The Neural Basis of Cognitive Development:

A Constructivist Manifesto

Steven R. Quartz[1] and Terrence J. Sejnowksi[2]

submitted to: Behavioral & Brain Sciences

[1]Computational Neurobiology Laboratory, and
The Sloan Center for Theoretical Neurobiology,
The Salk Institute for Biological Studies,
PO Box 85800
San Diego, CA 92186-5800
steve@salk.edu
(to whom correspondence should be addressed)

[2]Howard Hughes Medical Institute
The Salk Institute for Biological Studies, and
Department of Biology,
University of California, San Diego

Abstract

The neural structures and mechanisms underlying development undergo major changes during the acquisition of cognitive skills. How can these changes be characterized, and to what extent do they alter the learning properties of a developing system? By examining the neural basis for cognitive development, we suggest that cognitive and neural processes fundamentally interact during development. Through considering possible neural measures of representational complexity across development---synaptic numbers, axonal arborization, and dendritic arborization---we arrive at a constructivist view of development in which environmentally-derived activity plays a central role in regulating neuronal growth as a progressive increase in the representational capacities of cortex. This suggests a view of learning in development as a process of dynamic interaction between the informational structure of the environment and neuronal growth mechanisms, allowing the representational properties of cortex to be constructed by the nature of the problem domain confronting it. As we examine, this is a uniquely powerful and general learning strategy that undermines the central assumptions of classical learnability theory, which is premised on the assumption that the learning properties of a system can be deduced from a fixed computational architecture. By regarding the representational properties of cortex to be the outcome of a dynamic interaction between a structured environment and neuronal growth mechanisms, the constructivist view thus minimizes the need for prespecification, in agreement with recent neurobiological evidence indicating that developing cortex is largely equipotential and free of domain-specific structure. This suggests an evolutionary perspective in which cortical evolution should be viewed as the progression to more flexible representational structures, in contrast to the popular view of cortical evolution as an increase in specialized circuits that are innate.

1 Introduction

A fundamental question in the study of cognition is how neural mechanisms participate in, or underlie, cognitive development---what is the nature and extent of the interaction between cognitive and neural processes during development, and what are the consequences of this interaction for theories of acquisition? One of the most intriguing possibilities is that the structures underlying development are not fixed, but instead undergo changes during the course of the acquisition of cognitive skills. A further possibility is that these changes may not simply be intrinsically determined, but may depend on the nature of the interaction with the environment. In this case, the distinction between cognitive and neural development would be blurred since principles at both levels of explanation would be required in an integrated level of explanation. More specifically, a consequence of the interaction between cognitive and neural processes is that learning episodes may induce changes to the structures involved in learning, which would alter what could be learned during subsequent episodes. A system with such an ability is said to possess the principle of nonstationarity (e.g., Pinker 1979). This principle concerns what may be termed the ‘‘learning mechanism’’ or ‘‘acquisition device’’, roughly, the set of processes and structures that transform some set of input data into the mature state representing the mature competence. Nonstationarity makes the learning properties of such a system time-dependent and suggests that such factors as the order of presentation of information will have a significant effect on the acquisition properties of the system.

As a number of researchers have noted, the inclusion of nonstationarity, while increasing the learning capacities of a system, likewise introduces a number of special problems into an account of development; indeed, these problems are sometimes seen as so complicating that there have been a number of attempts to rule out nonstationarity on methodological grounds. For example, this tendency was illustrated in Chomsky’s (1980) idealization that development may be viewed as an instantaneous process without undercutting any legitimacy of the claims regarding language acquisition. Similarly, Pinker (1984), following Macnamara (1982), formulated the ‘‘continuity hypothesis," that children and adults should be viewed as qualitatively identical until forced to abandon this principle by some explanatory failure. This view thus contributed to the lack of interest in characterizing the neural substrates of acquisition.

We present evidence that leads us to question these assumptions. The resurgence of neural network research has placed learning and the creation of internal representations once again in the fore of cognitive science. Likewise, recent progress in cognitive neuroscience has made the claim that neurobiological evidence is irrelevant to theories of cognitive development untenable. Combined, these advances have led to the central assumption of cognitive and computational neuroscience: that meaningful and far-ranging connections can be made between cognitive and neural levels of description, and, furthermore, it is only through the mutual constraints that both levels impose on each other that a satisfactory theory can emerge. In this paper, therefore, we examine how nonstationarity in development can be characterized from within a neurocomputational perspective and the implications of this characterization for theories of development.

