

BOOK REVIEW

Brains, computation, and selection: An essay review of Gerald Edelman's *Neural Darwinism*

PAUL PATTON

University of Texas at Austin, Austin, Texas

and

THOMAS PARISI

Saint Mary's College, Notre Dame, Indiana

Neural Darwinism: The theory of neuronal group selection

By Gerald M. Edelman. 1987, New York: Basic Books. 371 pp. \$29.95.

Neuroscientists will have a difficult time ignoring the publication of *Neural Darwinism*, Gerald Edelman's exposition of his theory of neuronal group selection. Lavish brochures were sent to members of such groups as the Society for Neuroscience and the American Association for the Advancement of Science, and prominent advertisements heralded the publication in periodicals such as the *New York Review of Books*. In dust-jacket blurbs, Maxwell Cowan described the book as "perhaps the most original work on the nervous system in thirty years," and John Szentagothai remarked that "this is the first occasion that a real brain theory begins to take shape." To be sure, not all the hoopla has been positive. In a recent cover story in the *New York Times Magazine*, Gunther Stent said that he tried hard, but simply could not understand what Edelman is trying to say. H. B. Barlow (1988), reviewing the book for *Nature*, asked rhetorically whether Edelman's book presages a new era in neuroscience or is simply a hopeless muddle. Barlow clearly favors the latter view.

In a time of high specialization in the neurosciences, when much attention is devoted to reductionist approaches that emphasize the most molecular details of neural function, Edelman's book is a welcome attempt at a comprehensive synthesis. He has brought together research from many disparate areas to defend the notion that the brain works by a process of random variation and selection similar to that characteristic of Darwinian evolution. Unfor-

tunately, however, the book suffers on both substantive and stylistic grounds, ultimately failing to convince the reader that a *theory* of brain function has been either developed or presented. In what follows, we offer a summary and a critical discussion of Edelman's ideas as presented in *Neural Darwinism*.

The central aim of Edelman's book is to gain an understanding of the neural bases of perceptual categorization. The ideas that Edelman puts forth on this topic are clearly rooted in his earlier work in immunology, for which he won the Nobel Prize in 1972. Edelman's immunological work elucidated the structure of antibody molecules, and in doing so helped to resolve a long-standing puzzle about the way the immune system recognizes specific antigens (for an account of this work see Rosenfield, 1988). Prior to the work of Edelman and his colleagues, an hypothesis of immune function that had been put forward by Linus Pauling in 1940 was widely accepted. This hypothesis maintained that the immune system manufactures only one type of antibody molecule, which is then altered in its conformation according to the kinds of antigens it encounters. In this view, the immune system, at least in part, is built by *instruction* from the outside world (Jerne, 1967). Edelman helped demonstrate that this "instructionist" view is wrong. Rather than one kind of molecule, shaped or instructed by incoming information, he demonstrated that enormous variability exists among the active regions of antibody molecules, and that a particular antigen *selects* a particular antibody from among this diversity. The production of a specific immune response then, involves a form of recognition or categorization. Edelman's basic thesis in *Neural Darwinism* is that the processes of perceptual categorization in which the nervous system engages are based on similar "selectionist" mechanisms.

Edelman sees current approaches to the problem of perceptual categorization, which are based on cognitive psychology and computational neuroscience, as containing an implicit instructionist assumption that the organism is

The authors would like to thank Gerald Edelman, Bob Wyatt, Vinod Menon, and Torsten Minkwitz for discussions that were useful in the preparation of this review. Address all correspondence to Thomas Parisi, Department of Psychology, Saint Mary's College, Notre Dame, IN 46556.

instructed by fixed categories present in the environment. "Processors of information," Edelman writes, "must have information defined for them *a priori*" (p. 42). What Edelman seems to be saying is that if perceptual categorization is an information-processing task, then the environment must contain a clear and unique set of categories that the brain is prearranged to interpret correctly. "At one stage or another of such models," Edelman writes, "one sees evidence of typological thinking or essentialism—the positing of prior categories either in the environment or in the brain or both" (p. 43).

Having labeled information-processing models of brain function as both essentialist and instructionist, Edelman uses evidence drawn from cognitive psychology to argue that perceptual categories are not fixed or closed, and that these models must therefore be wrong. He notes three major features of perceptual categorization that he regards as incompatible with information-processing models. The first feature is that perceptual categories are not veridical, but rather are adaptive and modifiable by context. "The very same object," he notes, "may be classified differently at different times, and an animal may use different means to classify that object at different times" (p. 30). Second, Edelman notes that perception involves the capacity to generalize category membership on the basis of only a few examples. Referring to Herrnstein's (1979, 1985) work demonstrating such an ability in pigeons, he argues that this capacity must be independent of language.

