morphoregulatory molecule expression and neural structure by correlated neural activity itself. Under these influences, morphoregulatory molecules affect cell motion and process extension, leading to enormous local neuroanatomical variation at the finest ramifications of axons and dendrites. In certain regions, there is also a large amount of cell death, which occurs stochastically in particular neuronal populations. This unavoidable generation of diversity results in the formation within a given anatomical region of *primary repertoires*, consisting of large numbers of variant neuronal groups or local circuits. This occurs despite the fact that the overall pattern in any particular specialized region is similar from individual to individual.

Experiential Selection

After most of the anatomical connections of the primary repertoires have been established, the activities of particular functioning neuronal groups continue to be dynamically selected by ongoing mechanisms of synaptic change driven by behavior and experience. This selection occurs within populations of synapses, strengthening some and weakening others without major changes in anatomy. Experiential selection leads to the formation of *secondary repertoires* of neuronal groups in response to particular patterns of signals. Because of the changes that occur in synaptic efficacies, upon later encounters with signals of similar types, the previously selected circuits and neuronal groups in such secondary repertoires are more likely to be favored over others (Figure 1).

Unlike natural selection in evolution, which results

from differential reproduction, experiential selection results from differential amplification of synaptic populations. The synaptic changes do not represent information that is stored in individual connections between single neurons, as in connectionist models. Instead, signals act, often heterosynaptically, to select variant *populations* of synapses that connect cells within and between neuronal groups (Finkel and Edelman, 1987). Experiential selection thus involves statistical signal correlations between *groups* of pre- and postsynaptic neurons rather than the carriage of coded messages from one neuron to another. Nevertheless, if these statistical correlations are to serve adaptive behavior, they must reflect the spatiotemporal properties of signals arising in the real world. This is accomplished by reentrant signaling.

Reentrant Signaling

Although the perceptual categories constructed by the brain from sensory signals cannot be fixed prior to neuronal group selection, physical laws such as those of spatiotemporal continuity obviously must be obeyed. Neural mappings relating sensory receptor sheets to particular regions of the central nervous system provide one means to enforce such regularities. The activation of the resultant neural maps must be coordinated during adaptive behavior. Given the local variance in neural structure and connectivity and the variant statistics of synaptic change, however, specific tags or labels are *not* available to specify any given map position, as they are in a computer representation.

How then can different maps be coordinated? The theory of neuronal group selection proposes that mapped regions exchange and coordinate signals by a higher order selectional process called reentry (Figure 1). Reentry can be defined as ongoing parallel signaling between separate neuronal groups occurring along large numbers of ordered anatomical connections in a bidirectional and recursive fashion. Reentrant signaling can take place via reciprocal connections between and within maps (as seen in cortico-cortical, corticothalamic, and thalamocortical radiations), as well as via more complex arrangements seen in the connections among the cortex, basal ganglia, and cerebellum (Gerfen, 1992; Edelman, 1989). Reentry is a dynamic process that is inherently parallel and distributed.

Although it can occur within a single map, reentry usually involves correlative signaling between at least two maps, and it acts through ordered connections that sample these maps in both space and time. The simultaneous activation of neuronal groups in different maps by a given stimulus and the effects of previous reentrant activity both tend to strengthen some of the connections between those groups. This coordinated dynamic interaction across maps results in temporal correlation of the responses of a subset of groups to disjunct signals travelling in separate channels to each map. By these means, distinct operations in different maps that are related to the same percep-

tual stimulus can be linked to one another. A minimal arrangement of two reentrantly connected maps is called a "classification couple." As shown in Figure 1, after multiple encounters with a stimulus, particular patterns of groups are selected in each mapped area. This process is called reentry because the results of selection are shuttled back and forth between maps in a recursive fashion. With continued reentrant interactions across maps leading to temporally coordinated synaptic changes, responses to certain features or attributes of a stimulus object may become linked. Such reentrant activity is constructive: because of its reciprocal and recursive properties and its parallel structure, reentry leads to new neuronal responses and it can resolve conflicts arising between the synaptic activities of different mapped areas (Finkel and Edelman, 1989). It should be sharply differentiated from feedback. Feedback is concerned with error correction and defined inputs and outputs, whereas reentry has no necessary preferred direction and no predefined input or output function. Moreover, unlike reentry, which involves parallel processes, a given feedback loop, to be effective, involves only a single signal channel or pair of wired connections (Ashby, 1966).

Predictions and Supporting Evidence

To be useful, a global brain theory such as the theory of neuronal group selection should satisfy two major requirements: First, it should help to predict fundamental structural and functional relationships among the components of the central nervous system. Second, it should provide a basis for recording and understanding the wide variety of morphological, physiological, and behavioral observations obtained in various neuroscientific subdisciplines. Given that nervous systems have features which arose unpredictably in evolutionary history, the second requirement is by far the more important. Nevertheless, certain predictions of the theory are fundamental in the sense that, if they fail, the theory should be discarded. The central one, first proposed in 1978 (Edelman, 1978), concerns the existence of neuronal groups as targets of selection in the brain.