As we explore in detail, the principle of nonstationarity as applied to the learning mechanism results in the view of development as a process involving representational change. Since the locus of this change is the learning mechanism, it is important to characterize representational change during development in terms that correspond to structural changes in this mechanism. For this reason, we suggest that in order to understand nonstationarity in development it is necessary to characterize the following four elements: 1) changes to the neural substrates underlying acquisition, 2) the processes regulating these changes, 3) their correspondence to representational change, and 4) their implications for the learning properties of the developing system. These four elements suggest a methodology for understanding cognitive development that is centered around its neural basis. Recent research in developmental neurobiology makes it possible to begin to characterize the learning mechanism in these structural terms, which we suggest provides for a highly-constrained basis for assessing changes to that mechanism in terms of representational change. Hence, a first step toward characterizing the neural basis of cognitive development is to provide an appropriate measure of representational complexity in structural terms. This is one of our primary aims in this paper. Specifically, we first consider three candidate measures for changes in representational complexity in structural terms: changes in synaptic numbers, axonal arborization, and dendritic arborization. We examine the neurobiological evidence concerning the nature of changes in these measures during development, the involvement of intrinsic and extrinsic factors in regulating these processes, and the adequacy of each of these measures as an indicator of representational complexity.

We shall suggest that when these issues are considered in light of recent neurobiological research the best supported view of development is one as a progressive increase in representational capacity in which this increase depends fundamentally on interaction with a structured environment. On the basis of these two features, we will formulate a constructivist view of the neural basis of cognitive development. As its namesake implies, this model is sympathetic to the Piagetian tenet that development essentially involves an active interaction between the developing system and the structure of the environment in which the system is embedded. Also like the Piagetian program, the model we present emphasizes the constructive nature of this interaction, that is, one in which representational structures are progressively added during the course of development. Here, though, our primary concern will be to examine the nature of the processes involved in regulating these structural changes and their implications for representational change during development. Specifically, we shall suggest that dendritic development fulfills a number of important conditions for a nonstationary learning mechanism and shall propose a number of schemes for how dendritic development under the influence of environmentally-derived activity corresponds to cognitive principles for the construction of the representations that underlie the mature state.

In contrast to the view we present is the popular view of development as essentially a regressive phenomena. According to this view, development is seen as involving two discrete stages. The first involves an intrinsic construction of ‘‘prerepresentations’’ through both genetic and epigenetic means. The second stage involves the selective elimination of those prerepresentations through competitive stabilization mechanisms whereby a subset of that original structure is picked to underlie mature competences. Variants of this view have been defended by cognitive psychologists (e.g., Mehler 1985), psycholinguists (e.g., Lightfoot 1991; 1993), so-called selectionist brain theorists (e.g., Jerne 1967; Changeux and Danchin 1976) and those applying it to cognitive development (e.g. Piatelli-Palmarini 1989). Even what may be called the ‘‘standard view" of development in neurobiology, stemming from Hubel and Wiesel’s work on the visual system, emphasizes the notion of development as the selective elimination of exuberant initial growth. This has led to the common claim in the psychological literature that human development is characterized by an early synaptic overproduction followed by massive elimination (e.g., Siegler 1989; Pinker 1989). In evaluating the evidence for this view of development, we will conclude that these regressive models are largely unsupported by the bulk of recent neurobiological data and that regressive events as underlying cognitive development have been overstated.

In later sections, we extend this treatment of the neural basis of cognitive development to examine the learning-theoretic properties of the constructivist model, along with some alternatives. The best known characterization of the learning properties of a developing system stems from questions in language acquisition---what properties of a language a child could learn, what in the environment could serve as evidence for that learning, and ultimately, what must be prespecified by the child's genetic endowment. From these questions, 30 years of research has provided mainly negative results: the environment is impoverished and could provide only a limited source of information so that few interesting properties of a language could be learned. In the end, it was concluded that the child must bring most linguistic knowledge, in the form of a *universal grammar*, to the problem. Learning, on the other hand, could only set a few free parameters in that universal grammar to adapt to the specifics of the particular language confronting the child.

This striking view of language acquisition underlying the now general view of the mind as a collection of innately specified, specialized modules (e.g., Fodor 1983; Hirschfeld and Gelman 1994), while also based on numerous empirical studies of the nature of linguistic input (e.g., Brown 1973), would not be so compelling were it not for the widely-held assumption that it is also based on rigorous results in formal learning theory (FLT). According to FLT, development is essentially a learning problem and so is constrained by the learning-theoretic pressures that confront any learner. By attempting to characterize the general features of any learning problem, the proponents of this approach asked, how hard is the problem of deriving the mature cognitive state underlying a skill from the initial state on the basis of the evidence present in some environment (Gold 1967; Pinker 1979; Wexler & Culicover 1980; Osherson et al. 1986)? We shall suggest that a fundamental limitation of this theory is that it views the problem of acquisition in development solely as a response to learning-theoretic pressures. The extent to which the biological details of the developing system may influence the system's learning properties has been ignored and, again, the prospect of a nonstationary learning mechanism has been precluded. By incorporating changes to the learning mechanism according to the constructivist view we present, we shall argue that the role of cognitive processes of learning in regulating structural growth is such a central feature of human development that it undermines many of the well-known claims of learnability and results in a much richer view of learning in development than the traditional view allows. From this perspective, we shall thus argue that the learnability conclusions of classical cognitive science and the resulting view of development are fundamentally mistaken. As Pinker (1894, p.371) himself notes, the results of ''rigorous" learning theories demonstrating that complex skills cannot be learned ''assume that the nature of cognitive processes can be deduced from properties of a fixed computational architecture.'' The appeal of the constructivist position lies precisely in its rejection of this assumption and hence in possessing fundamentally different learning properties (Quartz 1993).