A third feature of perceptual categorization that Edelman believes to be incompatible with an information-processing view of perception is that "the classical summary description of a class defined by singly necessary and jointly sufficient conditions does not hold" (p. 31) for perceptual categories; most perceptual categories are polymorphous sets. To qualify for membership in such a set, an element must satisfy a specified number, but not all, of the rules governing membership in the category.

Edelman concludes that these three features of perception "together challenge any logic-based or 'information-driven' explanation of the data" (p. 30). He argues that the evidence clearly shows that "the categories of received signals can be defined only after the signals have been received, either because of evolutionary selection or as a result of somatic experience" (p. 43–44), and, therefore, that the necessary *a priori* code assumed by information-processing models cannot exist. Edelman clearly sees an analogy between perceptual categories and the Darwinian view of living species, in which species are regarded as changing and adaptable entities rather than as the fixed categories of the essentialist.

In Edelman's view, the structure of the nervous system is also incompatible with the perception of the brain as an information-processing system.

Although structures in a particular area of the brain are modally similar among conspecific animals, at the level of fine axonal and dendritic ramifications and connections, there

occurs a very large degree of individual variation in shape, extent, and connectivity. (p. 6)

Noting that computers and other electronic devices must be precisely wired in order to function properly, he argues that this variation "would doom any equivalent parallel computer to produce meaningless output even with the best of error correcting codes" (p. 42). Edelman believes that this variability is precisely what is needed, however, to fuel a Darwinian process of selection.

Given the need to account for adaptively varying perceptual categories and a source of variability upon which a process of selection might act in the nervous system, Edelman proceeds to lay out the specifics of his theory. In his theory, Edelman makes three specific assertions, which are, confusingly, sometimes referred to as premises and sometimes as claims. The first assertion is that the genome does not fully specify the details of the pattern of synaptic connections generated in the developing nervous system and, therefore, the details of connectivity in the nervous system are highly variant in the sense described above. As evidence for this assertion, Edelman cites neuroanatomical evidence as well as his own work on cell adhesion molecules (CAMs). He argues that CAMs guide neural morphogenesis, but lack the specificity necessary to determine fully the pattern of neural connectivity established during development. The developmental process thus produces "*primary repertoires* [italics added] of structurally variant neuronal groups" (p. 5). A neuronal group, as Edelman defines it, is "a collection of cells of similar or variant types, ranging in number from hundreds to thousands, that are closely connected in their intrinsic circuitry and whose mutual dynamic interaction may be further enhanced by increases in synaptic efficacy" (pp. 46–47). Each primary repertoire of neuronal groups is "degenerate" in that it includes "a significant number of non-identical variant groups, each of which could respond more or less well to a particular input" (p. 6).

Edelman's second assertion is that there is a competitive process of selection among neuronal groups within a primary repertoire, resulting in the selection of "combinations of those particular groups whose activities are correlated with various signals arising from adaptive behavior" (p. 5). This selection process occurs by changing synaptic efficacy with activity, as specified by pre-synaptic and postsynaptic rules. The competition can result in the growth of some neuronal groups by the capture of cells from other groups and by the shrinkage or disappearance of still others. The outcome of this process of selection and competition is the *secondary repertoire*, which consists of "functioning groups that are more likely to be used in future behavior" (p. 5). According to Edelman, the process is constant and ongoing throughout the life of the organism. As evidence for such plasticity, Edelman cites the findings of Michael Merzenich and his colleagues (Merzenich et al., 1983a, 1983b; Merzenich

et al., 1984), who have shown that the organization of the somatosensory cortex in a monkey can change following the amputation of a finger.

The third assertion of Edelman's theory is that functionally related groups within different secondary repertoires are linked by "re-entrant" connections. These connections provide for spatiotemporal coordination between interacting neuronal groups contained in different secondary repertoires. If one is observing a firecracker, for example, the loud bang and the flash of light activate different brain regions. Re-entrant connections, in Edelman's view, provide a way of linking these different aspects of an object or event into a coherent perception. Individual, electrophysiologically described maps are regarded by Edelman as being examples of secondary repertoires, and are referred to as "local maps." Re-entrant connections, both sensory and motor, between groups contained within multiple local maps, together with connections with non-mapped regions, form "global maps." Global maps embody spatiotemporally continuous representations of objects or events, and generate motor behaviors with respect to them:

Each repertoire within the local maps of a global mapping disjunctively samples various aspects or features of the environment. Connection of these local maps . . . in a global mapping serves to link samples by reentry so that various representations of features are *correlated* in space and time. None of the local maps would be adequate for perceptual categorizations or generalizations. A global mapping provides the minimal unit capable of such function. (p. 210)

According to Edelman, global mappings, and their formation, dissolution, and selective alteration with time, in turn provide the basis for learning and memory.