Although single neurons could occasionally serve as units of selection, the theory of neuronal group selection argues that neuronal groups, not single neurons, are the sufficient basis for the mapping interactions proposed by the theory. The existence of neuronal groups is supported by both logic and experimental evidence. The logical argument is as follows: in the presence of numerous densely connected interneurons in regions such as the cortex (see for example, Schmitt et al., 1981), it is difficult to imagine a neuron acting as an isolated or individual unit. Indeed, cooperative local signaling that couples neighboring neurons would seem to be inevitable, particularly in an anatomy such as that of the cortex, in which mapping and local connectivity predominate. Ana-

tomical arrangements of this type facilitate the convergence of multiple correlated inputs onto a single functional unit such as a neuronal group. Furthermore, the collective action of neurons organized in groups would enhance the reliability of the system, counteracting neuronal death and occasional nonadaptive selectional events affecting individual cells.

Most importantly, group organization provides a basic structural underpinning for the process of reentry, which, given the wide lateral spread of overlapping local dendritic and axonal arbors, could hardly be envisioned as occurring from single neuron to single neuron. Finally, group structure and variation is consistent with the events of embryonic development as well as with the plasticity of cortical neuronal function that is found even in the adult.

So much for logic. What about evidence? A great deal of supportive evidence can be generated from cortical anatomy itself. In regions of the central nervous system where specific roles can be assigned to neurons, local mosaic arrangements that provide a natural basis for a functional arrangement into neuronal groups are observed. These include ocular dominance columns, blobs, slabs, barrels, fractured somatotopies, etc. The characteristic patchy shape and sparse connectivity of cortical axonal arbors also accord well with the existence of groups.

Recent experiments have provided direct evidence for the existence of groups in the visual cortex (Gray and Singer, 1989). A moving, oriented bar was presented to a cat while simultaneously recording local field potentials and the firing of single neurons in area 17. Coherent oscillations at about 40 Hz were observed in both kinds of measurements, but only as long as the stimulus was present. Moreover, simultaneous presentation of two different stimulus objects revealed separate and independent patterns of correlations for each. In other experiments, it was shown that when a single oriented stimulus was presented, 40 Hz oscillations are correlated in regions of widely separated areas of the visual cortex, for example in areas 17 and 18. More recently, correlated oscillatory behavior was observed in the two cerebral hemispheres during the presentation of similar stimuli (Engel et al., 1991). Cutting the callosal connections abolished these correlations.

The strong temporal correlations between single unit firing and local field potentials in these experiments indicate that, upon presentation of a stimulus, neighboring neurons respond in a coordinated fashion and thus constitute a neuronal group. In addition, the separate observations of correlations between distant neuronal groups in the striate and extrastriate cortex and across the callosum provide direct evidence for the dependence of long-range selective interactions on the process of reentry.

Like Darwin's theory of natural selection, the theory of neuronal group selection is based on the continual generation of diversity, with selection occurring at various levels. In the embryonic and maturing brain,

variation and selection occur in migrating cellular populations, during cell death, as well as during synapse formation, and both processes are dramatically reflected in enormous synaptic loss. In the mature brain, variation and selection are seen mainly in the differential amplification of synaptic efficacies that result in neuronal group formation, a process that is continually modified by reentrant signaling. Given these population properties and the prevalence of continual variation and selection, the metaphor "neural Darwinism" is not inappropriate. The evidence suggests that recent criticisms (Crick, 1989; Barlow, 1988; Purves, 1988) of the proposals of neuronal group selection have been based on misconceptions about the nature of selection (Michod, 1989; Edelman, 1992).

Explanatory Power and Self-Consistency of the Theory of Neuronal Group Selection

In neuroscience, there is an extraordinary variety of phenomena demanding explanation in terms of fundamental principles of brain organization. Examples of such phenomena include the diversity of neurotransmitters and their receptors, the remarkably plastic properties of cortical maps, the problem of cortical integration and figure-ground segregation during perception, and the development of precise motor coordination in neuromuscular assemblies despite the presence of excess degrees of freedom across joints. In a series of detailed models, the theory of neuronal group selection has been shown to provide unifying insights into each of these phenomena.

Multiplicity of Neurotransmitters and Receptors

Work over the last decade has led to the discovery of a large number of neurotransmitters and neuropeptides, as well as a great diversity of receptors and channels. We lack satisfactory hypotheses to account for this profusion. An attractive idea consistent with the theory of neuronal group selection arises from the view that synapses function in populations. The existence of relatively precise biochemical and electrical requirements and, above all, the timing requirements for synaptic transmission and modification of synaptic efficacy indicate that there are strong constraints on the number of functioning circuits that can be supported by any single transmitter. An increase in the number of neurotransmitters and neuromodulators would, however, be expected to increase greatly the number of functional circuits that are combinatorially possible within a given anatomical network. A kind of "transmitter logic" can be defined in which combinations of transmitters, operating in a particular anatomical context, provide part of the diversity required for selection of interactive secondary repertoires (Finkel and Edelman, 1987). A rich pharmacology thus ensures a very rich set of functional network variants. The evolution of a large number of transmitter types is inexplicable in terms of functionalist theories that

emphasize software and consider the brain as an information processing system while ignoring its physical instantiation. In contrast, the existence of an increasing diversity of transmitters, channels, and new signaling modes is fully consistent with a selectional model of brain function.