In recent years, Chomsky's nativist view has been championed by evolutionary psychologists (e.g., Barkow et al. 1992). Spurred on by the view of development as an intrinsic process, they have sought a phylogenetic source for the intrinsically specified skills comprising the human cognitive repertoire. Whereas Chomsky remained uncommitted to the claim that complex mental structures evolved on the basis of selective pressures, this has become the banner of evolutionary psychology. Currently, it is popular to regard not only cognitive skills but also social and political attitudes as the expression of genetically encoded responses to phylogenetic pressures acting over the course of evolutionary history. Although the notion of specialized circuits is sometimes viewed from within a selectionist framework (e.g. Gazzaniga 1992) the selectionist model is strictly incompatible with the claim that specialized circuits have been selected by evolutionary pressures. The primary reason for this is that according to selectionism (e.g., Edelman, 1987) selective pressures operate on epigenetic variation during the ontogeny of the individual (in ''somatic'' time), not on a phylogenetic timescale. Pinker (1994) restates Roger Sperry's chemoaffinity hypothesis that address-encoding surface markers on individual cells underlies the connectivity of specialized circuits, but this is an extreme view of neural specificity that neurobiologists have not held for twenty-five years (see Easter et al. 1985 for a review). As we shall conclude in section 6, a major challenge to this view is to identify plausible mechanisms to underlie its developmental claims.

The constructivist view is in fundamental contrast to this account of cognitive skills as the expression of genetically encoded responses to selective pressures acting over the course of evolutionary history. We suggest that the view of human cognitive evolution as the progressive increase in specialized cortical structures reflects a basic misinterpretation of cortical evolution. In contrast, we suggest that the hallmark of cortical evolution is not the ever-increasing sophistication of dedicated or specialized cortical circuitry (Gazzaniga 1995); rather, it is marked by an increasing representational flexibility that allows environmental factors to fundamentally shape the structure and function of the human brain. In later sections, we suggest this is a uniquely powerful response to the central learning-theoretic pressure of determining the appropriate representation or hypothesis space for some particular domain. In addition, it minimizes the need for a prespecification for these sources, in contrast to the heavy burden nativism places on genetic mechanisms. Instead, the constructivist strategy transfers this burden to the informational structure of the environment, thereby viewing this structure as much richer than the ''poverty of the stimulus'' arguments of nativism allow, a richness of structure that neural networks have begun to uncover. This model thus makes the relation between changes in the environment, whether through natural means or by cultural innovation, and brain structure a direct one, and does not demand that these be mediated by novel specialized structures with an intrinsic specification.

We should note that this view is not a return to *Tabula Rasa* learning in that we view learning as a dynamic interaction between a structured environment and neural mechanisms that can both be modulated in fundamental ways by activity deriving from the informational structure of the environment but is also constrained by intrinsic properties. However, we view this interaction as sufficient to determine the mature representational properties of cortex without the need to postulate domain-specific predispositions that are somehow embedded *a priori* in the recipient cortex. The traditional dichotomy between a general-purpose, fixed computational architecture with generic learning procedures versus domain-specific knowledge embedded *a priori* in cognitive structures does not adequately transfer to the neural network perspective. According to our view, then, the process of the self-organization of cortex is a dynamic interaction between very general intrinsic processes and afferent activity carrying domain-specific information derived from environmental structure, an interaction that is sufficient to construct the mature computational architecture and its representational properties.

This paper is organized as follows. First, in section 2, we examine candidate neural measures of representational complexity in order to provide a structural basis for assessing the impact of nonstationarity in development. In section 3, we then consider dendritic development in these terms, a characterization that we extend in section 4 with a discussion of principles of constructivist growth. In section 5, we turn to a consideration of a learning-theoretic approach to development, an approach that has been influential in cognitive science. There we examine the learning properties of the constructivist view and suggest that it escapes the negative learnability arguments of classical learnability theory. In section 7, we present some general implications of the constructivist view.