Before turning to a critique of the substantive aspects of Edelman's book, we must first note that the book suffers from serious stylistic difficulties which often make it nearly impossible for the reader to understand the central elements of what Edelman is trying to say. One often encounters awkward, convoluted sentences. When faced with the choice between a simple and an obtuse term, Edelman either chooses the latter or invents a new one that is even more troublesome. To cite one example, Edelman introduces the term "instructionism" in an apparently novel context without ever defining it. We were able to infer what Edelman means by this term only after discovering its usage in the immunological literature. In addition, the book is repetitive: the main points are summarized many times. More attention to organization and some sense of audience would have resulted, we would hope, in a much briefer and more readable book. It is difficult to imagine how an editor could have allowed this book to be published in its present form.

In addition to problems of style, we must also remark that the book is, at times, marred by a seeming arrogance and condescension. The arrogance comes through most clearly in the first chapter of the book, in which Edelman traces the history of ideas leading up to his own theory. Edelman appears to view past theories of brain function

as nothing more than incomplete anticipations of his own work, and theories that lack essential features of his model are dismissed for this reason alone. For example, Edelman criticizes Hebb's (1949) theory because "it neglected Darwinian interpretations" (p. 13), and he rejects both behaviorism and cognitive psychology because

there is no hint in either view . . . that perhaps the problem [of generalization and categorization] is one of adaptively matching neural variants in a degenerate repertoire to external situations involving disjunct sets of features or properties, thus allowing generalization to occur through processes similar to those by which taxa are created during evolution. (p. 13)

Edelman's view of the history of theories of brain function and psychology is especially ironic in light of his emphasis on perception as adaptive and historically contingent rather than veridical.

Edelman often dismisses the views of others with disparaging remarks rather than with a cogent defense of his dismissal. For example, in discussing Ebbesson's theory of parcellation in development, Edelman dismisses the theory at least twice because it is "recapitulationist": "this ontogenetic explanation smacks too much of recapitulationist views and is much too narrow" (p. 155). Why is it recapitulationist? In what way is it too narrow? Edelman feels no need to explain.

Let us now turn to a consideration of the substantive issues raised by Edelman's book. We will begin with a critique of Edelman's metatheoretical assertions regarding the value of computational approaches to brain function, and will then turn to an analysis of the specific assertions of his theory of neuronal group selection.

The motive behind Edelman's theory, it seems to us, is a dissatisfaction with the idea of the computer as a metaphor for brain function. In recent decades, this metaphor has exerted a profound influence on all disciplines that deal with the brain, behavior, and cognition (Gardner, 1985). Edelman, however, believes that the problems of accounting for adaptively varying perceptual categories, perceptual generalization, the polymorphous character of perceptual categories, and the occurrence of reliable brain function in the presence of anatomical variance are so serious as to warrant the abandonment of the computer metaphor altogether and replacing it with his selectionism. Edelman tells us that he seeks "to account for [perceptual] categorization without assuming information processing or computing" (p. 4). "It is surprising," he writes, "to observe that neurobiologists . . . can . . . believe that precise algorithms are implemented and that computations and calculations of invariance can take place inside neural structures" (p. 42). "The selective behavior of ensembles or neurons," he states, "may be *describable* by certain mathematical functions . . . but it seems as unlikely that a collection of neurons carries out the computation of an algorithm as that interacting lions and antelopes compute Lotka-Volterra equations" (p. 235). Unfortunately, as we will see, Edelman's views regarding the computer metaphor are full of inconsistencies and

contradictions that are likely to thoroughly confuse even the most careful reader.

Edelman's specific point of attack is the computationalist approach to brain function developed by the late David Marr of the Massachusetts Institute of Technology in his book *Vision* (1982). "The present proposal," Edelman writes, "differs strongly from the computationalism of Marr: it rejects the idea that the nervous system computes a function" (p. 235). Marr's central concern in *Vision* was to explain how the brain generates an internal representation of a persisting three-dimensional world of surfaces and shapes from the shifting array of light intensities and wavelengths that impinge on the two-dimensional surfaces of the retinas of the eyes. Marr's definition of the problem of vision in this fashion, and his methods of approaching the problem, were novel. Although Edelman charges that in Marr's work "categories of natural objects in the physical world are implicitly assumed to fall into defined classes or typologies that are accessible to a program" (p. 38), in fact Marr did not address the problem of object categorization or recognition. "This theory," he wrote, "has nothing to say about semantic recognition, object naming, or function. . . . I think that an understanding of what we mean by the semantics of an object are fascinating, but I also think that they are very difficult indeed and at present much less accessible" (p. 355).

Marr's computationalism consists of a set of particular theoretical models of visual function and, more fundamentally, of a set of metatheoretical assertions about the way such problems should be addressed. Marr began his career as an electrophysiologist, but eventually realized that if brain function is to be understood, something more is needed than simply a description of the behavior and the properties of neurons in particular regions of the brain.

Trying to understand perception by studying only neurons is like trying to understand bird flight by studying only feathers. It just cannot be done. In order to understand bird flight, we have to understand aerodynamics; only then do the structure of the feathers and the different shapes of birds' wings make sense. (Marr, 1982, p. 27)

Marr rejected the reductionist notion that complex brain functions can be explained in terms of neurons alone, and instead proposed a multileveled explanatory scheme with neural implementation as only one level. He placed particular importance on the top level of his scheme, that of computational theory.