Such a model does not invoke coded signals acting in precisely predefined linear circuits, but instead emphasizes spatial and temporal correlations occurring within volumes of tissue. In line with this concept, my colleagues and I have recently proposed (Gally et al., 1990) that nitric oxide synthesized at sites of synaptic activity serves as a short-lived spatial signal that rapidly diffuses through the surrounding tissue. The suggestion that this signal might act to couple neural activity to local blood flow and to modulate synaptic strengths in accordance with local patterns of synaptic activity has subsequently received experimental verification in a number of laboratories (Shibuki and Okada, 1991; O'Dell et al., 1991; Böhme et al., 1991; Schuman and Madison, 1991; Haley et al., 1992; Northington et al., 1992). In computer simulations, we have shown how an activity-dependent spatial signal of this sort might also give rise by a selective process to topographic maps during cortical development (Montague et al., 1991; Gally and Edelman, 1992). The patterns of connectivity generated within these simulations define local groups of excitatory neurons that are strongly interconnected with one another and form sparse, reciprocal excitatory axonal contacts with other groups.

Plasticity of Cortical Maps

In maps of the somatosensory cortex of adult owl and squirrel monkeys, alterations of input due to nerve section or the repeated imposition of particular stimuli lead to changes in individual map boundaries, particularly in areas 3b and 1. (Merzenich et al., 1983; Kaas et al., 1983). These changes occur both acutely and chronically. Before and after the changes, the maps have receptive fields with sharp continuous borders no thicker than one or two cell diameters, despite the fact that the arbors of input axons from the thalamus extend over much wider distances (Landry and Deschenes, 1981). If single neurons were units of selection during such alterations, one would not expect continuous sharp boundaries, but rather a "salt and pepper" pattern or a more diffuse arrangement of receptive field boundaries. In each area in which correlated sets of inputs compete for cortical neurons (for example, those from the glabrous or dorsal skin of the hand), the responding neurons appear to segregate into groups, which at any one time are nonoverlapping and have sharp boundaries. A computer simulation (Pearson et al., 1987) based on group formation and competition neatly shows how these boundaries and their underlying neuronal groups may be dynamically sustained and yet change systematically with time under different stimulus conditions.

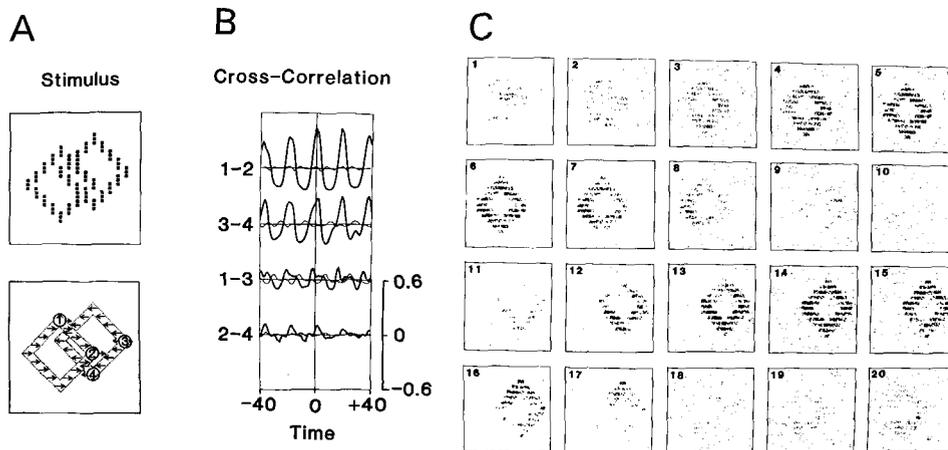


Figure 2. Consequences of Reentry in a Neural Model of Perceptual Grouping and Segmentation (Sporns et al., 1991)

(A) A stimulus consisting of two identical patterns, each composed of vertically oriented bars, is presented to the model. The two patterns overlap in visual space but move in different directions. In the top panel, the bars are shown at their starting positions; in the bottom panel, their corresponding directions of movement are indicated by arrows. Encircled numbers with arrows in the bottom panel refer to the locations of recorded neuronal activity; corresponding cross-correlations are displayed in (B). "Electrodes" 1 and 2 recorded from neurons responding to pattern 1, and "electrodes" 3 and 4 from neurons responding to pattern 2.

(B) Cross-correlograms computed over a 10 ms sample period and subsequently averaged over 10 trials. Numbers refer to the locations of responding direction-selective repertoires containing neuronal groups that are analyzed for their correlations (see [A]). Four correlograms computed between 201 and 300 ms after stimulus onset are shown. The correlograms are scaled, and shift predictors (thin lines; averages over nine shifts) are displayed for comparison.

(C) Frames taken from a movie showing the responses of direction selective groups in the model to the stimulus in (A). The frames show a continuous period of 20 (20 iterations) recorded about 150 ms after stimulus onset. Each frame displays the model's entire array of neuronal groups (16×16 for a total of 50,720 cells) selective for motion to the right and to the left and arranged in an interleaved fashion (this accounts for the striped pattern). Each small dot within the array is an active neuron. For the first 10 ms (frames 1-10), groups responding to the pattern moving right are mainly active; subsequently these groups are silent, and groups responsive to the other pattern become active (frames 11-20). Note that neuronal activity is strongly correlated both within groups as well as over the entire extent of each pattern.

(A and B) Modified from Sporns et al., 1991; reproduced with permission. (C) Modified from Tononi et al., 1992; reproduced with permission.