2 Measures of representational complexity

A first step toward understanding the neural basis of cognitive development is to constrain changes to the learning mechanism in structural terms. As we briefly indicated above, there are two major alternative views of nonstationarity in development, the selectionist and the constructivist. The selectionist learning mechanism is one whereby development involves a reduction in circuit complexity through selective mechanisms operating on redundant structures. On the other hand, the constructivist view we propose regards development as consisting of a directed construction of representations through interaction with a structured environment, and so as involving a progressive increase in representational capacity. These two positions fundamentally differ over the view of development as changes in representational or circuit complexity and the mechanisms responsible for these changes. These differences thus place a premium on finding an appropriate measure of circuit or representational complexity. As we also mentioned above, there are three main candidates for this measure: synaptic numbers, axonal arborization, and dendritic arborization. Although these are all related, it is worthwhile to treat them separately, as a change in one does not necessarily imply a change in the others. In the remainder of this section, we consider both the empirical support for changes in these measures along with their adequacy as measures of representational capacity. Again, our main aim in so doing is to determine the appropriate structural measures that will serve as the basis for understanding these changes in cognitive terms as representational change.

2.1 Manipulation of activity: Effects on development

The notion that development involves changes in representational capacity is closely tied to the question of the nature of the mechanisms that underlie such changes. According to the constructivist model, a general feature of development involves a progressive increase in representational capacity. What makes progressive growth interesting from a theoretical perspective is that it raises the possibility that this growth is activity-dependent and therefore sensitive to the structure of the animal's environment. In this case, development would be directed in the sense that cortical structural development would be responsive to environmental demands. Activity resulting from those demands would effectively modulate the complexity of neural structure in those regions, building up circuits in an experience-dependent manner. This process would ensure a match between the informational structure of the input space and the structures which sample that space. In slightly more cognitive terms, this growth would correspond to allowing the structure of the input to determine the architecture that is to underlie the representational properties of the region (Quartz 1993). This suggests that the question of the activity-dependence of development is as central as the issue of whether there is a net addition of structure. The standard and selectionist models do not postulate directed sampling mechanisms in this sense but only allow for the functional validation of structure whose construction is insensitive to the structure of the environment. In contrast, as we just noted, the constructivist model incorporates mechanisms that make the initial structure itself sensitive to the activity of a system, and thereby sensitive to the structure of the environment. Hence, alongside our examination of candidate measures of representational complexity, we will consider the evidence for this mode of development.

2.2 Synaptic numbers

2.2.1 Synaptic numbers over development

We first consider changes in synapse numbers as a measure of changes in representational capacity. A widely-held assumption in neuroscience is that synapse numbers are the appropriate measure of circuit complexity. For example, studies of structural plasticity following learning typically subscribe to the ''more is better" doctrine (e.g., Turner & Greenough 1985). In development, a number of studies have quantitatively examined changes in synaptic numbers, despite the technical difficulties such studies present, both in obtaining unbiased counts (rev. in Coggeshall 1992) and in the amount of labor demanded by the techniques involved, such as serial reconstruction in electron microscopy (Huttenlocher et al. 1982; O'Kusky and Collonnier

1982a; 1982b; Rakic et al. 1986; Huttenlocher & de Courten 1987; Bourgeois et al. 1989; Zecevic & Rakic 1991).

Among these, the Rakic et al. (1986) study has been the most influential and is often cited as support for a selectionist model of development. In that study, synaptic density was examined in motor, somatosensory, prefrontal, visual, and limbic cortices of the rhesus monkey. A striking finding was that synaptogenesis proceeded uniformly across all regions examined, with a peak between two and four months of age, followed immediately by a dramatic decline. This concurrent synaptogenesis was taken to indicate that a single intrinsic signal regulated the process of cortical synaptogenesis. A limitation of this study, however, was that synaptic density was related only to general characteristics of the neuropil and could not be correlated with particular cell types. If particular cell types matured at different rates, these differences might become apparent only by directly examining their development. An example of such a study is that of Lund et al. (1990a). They found that layer 4c alpha and beta stellate cells, cells in visual cortex in adjacent sublamina, matured at different rates, reaching various landmarks almost 4 weeks apart. This indicated that even nearby cells do not necessarily follow a similar, intrinsic timetable of development. What could then account for the differing maturation? Although these cell types are nearby one another, they receive input from distinct pathways, the magno- and parvocellular. Under various deprivation conditions that minimized differences in the activity of these pathways, these two cell types showed a more synchronous development (Lund et al. 1990b). As we return to in detail in section 2.2.2, this suggests that differences in afferent drive, and not just intrinsic mechanisms, may play an important role in regulating cellular maturation.