Just as a theory of aerodynamics is essential for understanding bird flight, Marr maintained, so must one develop an abstract theory of how a particular information-processing task can be performed in principle if one wishes to understand how the brain performs that task. This computational theory must, by its nature, be mathematical. One example, which Marr studied in great detail, is the task of binocular stereopsis. To develop a theory of the task of stereopsis, one must deal in the trigonometry of binocular parallax. One must also have mathematical rules for matching corresponding points in the two

eyes' images. The mathematics of a valid computational theory of binocular stereopsis must, in some sense, be embodied in any system that successfully solves the problem of binocular stereopsis (whether by a straightforward route or by an awkwardly convoluted one), whether that system is a computer or a brain.

Edelman is not alone in criticizing Marr's computational theories of vision. Patricia Kitcher (1988) has noted that Marr's theories contain an implicit assumption that computations are carried out by the brain in the most optimal way. Citing the arguments of Gould (1980) and of Gould and Lewontin (1979), she points out that such an assumption is unlikely to be valid in a system produced by an historical process of random mutation and natural selection without an intelligent designer. Paul Grobstein (in press-a) noted that throughout much of *Vision* Marr seemed to advocate computational theory over direct investigation of the nervous system. Grobstein has stressed the importance of direct investigations in the identification and definition of the computational tasks performed by the nervous system, and in the discovery of new modes of neural information processing that may not have been imagined by computer scientists. He has suggested that computational investigations can best be guided by nervous-system studies involving brain lesioning and behavioral analysis. Unlike Edelman, however, neither of these critics has advocated the abandonment of the computational level of investigation.

Let us now consider whether Edelman has been successful in consistently rejecting computationalism. What does it mean to say that the brain computes an algorithm? An algorithm is defined as a set of unambiguous rules for solving a problem or for performing a procedure in a finite number of steps (Knuth, 1969). Like computers, brains can solve complex problems within a finite period of time, and there must be some orderly process by which they do so. At its most basic level (what Marr would call the level of implementation), an electronic computer is nothing more than a machine that gates electrical impulses in a complex pattern. On this level, a computer no more computes than do the lions and antelopes of Edelman's example. To understand the orderly process that is going on in the computer, and to develop new computer programs, however, computer scientists must make use of algorithms and computational theories. Computationalism asserts that it is also valuable to make use of algorithms and computational theories to describe the tasks the brain performs and the orderly process by which it performs them. Computationalism no more requires that an algorithm somehow exist objectively inside the brain than it requires that an algorithm exist objectively inside the computer. It only requires that the tasks that the brain performs be *describable* by a computational theory, and that the methods that the brain uses to perform them be describable using algorithms (Sejnowski, Koch, & Churchland, 1988).

Can the solutions to problems such as perceptual categorization be stated in a finite number of unambiguous steps? Currently, computers are generally used to

solve problems that possess a clear logical structure and a single correct solution. An example of such a problem is the computation of the logarithm of a number. For such problems, a relatively compact algorithm can be stated that permits the computation of a solution in only a few steps. However, Edelman doubts that the problem of perceptual categorization or object recognition is one of this sort. The world, he comments, "is initially an unlabeled place" (p. 3), by which he means that it contains no prior logical categorical structure upon which a succinct algorithm for perceptual categorization might be based.

Many of Edelman's central concerns about perceptual categorization have also been raised by some computationalists, who believe that neural pattern recognition may involve the solution of problems that lack a clear logical structure (Abu-Mostafa & Psaltis, 1987). These investigators distinguish between two different sorts of computational problems. The first sort of problems are structured problems, like those discussed above, that can be solved using a succinct algorithm. The second sort of problems, known as random problems, lack sufficient logical structure to permit a simple algorithmic solution. Abu-Mostafa and Psaltis, like Edelman, regard the problem of object recognition as one of this sort. They do not, however, see this as reason for abandoning the notion that such problems can be solved by a process that involves a finite number of unambiguous steps. They propose instead that the algorithm for solving such a problem should consist of a stored list of all possible solutions to the problem along with rules for choosing among them.

Consider, for example, the problem of recognizing a picture of a tree. One may attempt to solve this problem as a structured problem by using a brief list of stored rules that distinguish a tree from all other objects. This is clearly the sort of approach to which Edelman objects when he denies that the world possesses a prior categorical structure. Indeed, computer models of object recognition based on such an approach have had little success. Another approach would be to store the images of a large number of trees in the computer, along with rules for comparing any input image with the stored images. Abu-Mostafa and Psaltis (1987) have reported some success in using such a strategy for pattern recognition in experiments aimed at developing pattern-recognizing parallel computers using optical technology. Their strategy closely parallels that employed in Edelman's computer model of perceptual generalization: Darwin I.