Neural Integration of Functionally Segregated Maps

The organization of the cerebral cortex is such that, even within a single sensory modality such as vision, there is a multitude of specialized areas and functionally segregated maps. What we are aware of, however, is a unitary and coherent perceptual scene that seems to be a prerequisite for adaptive behavior. The unitary appearance of the perceived world and our ability to act coherently in the presence of diverse and often conflicting sensory stimuli require a process of neural integration that must occur at many different levels of organization.

A key question is how such integration takes place. Within a single cortical area, "linking" must occur among the responses of neuronal groups that belong to the *same* feature domain. Perceptual grouping within a single submodality such as color or movement provides an example of integrative linking at an early level. At a higher level, "binding" must take place among the responses of neuronal groups found in *different* feature domains that are distributed in different cortical areas. An example is the integration of neuronal responses to a particular object contour with its color, position, and direction of movement. Such perceptual and behavioral integration across functionally segregated maps can occur in times rang-

ing from 50 to 500 ms, placing strong temporal constraints on any proposed mechanism.

Reentry, occurring both locally and between maps, is the basic means by which integration occurs in the absence of any single cortical "master map." The self-consistency of this proposal has been tested extensively in a series of computer models (Sporns et al., 1989, 1991; Tononi et al., 1992). It has been shown that reentrant interactions within a single cortical area can give rise to temporal correlations between neighboring as well as distant groups with a near-zero phase lag, as observed in cats and monkeys (Gray and Singer, 1989; Engel et al., 1991; Kreiter and Singer, 1992). An early computer simulation (Sporns et al., 1989) showed how linking can be mediated by reentry. In agreement with experimental data, when a continuous long bar was presented as a moving stimulus to the model, correlations were found between units in groups with nonoverlapping receptive fields. These distant correlations disappeared if two colinear short bars that were separated by a gap were moved with the same velocity. A more extended model (Sporns et al., 1991) was presented with a pattern composed of several bars moving coherently together but embedded in a background of vertical and horizontal bars that were moving at random to the right and left, or up and

down. The neuronal groups responding to the bars that moved in the same direction were rapidly linked by coherent oscillations through reentry even though the lateral spread of the anatomical connections from each neuronal group was much smaller than the projected size of the "object." This model was shown to segregate a figure from another overlapping figure or from a coherent background of identical texture moving in a different direction (Figure 2).

All of these simulations that demonstrated linking depended strongly upon the occurrence of rapid changes in synaptic efficacy. The neural mechanism for integration and segregation of elementary features into objects and background is based on the pattern of temporal correlations and phase relationships among neuronal groups. These correlations depended critically upon reentry and disappeared when the underlying connectivity was disrupted. The resulting figural grouping and segregation are consistent with the Gestalt laws (Köhler, 1947) of continuity, proximity, similarity, common orientation, and common motion, and this work thus suggests a neural basis for these laws.

What about the centrally important binding problem? In a much larger simulation (Tononi et al., 1992), it has been shown that a model with interareal reentry among nine functionally segregated areas, divided into three anatomical streams for form, color, and motion, could distinguish two or more objects present in the same visual scene regardless of position. Reentry and temporal correlation were shown to be sufficient to solve the binding problem, without any need to call upon a programmed superordinate executive area. While the chosen modality in this model was visual, there is every reason to expect that similar principles operate in other modalities throughout the brain.

Complex Motor Control and the Bernstein Problem

Experiments on neural responses of monkeys reaching for targets indicate that a given reaching movement results from the contributions of multiple populations of neurons, each tuned to a particular direction of movement (Georgopoulos et al., 1986.) A particular movement can thus be explained in terms of the activity of appropriate combinations of neuronal groups. But the development of coordinated movements poses an additional problem that cannot easily be explained by purely cybernetic models of movement control. The presence of multiple degrees of freedom in the joints of the arm and hand indicates that the system is dynamically underdetermined and therefore that certain constraints are necessary to account for the precise and rapid targeting of movements. This challenging problem of inverse kinematics, first posed by Bernstein (1967), can be solved in terms of the selection of appropriate movements from a repertoire of variant movements that result from underlying neuronal group selection (Sporns and Edelman, 1992). This has been explicitly demonstrated in a com-

puter model, called Darwin III, to be considered in the next section.

In each of the simulations briefly described above, phasic reentrant signaling (Edelman, 1978) was taken to be a central principle governing the spatiotemporal integration of higher brain functions. It is a prediction of the theory of neuronal group selection that many such dynamic reentrant organizations reflecting integration by selection will be found to exist in the brain. The development of new differential brain scanning techniques with high spatial and temporal resolution should allow this hypothesis to be tested stringently. In the meanwhile, reentry has been shown to have considerable explanatory power as a basis for cortical integration (Finkel and Edelman, 1989; Tononi et al., 1992). With this as a basis, we are now in a position to consider how the theory of neuronal group selection deals with the fundamental problem of perceptual categorization mentioned at the beginning of this article.

From Reentrant Maps to Perceptual Categorization

How does an animal with a variable nervous system partition an unlabeled world into objects and events? When an animal moves, local visual maps in the brain interact with those of other sensory modalities as well as with those guiding motor output. How are all of these maps coordinated with movement to allow rich perceptual categorization? Perceptual categorization does not occur solely in a particular sensory area, which then executes a program to activate motor output. Instead, the consequences of continual motor activity are an essential part of the perceptual process itself. Perception depends upon action. This implies that the neural structures that carry out various categorizations must contain multiple sensory and motor maps, forming what I have called global mappings (Edelman, 1987, 1989). Neuronal group selection within such mappings occurs in a set of dynamic reentrant loops that continually match gestural and postural movements to several kinds of sensory signals.