If concurrent synaptogenesis does not seem to hold as a general property, then how has the claim of initial synaptic overproduction followed by selective elimination held up? A recent study by Rakic and colleagues (Bourgeois et al. 1994) on synaptic development in the prefrontal cortex of macaque monkeys has forced a reinterpretation of this model. Specifically, with a larger sample they did not find an early synaptic overproduction followed immediately by elimination, as in their earlier studies. Instead, they found that synaptic density reached a peak around two months of age and did not begin to decline until puberty. Even then, the decline was a gradual one lasting until 20 years of age (their last sample point). In such a case it is difficult to reconcile this finding with the claim that selective elimination underlies cognitive development since no such process appears to be operative during the most formative years of cognitive development. This is made more problematic by evidence for volume increases during this period, particularly in prefrontal cortical areas (see Dekaban 1978; Jernigan et al. 1991 and references therein).

Analogous studies in humans have played a central role in forming a view of human brain development that has been popular in psychology (e.g., Siegler 1989; Pinker 1990). As in Rakic et al. (1986) study, this view regards human brain development as comprised of an initial overproduction of synaptic numbers to one or two years of age followed by the selective elimination of this overproduction. The actual data on human cortical development, however, is scarce. In fact, the only published quantitative study of synaptic measures in developing human non-visual cortex is Huttenlocher (1979) (Figure 1).

Although the Huttenlocher study is an important one, and more basic human neuroanatomical work is an important goal, it is worthwhile to point out some limitations in this study. Because of the difficulty in obtaining such samples, the Huttenlocher (1979) study involves a total of 21 brains (with data only for layer III) from subjects ranging in age from newborn to 91 years of age (with no gender distinctions made). The major limitation, though, in attempting to infer general timetables of human prefrontal development lies in the fact that this study lacks samples between the ages of 1--5 and 8--15. This alone severely undermines any attempt to reach a conclusion as the ages of interest are not represented by the data sample. From what data there is, however, there are no significant differences among synaptic density measures between one year and eight years of age. Whether any decreases become evident between 8 and 15 is unclear as there are no data points for this range. Also, the data from prefrontal cortex displays a significantly different timecourse from that in visual cortex, suggesting that concurrent synaptogenesis does not hold between these regions (Huttenlocher 1979; 1990) .

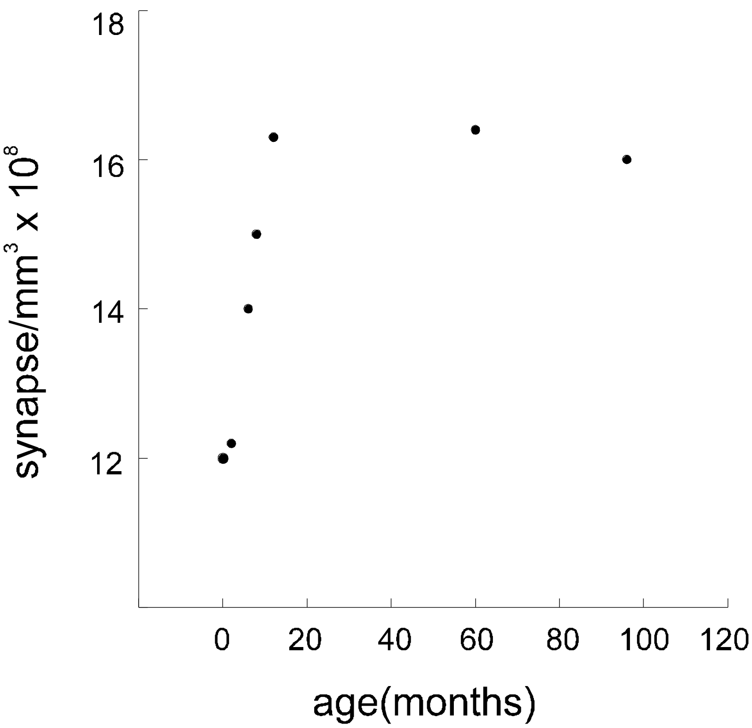

Figure 1. Synaptic density in layer III, human middle frontal gyrus from birth to 8 years. Data from Huttenlocher (1979). Note the absence of data points between 1 and 5 years of age.

The study of Bourgeois et al. (1994) mentioned above helps to clarify some of the confusion surrounding claims of human cortical development as a regressive process (see also Rakic et al. 1994). Specifically, by matching their data with conceptual ages in the human data, Bourgeois and colleagues suggest that synaptic development in human prefrontal cortex does not undergo any significant reduction prior to the onset of puberty. Again, this makes the claim that synaptic elimination underlies the acquisition of cognitive skills (Changeux & Danchin 1976; Edelman 1987) increasingly problematic. The revised view of synaptic development, therefore, does not agree with the notion that selective elimination of exuberant structure underlies cognitive development, as the popular selectionist position supposes. Instead, to the degree that there is a reduction in synaptic density estimates, these take place well after the formative years of cognitive development. Also, studies in orhter areas that have obtained volume estimates, and thereby could estimate the total number of synapses, do not find evidence for declines in synapse numbers (e.g., Moore & Bernstein 1989; Pomeroy et al. 1990). We conclude, then, that while these results do not imply that synaptic overproduction is not a real phenomena, its generality has been overstated and the identification of cognitive development with synaptic elimination in cerebral cortex is dubious. In addition, as we consider in more detail below, the intrinsic source of structure that the selectionist model posits is likewise overstated.