Abu-Mostafa and Psaltis (1987) have pointed out that if the problem of recognizing a picture of a tree is treated as a random problem, the need for any prior knowledge of trees is eliminated. The programmer of a computer designed to solve random problems (which Abu-Mostafa and Psaltis refer to as a "neural" computer)

does not have to understand in a formal mathematical sense the problem for which he or she is programming. The programmer only has to provide enough "training" data (consisting of possible solutions) to the computer and allow it to set up a unique pattern of connections for each solu-

tion. In other words, it would be possible for a neural computer to program itself. (Abu-Mostafa & Psaltis, 1987, p. 93)

This is precisely what Edelman is claiming for the brain in his theory of neuronal group selection. Edelman's concerns about the limitations of algorithms seem meaningful only if one accepts a far more restrictive definition of an algorithm than that adopted by Abu-Mostafa and Psaltis. However, the usefulness of such a restriction is unclear.

Given Edelman's stated aim of accounting for perceptual categorization "without assuming information processing or computing," most readers of *Neural Darwinism* will be surprised to discover that Edelman presents a number of computer models that illustrate various aspects of his theory of perceptual categorization. In chapter 3, for example, Edelman presents a computer model, called Darwin I, that illustrates some aspects of his theory of perceptual categorization. This program is presented as a model for several aspects of perceptual categorization considered in the abstract, without direct reference to the nervous system. It appears to us that, without knowing it, Edelman is attempting to formulate a computational theory for the task of perceptual categorization. We have the same impression of Edelman's model of pattern recognition, Darwin II. Here, Edelman tells us that pattern categorization can be accomplished by means of two neural networks, linked as a classification couple. One of the networks serves as a feature detector and the other as a feature correlator. The assertion that coupled networks that possess these characteristics are sufficient for pattern recognition seems to us to be an abstract statement about the task of pattern recognition rather than a statement about the brain itself. In other words, it seems to be an assertion concerning the computational theory of the task of pattern recognition. Perhaps there is some way this apparent contradiction can be justified in terms of computer theory. If there is, Edelman makes no serious attempt to present it. In the absence of such an explanation, few readers are likely to find his rejection of computationalism convincing. What does seem clear is that if aspects of Edelman's theory can be embodied in computer models, then those aspects must be storable in terms of algorithms, or the necessary computer programs cannot be written. Edelman thus does not seem to have escaped successfully from the assumption that the brain computes in an algorithmically describable fashion.

Although Edelman does not consistently seem to have rejected the computer metaphor in any deep sense, his approach to brain function clearly differs in important ways from that of traditional artificial intelligence or computational neuroscience. Marr and other computationalists have concerned themselves with the question of how the brain performs its functions and have devised algorithms and computational theories that incorporate knowledge about such functions. Edelman instead asks how the brain arrives at the ability to perform its functions without prior knowledge. We do not currently know of any way that

algorithms or computer programs can be generated without a computer programmer. Because Marr and other computationalists have not addressed the problem of where the brain gets its algorithms, Edelman charges that these investigators implicitly assume a homunculus in the brain to serve as its programmer.

Although Edelman's statement on the homunculus problem does not, in our view, invalidate the computationalist program, it does highlight an area of major importance that computational neuroscience has so far left largely unexplored. Edelman's answer to the homunculus problem is that orderly, adaptive function in the course of Darwinian evolution must arise from the bottom up by natural selection and by a process of neuronal group selection acting during the lifespan of a single organism. The basis for generating this order, in Edelman's view, is the presence of small-scale anatomical variability in the nervous system, sensory input from the environment, and a set of rules governing the modification of synaptic strengths with activity. It should be noted that these rules, unlike the algorithms generally postulated by computational neuroscientists, contain no information about the task that the network is to perform. Therefore, this sort of model is known as a noninstructed neural-network model. Edelman claims that such a network, if properly organized, can arrive at adaptive function by means of a competitive process of neuronal group selection that arises as a result of the operation of the synaptic efficacy rules. At its deepest level, such a network is a computational procedure like any other, with rules governing the modification of synaptic efficacy with activity. The operation of these rules, however, makes the system capable of self-organization in response to activity, and thus makes it capable, in Edelman's view, of generating adaptive behaviors without built-in instructions for those behaviors. For many neurobiologists, a model of this general sort will hold great appeal. Grobstein (1988b, 1988c), for example, has argued cogently that the phenomenon of adaptive plasticity in brain and behavior is of central importance to any account of behavioral brain function. The merit of Edelman's particular development of this idea, however, depends on the specifics of his proposal. Let us now turn to an examination of these specifics as they are presented in *Neural Darwinism*.

The first major assertion of Edelman's theory of neuronal group selection is that the genome does not fully determine the details of neural connectivity. This indeterminism, together with the tendency of axonal terminal arbors to overlap broadly, provides a source of random variability upon which a selective process may act. The opposing view, that the nervous system is hardwired by the genome, is generally associated with the chemoaffinity hypothesis, commonly attributed to Roger Sperry (see discussion in Grobstein, 1988c).