Large portions of the nervous system are involved in these loops. It is therefore not a simple matter to trace the activity of all of the neurons involved in these mappings, nor to imagine the sequential patterns generated by such structures, even when individual activities are known. For this reason, it has been heuristically valuable to test the self-consistency of the theory of neuronal group selection using an approach (Reeke et al., 1990; Edelman et al., 1992) that my colleagues and I have called synthetic neural modeling (SNM). SNM consists of the large scale computer simulation of all levels of the nervous system and the phenotype of a particular designed "organism" that is behaving in a changing and unpredictable environment. SNM permits the interactions of all simulated structural and functional levels, ranging from molecular to behavioral, to be analyzed as behavior develops. Given the historical properties of selective systems, this ap-

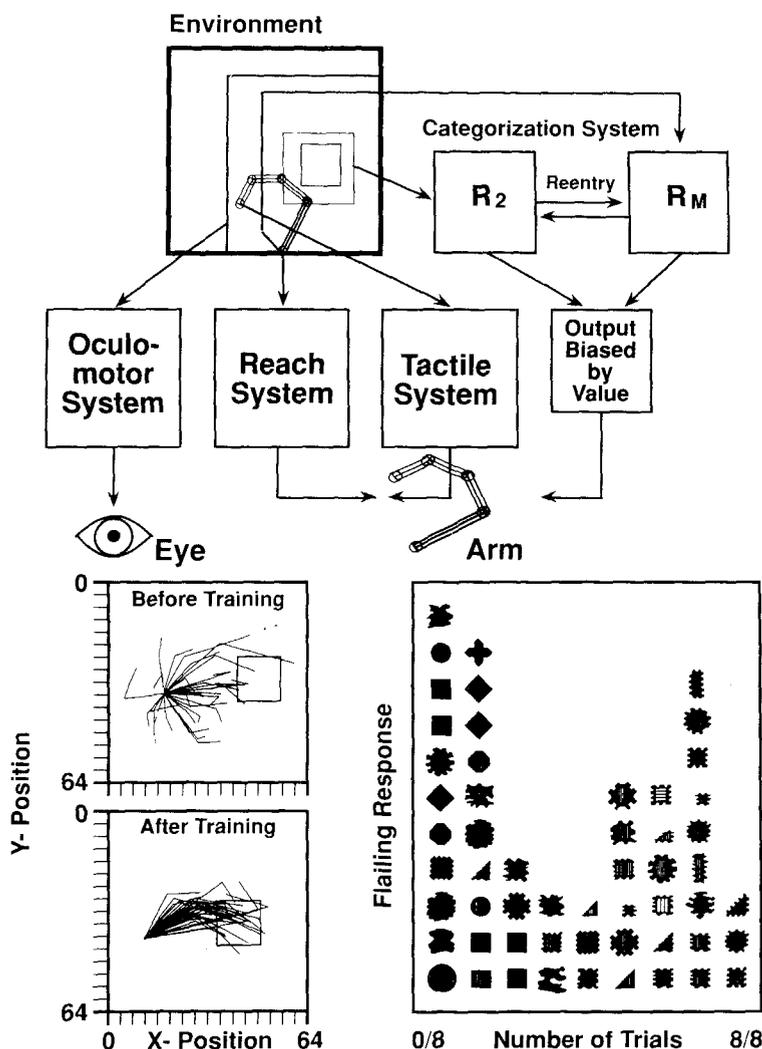


Figure 3. Darwin III, a Recognition Automaton That Performs as a Global Mapping Simulated in a Supercomputer

It has a single movable eye, a four-jointed arm with touch at the last joint, and kinaesthesia signaled by neurons in its joints as they move. Its nervous system is organized into several subsystems, each responsible for different aspects of its behavior (top). What is *programmed* in the simulation is the "evolutionary" phenotype, including neuroanatomy; the behavior of the simulation is *not programmed*. After experience with randomly moving objects that it "sees," its eye will follow any object. Similarly, its arm reaches out to "touch" an object, and with each selection of movements, it increases its success in achieving this touching (lower left). In the experiments shown at the lower left, the tip of the arm always starts in a standard location (the point of origin of the traces) and its motion toward a target area (the square box) is plotted. Notice that before training, the arm moves in many directions. After training involving selection (bottom set of traces), its movements are targeted. Darwin III was confronted (lower right) with 55 different objects and was given eight trials in which to categorize each object. The results, plotted as a positive flailing response versus the number of successful trials, indicate that Darwin III divided this collection of objects into two classes.

proach has distinct advantages over mathematical analysis alone.

The effectiveness of SNM has been tested by construction of a series of selective recognition automata. Later versions of these automata act in real environments and have some of the autonomy of biological organisms. An early automaton, Darwin III, is an explicit model of a global mapping (Reeke et al., 1990). Darwin III consists of a simple, sessile organism with a moveable eye and a four-jointed arm (Figure 3). It possesses neurons subserving contrast vision, light touch, and kinaesthesia. A detailed set of neuroanatomical and neurochemical constraints are embedded within its structure, which was designed to embody a series of evolutionary and developmental steps that might have yielded such a phenotype.