2.2.2 Environmental conditions and synaptic numbers

As we noted in the previous section, the study of Lund et al. (1990b) suggested that the differences between the maturation rate of two nearby cell types might be due to differences in afferent drive. This raises the question of whether activity-dependent mechanisms might be involved in the establishment of neuronal structure. The most direct source of evidence regarding the degree of activity-dependence in the establishment of neuronal structure comes from differential rearing/environment studies. These studies are important sources of evidence for distinguishing between competing models of development since they can distinguish between activity-dependent and independent modes of growth throughout the lifespan and, in some cases, can correlate specific structural changes that correspond to a single manipulated parameter (e.g., Tieman & Hirsch 1982). As we describe, analogous findings in mature animals rule out the possibility that these changes reflect only increased or decreased stabilization of present structure.

A number of early studies examined the effects of differential rearing conditions on spine numbers and density, and thus the relation between the complexity of an animal's environment and neuronal structure (reviewed in Boothe et al. 1986). These studies, summarized in Table 1, found systematic alterations in these measures that mirrored the change in the animal's environment. Of particular interest was the finding of Valverde (1971), who demonstrated that these effects were largely reversible.

This paradigm has also been used to examine whether mature forms of plasticity involve changes in synapse number, also summarized in Table 1. One important study, that of Black et al. (1990), has recently demonstrated synaptogenesis to accompany acquisition of motor learning tasks in behaving animals. To distinguish between the possibility that motor activity alone caused the increases in synapse number, a group of animals underwent exercise on tasks that were easily mastered and required little learning. In these cases, there were no significant increases in cortical synapses number but there was angiogenesis (increased capillary support). In contrast, rats which learned complex locomotor tasks had a 25% increase in the number of synapses per Purkinje cell in the cerebellum. These studies strongly indicate that synaptic development appears to have an important activity-dependent component, which remains as a lifetime capacity. This is likely another example of a conserved biological mechanism, originally serving a development role, that was then utilized in mature forms of plasticity.

2.2.3 Synapse number as a measure of circuit complexity

In addition to the difficulties we have considered above regarding the claim of a general role for activity-independent synaptogenesis followed by synapse elimination, there is a further and perhaps more telling difficulty in the inference from a decrease in synaptic numbers to a decrease in circuit complexity. This is, what is the relation between the number of synapses of a cell or circuit and its computational or representational power? In order to hold that there is a direct relation between these two measures, a number of functional assumptions regarding the relation between synaptic numbers and the signal integration properties of cells must be posited. In connectionist models, where each input and weight effectively serves as a parameter or degree of freedom, there is a direct relation between numbers and complexity. This idealization assumes that the cell is essentially a point neuron and that the spatial properties of integration and possible nonlinear conductance properties of dendrites must be disregarded. In essence this assumes that there is a monotonic relation between synaptic numbers and signal integration. As is well known, the addition of active conductance properties suggests that the spatial arrangement of pre- and postsynaptic elements is more crucial to the determination of the response properties of a cell than absolute synaptic numbers, so the assumption is almost

Study	System	Remarks
Valverde, 1967; 1968	mouse visual cortex, stellate cells	decrease in spine density, numbers in dark rearing
Globus and Scheibel, 1967	rabbit visual cortex	visual deprivation resulted in deformed spine morphology
Valverde, 1971	mouse visual cortex, layer V pyramidal cells	mice raised in darkness for 20 days attain normal spine values when returned to normal environment
Globus et al., 1973	rat visual cortex	increase in spine density, numbers in complex environments
Cragg, 1975	cat visual cortex	bilateral suture or deafferentation, 30% reduction in the number of synapses per neuron.
Lee et al., 1980	CA1, hippocampus	increased synapse numbers with long-term potentiation
Winfield, 1981	cat visual cortex	bilateral eye suture slows rate of development and decreases synapses per neuron by 32%
Chang and Greenough, 1984	CA1, hippocampus	increased synapse numbers with long-term potentiation, but not non-LTP inducing stimulation
Turner and Greenough, 1985	rat visual cortex	used electron microscopy to quantify synapse/neuron; highest numbers in complex environments, lowest in isolated environment
Bailey and Chen, 1988a, b	Aplysia	sensitization led to 50% increase in synapse/neuron; habituation led to a 35% decrease
Black et al., 1990	rat cerebellum	motor learning led to 25% increase in synapses/neuron whereas motor activity alone caused no increase
Chang et al., 1991	CA1, hippocampus	long-term potentiation increased synaptic numbers in aged (two year old) rats

Table 1. Representative studies demonstrate the effects of rearing and manipulations to activity on synaptic numbers. See text for details.

certainly false. Nonlinear dendritic conductance properties, now well established (Stuart & Sakmann 1994), thus shift the emphasis from absolute synaptic numbers to the spatial arrangement of synapses and the branching patterns of postsynaptic structures. This suggests that axonal and dendritic arborization may be more central determinants of circuit complexity than absolute synaptic numbers. Hence, we consider these two measures in turn below.