Edelman seems to regard the mere existence of anatomical variability in the nervous system as evidence for genetic indeterminism. He notes, for example, that in repeated neural structures no two cells within a given structure exhibit an arbor with precisely the same shape,

and that no two individual brains exhibit an identical anatomy. He also points out that cortical sensory maps, as mapped electrophysiologically, show an extraordinary degree of variability from individual to individual. However, this evidence is hardly compelling. The fact that variability exists in repeated structures tells us nothing about whether the source of that variability is genetic or epigenetic. Since different individuals have different genomes, individual differences in cortical maps might be due to genetic, rather than epigenetic, influences.

As additional evidence against complete genetic specification, Edelman presents a detailed review of the role of CAMs in development and of studies of the formation of cortical maps. He argues that the picture of development as presented by modern developmental biology suggests that the development process lacks sufficient specificity to allow the genome to determine neural connectivity fully. However, an assertion that the *known* mechanisms governing development lack the specificity to determine neural connectivity fully is not in itself compelling evidence that such specificity could not exist.

On a theoretical level, Changeux and Danchin (1976) have argued that the genome cannot contain sufficient information to specify neural connectivity fully. In a recent review, Grobstein (1988c) cited a considerable body of evidence, derived from studies of growth, metamorphosis, and sensory experience, that cell-cell recognition processes are not by themselves sufficient to determine network connectivity uniquely, and argued that additional epigenetic influences are needed. Given the strength of these theoretical and empirical arguments, it seems reasonable to grant Edelman's assertion that the genome does not fully determine the pattern of neural connectivity. However, such a conclusion cannot be reached on the strength of the evidence that Edelman himself emphasizes.

Edelman reviews some evidence that is suggestive of a role for competitive and selective interactions during the formation of what he calls the primary repertoire. He notes the widespread occurrence of cell death during development, along with evidence that patterns of cell death are governed by the success of individual cells in forming appropriate connections rather than being genetically preprogrammed. He also notes evidence of competition in synapse formation at neuromuscular junctions.

Edelman devotes a much greater effort to the development of selectionist models related to the second major assertion of his theory of neuronal group selection: the formation of a secondary repertoire of neuronal groups by selection from the primary repertoire. Here, however, we encounter problems that have already been noted by Barlow in his review of *Neural Darwinism* (Barlow, 1988). A major problem is that Edelman's definition of a neuronal group is confusing and imprecise. At one point, he defines a cortical neuronal group as "an ensemble of cohesively interconnected cells, all of which express the same receptive field" (pp. 164-165). In the visual cortex, as Barlow has noted, there is no experimental evidence to support the claim that nearby cells have identical receptive fields. When questioned on this matter by

one of the present authors, Edelman replied that he meant that the cells express similar receptive fields. Edelman's carelessness in describing such crucial aspects of his theory is a major barrier to understanding and accepting it.

We also wonder, as did Barlow (1988), what functional advantage could be conferred by having a set of cells with a similar or identical function that grows to incorporate more cells if it is frequently activated? Much of Edelman's discussion of the formation of the secondary repertoire is directed specifically toward accounting for the plasticity of topographic somatosensory maps reported by Merzenich and his colleagues (Merzenich et al., 1983a, 1983b; Merzenich et al., 1984), rather than dealing with the question of the functional significance and origin of the varied and often complex receptive-field properties found in the cells that form cortical maps.

In Edelman's theory, the appearance, selection, and competitive growth of neuronal groups is an outcome of the operation of pre- and postsynaptic rules that act independently to alter the efficacy of pre- and postsynaptic terminals. The novel feature of these rules is that changes in the efficacy of a synaptic terminal can be influenced by activity at other synaptic terminals of the same cell. In the postsynaptic rule, changes in the efficacy of a postsynaptic terminal occur when its activation closely corresponds in time to that of other synaptic terminals of the same cell. The influences of other terminals are conveyed to the terminal in question either by electronic spread or by diffusible messenger molecules. Edelman's presynaptic rule states that the efficacy of presynaptic terminals is altered based on the long-term average of the instantaneous presynaptic efficacy of the cell. As elsewhere in his book, this vital discussion is obscured by stylistic problems. We were not able to determine precisely what governs changes in the instantaneous presynaptic efficacy of the cell in Edelman's model.