The environment of Darwin III consists of various objects of different shapes appearing and moving across its visual field. These objects are chosen and driven by a random number generator. A naive individual automaton, with its initial neural activity driven by a separate random number generator, was exposed

to these stimuli. After this exposure, selection among its neuronal repertoires under the constraint of various criteria of value resulted in consistent patterns of visual tracking, reaching with the arm, and discrimination among different simulated objects.

After "being born" into an environment of moving objects, Darwin III begins to track and fixate particular objects and to reach out to touch and trace them. With suitable experience leading to synaptic selection, the eye of Darwin III begins to make appropriate saccades and fine tracking movements with no further specification of its task than that implicit in a value scheme, to be discussed below. In a similar fashion, the arm of Darwin III can be trained to reach for and touch objects that are first detected and tracked by the visual system (Sporns and Edelman, 1992). This performance, which entails the coordination of gestural motions involving various joints (Bernstein, 1967), involves participation of a whole series of neural repertoires that perform functions similar to those carried out in the real nervous system. After appropriate experience, the global mapping represented by Dar-

win III was capable of perceptually segregating striped bumpy "objects" of various shapes from smooth or non-striped "objects," despite the fact that its specific behavior was not driven by an explicit program (Figure 3).

Because there is no program explicitly controlling their behavior, somatic selectional systems such as Darwin III must contain structures and constraints in their phenotype that reflect prior evolutionary selection for what we have termed "values" (Reeke et al., 1990). Accordingly, specialized networks in Darwin III were designed to respond to the relative adaptive value to the automaton of its motor actions in a manner that does not require prior experience. Values are arbitrary: while specific in a given example of Darwin III, they correspond to various possible kinds of evolutionarily determined characteristics that contribute to phenotypic fitness. An example of a low level value is "seeing is better than not seeing," which is translated as "when the central retina becomes strongly stimulated, selectively increase the strengths of those synapses that were active in the recent past and thus potentially involved in the behavior that brought about the increased stimulation." It is important to stress that, in Darwin III, detection of central retinal stimulation can be accomplished by a simple anatomical adaptation that contains no programmed specification of *how* light-tracking saccades are to be accomplished. The selective amplification of synaptic strengths depends on adaptive value as registered by such internal structures and not on any measure of error supplied by an external observer or device. As expected, if the value networks are removed from the simulation, the automata show no convergent behavior. While categorization depends on value, it is not explicitly specified by it; instead, it emerges from selection based on behavior. It should be emphasized that the resulting synaptic changes in a value-dependent system are probabilistic. No two versions of Darwin III show identical behavior. Nevertheless, their behavior tends to converge in directions favored by their intrinsic value systems and phenotype.

Although the use of independent random number sequences in both the nervous system and the environment of Darwin III helped us avoid explicit instructional biasing of its behavior, various choices made in the concurrent design of neuronal networks and environmental stimuli could inadvertently have led to such a bias. Moreover, because of the simplified and specialized nature of the environment to which it was exposed, the behavior of Darwin III could not be readily compared with that of experimental animals. To avoid these limitations, we have recently (Edelman et al., 1992) applied SNM techniques to a real-world artifact, NOMAD (neurally organized multiply adaptive device). Adoption of this tactic restricts the computer simulation solely to the organization and dynamics of the nervous system; the environment and mobile artifact are real. NOMAD, a 1 m high wheeled device with a CCD camera for an eye and a sensory

snout connected to an electromagnet, can move about in its environment, providing visual and other sensory inputs to a simulated nervous system (called Darwin IV) in a supercomputer. NOMAD receives telemetered signals from the neural portions of the Darwin IV simulation that govern its behavior.

Darwin IV can execute built-in reflexes as well as several modes of ongoing sensor-guided motion that are subject to selective amplification of synaptic strengths. These behavioral modes and reflexes are combined during experience to perform a number of exemplary tasks. Under suitable constraints, Darwin IV can be trained to track a light moving in a random path and to approach colored blocks placed in various positions and collect them to a home position. Following a series of contacts with such blocks, value signals received through a conductive snout allow it to categorize and sort these blocks on the basis of differences in reflectance regardless of where the blocks were placed. This challenging task requires the successful combination of sequences of both reflex and acquired adaptive behaviors. As in Darwin III, this behavioral combination occurs as a result of selection based on value leading to perceptual categorization. When the value circuits are cut or interfered with, the behavior does not occur. The real-world approach to SNM that is embodied in Darwin IV retains all the advantages of the technique but avoids the difficulties and pitfalls faced by attempts to simulate a rich environment in addition to a brain. It provides a valuable means of testing the psychological implications of the theory of neuronal group selection, particularly those related to learning (Kandel and Schwartz, 1982).

In this brief review, I have not dealt in detail with many important subjects addressed by the theory of neuronal group selection, including aspects of memory, learning, and consciousness. The interested reader may consult several publications (Edelman, 1987, 1989, 1992) that describe how, without further assumptions, the theory can account for key features of such higher order functions of the brain. Here, my main purpose has been to stress the explanatory and predictive usefulness of the theory of neuronal group selection by addressing several fundamental phenomena that must be explained before considering those subjects. It does not seem likely that neuroscientific research can lead to a deep view of how the brain functions unless such global theories and the models that reflect them are available to bridge the experimental results obtained in a variety of disparate fields ranging from molecular biology to psychology.