2.3 Axonal arborization

Since it is easier to measure than synaptic numbers, patterns of axonal arborization have been widely employed as a measure of representational complexity. Indeed, studies of axonal arborization have largely contributed to what we are calling the standard model of development, as we consider next.

23.1 Axonal arborization over development

Central to modern neuroscience is the electrophysiological, developmental, and anatomical study of early vision. In particular, the retinogeniculocortical pathway, and especially the development of ocular dominance (OD) columns of layer IV, area 17 (Wiesel & Hubel 1962; 1963; 1965; Shatz & Stryker 1978; LeVay et al. 1980) have been important sources for the standard model of development. Ocular dominance columns are alternating, anatomically defined regions of input from one primarily eye. According to the standard view in the case of ocular dominance column formation, geniculocortical axons initially arborize in overlapping patterns. During development, activity segregates these diffuse projections through selective stabilization and elimination into interdigitating regions.

Since the standard model, and the model of OD formation in particular, has been so influential, especially for computational models (Miller et al. 1989), it is important to review the evidence for it. As Stryker notes (1991), almost all of the evidence for this standard model comes from autoradiography by transneuronal transport of ^{3}H-sugars or amino acids (Hubel & Wiesel 1972; Shatz et al. 1977). In this technique, a radiolabeled sample injected into the ipsilateral eye reveals an initially uniform pattern throughout the tangential extent. As development continues, a patchy pattern becomes evident, indicating the segregation of input. Along with some technical limitations of the technique, autoradiography has a limited resolution, limiting the ability to identify cellular components. Because of these limitations, and since the claim is about axonal arbor extent, it would be desirable to have anatomical evidence for this model that is based on more direct observations, such as intracellularly filled axons. This sort of evidence would document the extent of single axonal arbors and could be used in conjunction with physiological studies.

Rather than there being a large sample of identified cells, a single Golgi stained cell at early and late development (LeVay & Stryker 1979) is typically cited among the literature (e.g., Miller 1990); Antonini and Stryker (1993, p.3570) state that this cell may actually be atypical). A reason for the lack of intracellularly filled cells is the technical difficulty of filling these cells at an early age due to their small axonal diameter, thereby making direct studies prohibitive. Recently, however, Anderson et al. (1992) have managed to intracellularly fill a small sample of physiologically characterized, layer IV cells in developing kitten. The fact that these cells have been physiologically characterized is an important one, as it allows for the determination of cell type. Unlike the interpretation based on autoradiography, Anderson et al. (1992) found no evidence for initial exuberance of X-pathway axons, but instead found that the area of axonal arbor fell on the lower range of adult arbor areas, consistent with their more detailed studies in area 18 (see Table 2).