Although Edelman's discussion of the formation of secondary repertoires seems geared primarily towards explaining the plasticity of topographic representations reported in experiments such as those of Merzenich and his colleagues (Merzenich et al. 1983a, 1983b; Merzenich et al. 1984), broader functional considerations are more directly addressed in the third assertion of Edelman's theory of neuronal group selection. This assertion states that individual secondary repertoires are linked into global mappings by reentrant connections, and that these global mappings form the minimal units of perceptual categorizations. Edelman's major concern in proposing reentrant connections is that different properties of a particular object or event in the sensory environment are represented in separate, spatially distinct, topographic maps. These maps are Edelman's answer to the question of how all these separately registered, but simultaneously occurring, properties are linked together and placed in a single category. Global mappings also link sensory structures to motor structures. Edelman believes that perceptual categorization is an active process that requires coordi-

nated motor acts to acquire the information necessary to form perceptual categories.

Edelman bolsters these claims by presenting a computer model of perceptual categorization known as Darwin II. This model, which uses printed characters as its input stimuli, consists of a digital computer simulation of two interconnected neural networks known as "Darwin" and "Wallace." The Darwin network is a feature-detector network. The first level of the network consists of an array of detectors sensitive to line segments of various orientations or with particular bends. The Wallace network is a feature-correlator network, the first level of which responds to continuity properties of objects. The individual elements of these networks represent neuronal groups, each network corresponds to a set of secondary repertoires, and the system as a whole corresponds to a global mapping. Edelman presents evidence that this model exhibits a limited capability to form categories and to generalize category membership using letters of the alphabet as stimuli.

Edelman believes that global mappings are sufficient to account for perceptual categorization, and that they are the foundations for higher order brain functions such as learning and memory. More generally, he asserts that peripheral sensory-receptor systems "abstract major adaptive properties and that the function of higher order neural networks is not to 'compute' but to correlate these abstracted properties by further reentrant selection among selected populations of participating neuronal groups" (p. 234). Perception, however, involves more than simply categorizing groups of sensory stimuli. Studies of many animals, for example, have provided evidence that they are capable of extracting information about the distribution and three-dimensional structure of shapes and surfaces in their environment from the patterns of light falling on their retinas (Collett & Harkness, 1982; Marr, 1982). It may be that, for some animals, the generation of such an internal representation of their three-dimensional environment is a necessary first step or a valuable aid to perceptual categorization. The question of how this representation is formed was the central concern of Marr's computational investigation of vision. Marr showed that to generate an internal representation of this sort, the brain must solve computational problems of considerable difficulty. Edelman has little to say as to whether his global mappings can solve these problems, and if so, how.

There are many instances in which animals must deal with the geometry of their world, and the evidence suggests that the brain mechanisms that have evolved to carry out these tasks are complex indeed. Orientation of the head or body with respect to a stimulus, for example, was once thought to be a simple matter of establishing a fixed correlation between a locus on a sensory map corresponding to the direction of the stimulus, and a locus on a motor map capable of triggering an appropriately directed movement. The study of oculomotor orientation in the mon-

key (Mays & Sparks, 1980; Sparks & Mays, 1983) and, in particular, of visually evoked prey orienting movements in the frog (Grobstein, 1988a, in press-b; Kostyk & Grobstein, 1987), have demonstrated that the control of such movements involves computations of far greater complexity than the simple correlation of sensory and motor maps. In the case of the frog, there is evidence that generating an orienting turn involves an intermediate processing level in which signals are coded in neither a retinal- nor a movement-coordinate frame, but rather in some more generalized frame of reference (Grobstein, 1988a). Edelman's theory deals with how the brain categorizes sensory stimuli and motor acts, but he has very little to say about how the brain deals with the geometrical aspects of either perception or movement. In neglecting to consider such matters explicitly, he risks building a global theory of perception and action that fails to capture essential elements of both.

A driving force behind Edelman's book is his insistence on rejecting the computational approach to brain function developed by Marr and others. Although he supposedly rejects computationalism, Edelman seems, paradoxically, to be groping towards a computational theory of perceptual categorization. His argument that selectionism can replace computationalism seems to us as misplaced as an argument that evolutionary theory could replace physiology. Edelman's presentation is so confused and confusing that few readers will be convinced that a true theory of brain function has been presented. Many readers may nevertheless find aspects of the basic thesis of *Neural Darwinism* to be highly intriguing. The need to provide an explanation of the neural basis for the ability of an individual organism to adapt its behavior to its environment, as well as the need to account for adaptive plasticity in the nervous system, is widely recognized. Darwinian evolution is the only process in nature that we know of that produces adaptive order without intelligent design. It thus seems natural to suggest that a similar process, operating within the brain of an individual organism, might adapt the perceptions and actions of that organism to its environment. Edelman's book serves the valuable function of raising these issues for a broad scientific audience for the first time. Paradoxically, it seems likely that these ideas will ultimately add legitimacy to the very metaphors of computationalism that Edelman is trying to undo, by suggesting ways by which brains, or computers, can arrive at adaptive function.