Acknowledgments

The work performed at The Neurosciences Institute cited here was supported by the Neurosciences Research Foundation.

References

Ashby, W. B. (1966). *An Introduction to Cybernetics* (New York: Wiley).

- Barlow, H. B. (1988). Neuroscience: a new era? *Nature* 331, 571.
- Bernstein, N. (1967). *The Coordination and Regulation of Movements* (Oxford: Pergamon Press).
- Böhme, G. A., Bon, C., Stutzmann, J.-M., Doble, A., and Blanchard, J.-C. (1991). Possible involvement of nitric oxide in long-term potentiation. *Eur. J. Pharmacol.* 199, 379–381.
- Changeux, J.-P., and Danchin, A. (1973). Selective stabilization of developing synapses as a mechanism for the specification of neuronal networks. *Nature* 264, 705–711.
- Cowan, W. M. (1978). Aspects of neural development. *Int. Rev. Physiol.* 17, 150–191.
- Crick, F. (1989). Neural Edelmanism. *Trends Neurosci.* 12, 240–248.
- Eckhorn, R., Bauer, R., Jordan, W., Brosch, M., Kruse, W., Munk, M., and Reitboeck, H. J. (1988). Coherent oscillations: a mechanism of feature linking in the visual cortex? Multiple electrode and correlation analyses in the cat. *Biol. Cybernet.* 60, 121–130.
- Edelman, G. M. (1978). Group selection and phasic re-entrant signalling: a theory of higher brain function. In *The Mindful Brain*, G. M. Edelman and V. B. Mountcastle, eds. (Cambridge, Massachusetts: MIT Press).
- Edelman, G. M. (1987). *Neural Darwinism: The Theory of Neuronal Group Selection* (New York: Basic Books).
- Edelman, G. M. (1988). *Topobiology: An Introduction to Molecular Embryology* (New York: Basic Books).
- Edelman, G. M. (1989). *The Remembered Present: A Biological Theory of Consciousness* (New York: Basic Books).
- Edelman, G. M. (1992). *Bright Air, Brilliant Fire: On the Matter of the Mind* (New York: Basic Books).
- Edelman, G. M., and Crossin, K. L. (1991). Cell adhesion molecules: implications for a molecular histology. *Annu. Rev. Biochem.* 60, 155–190.
- Edelman, G. M., Reeke, G. N., Jr., Gall, W. E., Tononi, G., Williams, D., and Sporns, O. (1992). Synthetic neural modeling applied to a real-world artifact. *Proc. Natl. Acad. Sci. USA* 89, 7267–7271.
- Engel, A. K., König, P., Kreiter, A. K., and Singer, W. (1991). Interhemispheric synchronization of oscillatory neuronal responses in cat visual cortex. *Science* 252, 1177–1179.
- Finkel, L. H., and Edelman, G. M. (1987). Population rules for synapses in networks. In *Synaptic Function*, G. M. Edelman, W. E. Gall, and W. M. Cowan, eds. (New York: Wiley), pp. 711–757.
- Finkel, L. H., and Edelman, G. M. (1989). The integration of distributed cortical systems by reentry: a computer simulation of interactive functionally segregated visual areas. *J. Neurosci.* 9, 3188–3208.
- Gally, J. A., and Edelman, G. M. (1992). Nitric oxide: linking space and time in the brain. *Proc. Natl. Acad. Sci. USA* 89, 11651–11652.
- Gally, J. A., Montague, P. R., Reeke, G. N., Jr., and Edelman, G. M. (1990). The NO hypothesis: possible effects of a short-lived, rapidly diffusible signal in the development and function of the nervous system. *Proc. Natl. Acad. Sci. USA* 87, 3547–3551.
- Georgopoulos, A. P., Schwartz, A. B., and Kettner, R. E. (1986). Neuronal population coding of movement direction. *Science* 233, 1416–1419.
- Gerfen, C. R. (1992). The neostriatal mosaic: multiple levels of compartmental organization in the basal ganglia. *Annu. Rev. Neurosci.* 15, 285–320.
- Gray, C. M., and Singer, W. (1989). Stimulus-specific neuronal oscillations in orientation columns of cat visual cortex. *Proc. Natl. Acad. Sci. USA* 86, 1698–1702.
- Haley, J. E., Wilcox, G. L., and Chapman, P. F. (1992). The role of nitric oxide in hippocampal long-term potentiation. *Neuron* 8, 211–216.
- Hebb, D. O. (1949). *The Organization of Behavior: A Neurophysiological Theory* (New York: Wiley).
- Herrnstein, R. J. (1985). Riddles of natural categorization. *Phil. Trans. Roy. Soc. Lond. (B)* 308, 129–144.
- Jones, F. S., Prediger, E. A., Bittner, D. A., De Robertis, E. M., and Edelman, G. M. (1992a). Cell adhesion molecules as targets for *Hox* genes: neural cell adhesion molecule promoter activity is modulated by cotransfection with *Hox-2.5* and *-2.4*. *Proc. Natl. Acad. Sci. USA* 89, 2086–2090.
- Jones, F. S., Chalepakis, G., Gruss, P., and Edelman, G. M. (1992b). Activation of the cytotactin promoter by the homeobox-containing gene *Evx-1*. *Proc. Natl. Acad. Sci. USA* 89, 2091–2095.
- Kaas, J. H., Merzenich, M. M., and Killackey, H. P. (1983). The reorganization of somatosensory cortex following peripheral-nerve damage in adult and developing mammals. *Annu. Rev. Neurosci.* 6, 325–356.
- Kandel, E. R., and Schwartz, J. H. (1982). Molecular biology of learning: modulation of transmitter release. *Science* 218, 433–443.
- Köhler, W. (1947). *Gestalt Psychology* (New York: Liverwright).
- Kreiter, A. K., and Singer, W. (1992). Oscillatory neuronal response in the visual cortex of the awake macaque monkey. *Eur. J. Neurosci.* 4, 369–375.
- Landry, P., and Deschenes, M. (1981). Intracortical arborizations and receptive fields of identified ventrobasal thalamocortical afferents to the primary somatic sensory cortex in the cat. *J. Comp. Neurol.* 199, 345–371.
- Lashley, K. S. (1947). Structural variation in the nervous system in relation to behavior. *Psychol. Rev.* 54, 325–334.
- Merzenich, M. M., Kaas, J. H., Wall, J. T., Nelson, T. J., Sur, M., and Felleman, D. J. (1983). Topographic reorganization of somatosensory cortical areas 3b and 1 in adult monkeys following restricted deafferentation. *Neuroscience* 8, 33–55.
- Michod, R. E. (1989). Darwinian selection in the brain. *Evolution* 43, 694–696.
- Montague, P. R., Gally, J. A., and Edelman, G. M. (1991). Spatial signaling in the development and function of neural connections. *Cerebral Cortex* 7, 199–220.
- Northington, F. J., Matherne, G. P., and Berne, R. M. (1992). Competitive inhibition of nitric-oxide synthase prevents the cortical hyperemia associated with peripheral-nerve stimulation. *Proc. Natl. Acad. Sci. USA* 89, 6649–6652.
- O'Dell, T. J., Hawkins, R. D., Kandel, E. R., and Arancio, O. (1991). Tests of the roles of two diffusible substances in long-term potentiation: evidence for nitric oxide as a possible early retrograde messenger. *Proc. Natl. Acad. Sci. USA* 88, 1285–1289.
- Pearson, J. C., Finkel, L. H., and Edelman, G. M. (1987). Plasticity in the organization of adult cortical maps: a computer simulation based on neuronal group selection. *J. Neurosci.* 7, 4209–4223.
- Purves, D. (1988). *Body and Brain: A Trophic Theory of Neural Connections* (Cambridge, Massachusetts: Harvard University Press).
- Pylyshyn, Z. W. (1984). *Computation and Cognition* (Cambridge, Massachusetts: MIT Press).
- Rakic, P. (1988). Specification of cerebral cortical areas. *Science* 241, 170–176.
- Reeke, G. N., Jr., Finkel, L. H., Sporns, O., and Edelman, G. M. (1990). Synthetic neural modeling: a multilevel approach to the analysis of brain complexity. In *Signal and Sense: Local and Global Order in Perceptual Maps*, G. M. Edelman, W. E. Gall, and W. M. Cowan, eds. (New York: Wiley).
- Schmitt, F., Worden, F. G., Edelman, G., and Dennis, S. G., eds. (1981). *The Organization of the Cerebral Cortex* (Cambridge, Massachusetts: MIT Press).
- Schuman, E. M., and Madison, D. V. (1991). A requirement for the intercellular messenger nitric oxide in long-term potentiation. *Science* 254, 1503–1506.
- Shibuki, K., and Okada, D. (1991). Endogenous nitric oxide release required for long-term synaptic depression in the cerebellum. *Nature* 349, 326–328.
- Smith, E. E., and Medin, D. L. (1981). *Categories and Concepts* (Cambridge, Massachusetts: Harvard University Press).

Sporns, O., and Edelman, G. M. (1992). Solving Bernstein's problem: a proposal for the development of coordinated movement by selection. *Child Dev.*, in press.

Sporns, O., Gally, J. A., Reeke, G. N., Jr., and Edelman, G. M. (1989). Reentrant signaling among simulated neuronal groups leads to coherency in their oscillatory activity. *Proc. Natl. Acad. Sci. USA* *86*, 7265-7269.

Sporns, O., Tononi, G., and Edelman, G. M. (1991). Modeling perceptual grouping and figure-ground segregation by means of active reentrant connections. *Proc. Natl. Acad. Sci. USA* *88*, 129-133.

Takeichi, M., Inuzaka, H., Shimamura, K., Fujimori, T., and Nagafuchi, A. (1990). Cadherin subclasses—differential expression and their roles in neural morphogenesis. *Cold Spring Harbor Symp. Quant. Biol.* *55*, 319-325.

Tononi, G., Sporns, O., and Edelman, G. M. (1992). Reentry and the problem of integrating multiple cortical areas: simulation of dynamic integration in the visual system. *Cerebral Cortex* *2*, 310-335.

Underwood, G. V. (1978). *Strategies of Information Processing* (London: Academic Press, Inc.).

Zeki, S. (1981). The mapping of visual functions in the cerebral cortex. In *Brain Mechanisms of Sensation: The Third Taniguchi Symposium on Brain Sciences*, Y. Katsuki, R. Norgren, and M. Sato, eds. (New York: Wiley), pp. 105-128.