REFERENCES

- ABU-MOSTAFA, Y. S., & PSALTIS, D. (1987). Optical neural computers. *Scientific American*, **256**, 88-95.
- BARLOW, H. B. (1988). Neuroscience: A new era? *Nature*, **331**, 571.
- CHANGEUX, J. P., & DANCHIN, A. (1976). Selective stabilization of developing synapses as a mechanism for the specification of neuronal networks. *Nature*, **267**, 705-711.
- COLLETT, T. S., & HARKNESS, L. I. K. (1982). Depth vision in animals. In D. J. Ingle, M. A. Goodale, & R. J. W. Mansfield (Eds.), *Analysis of visual behavior* (pp. 111-176). Cambridge, MA: MIT Press.
- GARDNER, H. (1985). *The mind's new science: A history of the cognitive revolution*. New York: Basic Books.
- GOULD, S. J. (1980). *The panda's thumb*. New York: Norton.
- GOULD, S. J., & LEWONTIN, R. (1979). The Spandrels of San Marco and the panglossian paradigm: A critique of the adaptationist program. *Proceedings of the Royal Society of London*, **B205**, 581-598.
- GROBSTEIN, P. (1988a). Between the retinotectal projection and directed movement: Topography of a sensorimotor interface. *Brain, Behavior, & Evolution*, **31**, 34-48.
- GROBSTEIN, P. (1988b). From the head to the heart: Some thoughts on similarities between brain function and morphogenesis, and on their significance for research methodology and biological theory. *Experientia*, **44**, 960-971.
- GROBSTEIN, P. (1988c). On beyond neuronal specificity: Problems in going from cells to networks and from networks to behavior. In P. G. Shinkman (Ed.), *Advances in neural and behavioral development: Vol. 3* (pp. 1-58). New York: Albex.
- GROBSTEIN, P. (in press-a). Strategies for analyzing complex organization in the nervous system: 1. Lesion experiments: The old re-discovered. In E. Schwartz (Ed.), *Systems development foundation benchmark series in computational neuroscience*. Cambridge, MA: MIT Press.
- GROBSTEIN, P. (in press-b). Strategies for analyzing complex organization in the nervous system: 2. A case study: Directed movement and spatial representation in the frog. In E. Schwartz (Ed.), *Systems development foundation benchmark series in computational neuroscience*. Cambridge, MA: MIT Press.
- HEBB, D. O. (1949). *The organization of behavior: A neuropsychological theory*. New York: Wiley.
- HERRNSTEIN, R. J. (1979). Acquisition, generalization and discrimination of a natural concept. *Journal of Experimental Psychology & Animal Behavior Processes*, **5**, 116-129.
- HERRNSTEIN, R. J. (1985). Riddles of natural categorization. *Philosophical Transactions of the Royal Society of London [Biology]*, **308**, 129-144.
- JERNE, N. K. (1967). Antibodies and learning: Selection versus instruction. In G. C. Quarton, T. Melnechuk, & F. O. Schmitt (Eds.), *The neurosciences: A study program* (pp. 200-205). New York: Rockefeller University Press.
- KITCHER, P. (1988). Marr's computational theory of vision. *Philosophy of Science*, **55**, 1-24.
- KNUTH, D. E. (1969). *The art of computer programming: Vol. 1. Fundamental algorithms*. Reading, MA: Addison-Wesley.
- KOSTYK, S., & GROBSTEIN, P. (1987). Neuronal organization underlying visually elicited prey orienting in the frog: 1. Effects of various unilateral lesions. *Neuroscience*, **21**, 41-55.
- MARR, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. New York: Freeman.
- MAYS, L. E., & SPARKS, D. O. (1980). Saccades are spatially, not retinocentrically coded. *Science*, **208**, 1163-1165.
- MERZENICH, M. M., KAAS, J. H., WALL, J. T., NELSON, R. J., SUR, M., & FELLEMAN, D. J. (1983a). Progression of change following median nerve section in the cortical representation of the hand in areas 3b and 1 in adult owl and squirrel monkeys. *Neuroscience*, **10**, 639-665.
- MERZENICH, M. M., KAAS, J. H., WALL, J. T., NELSON, R. J., SUR, M., & FELLEMAN, D. J. (1983b). Topographical reorganization of somatosensory cortical areas 3b and 1 in adult monkeys following restricted deafferentation. *Neuroscience*, **8**, 33-55.
- MERZENICH, M. M., NELSON, R. J., STRYKER, M. P., CYNADER, M., SCHOPPMAN, A., & ZOOK, J. M. (1984). Somatosensory cortical map changes following digit amputation in adult monkeys. *Journal of Comparative Neurology*, **224**, 591-605.
- ROSENFELD, I. (1988). *The invention of memory: A new view of the brain*. New York: Basic Books.
- SEJNOWSKI, T. J., KOCH, C., & CHURCHLAND, P. S. (1988). Computational neuroscience. *Science*, **241**, 1299-1306.
- SPARKS, D. L., & MAYS, L. E. (1983). Spatial localization of saccade targets: 1. Compensation for stimulus-induced perturbations in eye position. *Journal of Neurophysiology*, **49**, 64-74.