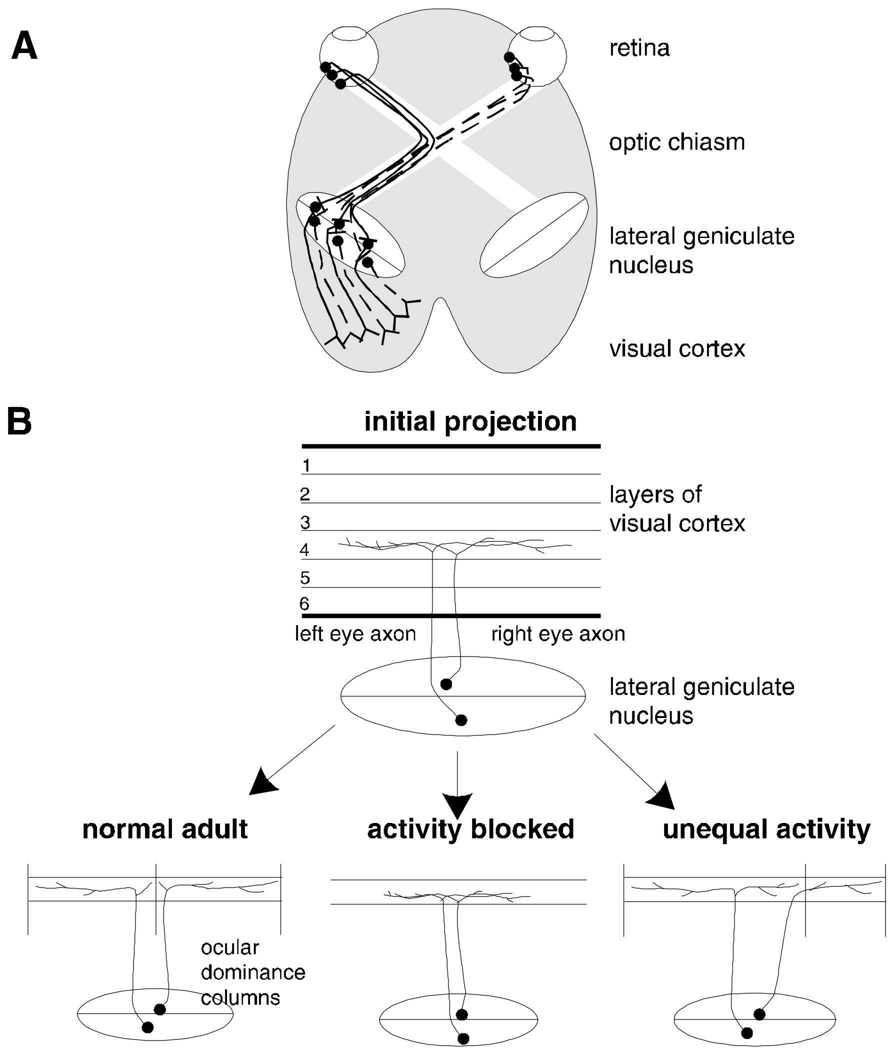

Figure 2. The generation of specificity according to the standard model. A, schematic representation of the connections between retinal ganglion cells and the lateral geniculate nucleus of the thalamus and connections to primary visual cortex. B, the two phased process of development and perturbations to that development. An initial projection shows a wide overlap in axonal arbors from both eyes, followed by activity-dependent segregation in normal development or a failure to segregate in the case of blockade of activity. In the case of unequal activity, as in monocular deprivation, the eye with the most activity is stabilized preferentially. (modified from Goodman & Shatz 1993).

Recently, the standard model has come under further revision from two of its earlier proponents. Specifically, Antonini and Stryker (1993) examined axonal development in primary visual cortex of the cat through a technique known as bulk-filling axons. Although this is not as precise as intracellular fills, and does not allow for a physiological characterization of identified cells, it does allow for visualization at earlier ages where intracellular filling is not feasible. While they found some retraction of early axonal structure, they also found substantial progressive increases in axonal measures throughout development. Strikingly, during the first week of segregation between left and right eye input, the number of branch points of an axon increase from a median value of 34.5 to 96, while total arbor length increased from 7538 to 12848 microns. Similarly, coverage area of the arbor also increased during this time, an increase which becomes significant by postnatal day 39, which is still about half the coverage area of a mature axonal arbor (Humphrey et al. 1985). Antonini and Stryker conclude (1993, p.3572) that, ''[N] development thus appears to involve both selective elimination of widely extended branches and considerable growth and elaboration.'' It is indeed difficult to reconcile these findings with a strictly regressive model.

Study	System	Remarks
Sur et al., 1982	cat retinogeniculate axons, Y-pathway	monocular deprivation caused a failure to grow of Y- pathway axonal arbors, whereas X-cells expanded
Friedlander et al., 1985	cat retinogeniculate axons, Y-pathway	progressive expansion of terminal arbors
Kalil et al., 1986	cat retinogeniculate axons, Y-pathway	elimination of action potentials blocks structural development
Friedlander and Martin, 1989	cat Y-pathway, area 18	progressive expansion of terminal arbors
Callaway and Katz, 1991	cat horizontal connections, layer 2/3 area 17	progressive axonal growth, particularly at distal segments
Friedlander et al., 1991	cat Y-pathway, area 18	monocular deprivation caused a failure to grow
Callaway and Katz, 1992	cat layer 4 spiny neurons	progressive axonal expansion
McCasland et al., 1992	rodent somatosensory cortex	decline in outgrowth of intracortical axons following deafferentation
Anderson et al., 1992	cat X-pathway, area 17	arbor area falls on the lower limit of adult measures
Antonini and Stryker, 1993	cat, X and Y-pathway, area 17	expansion of arbor area and branch numbers

Table 2. Representative studies of progressive axonal expansion and the effects of manipulations of activity. See text for details.

These recent results suggest that the standard model has overemphasized the regressive nature of development even within the very model system on which the theory was based, the OD columns. It is, therefore, worthwhile to consider the development of other systems to determine the support for this view, some of which are summarized in Table 2.

2.4 Environmental conditions and axonal development

Some studies indicating an activity-dependent component of axonal development are also summarized in Table 2. Among these, the Friedlander et al. (1991) study was an important one, as it examined the effects of monocular deprivation on Y-type geniculocortical axons. According to the standard account, monocular deprivation causes a shift in OD column formation through an expansion of the non-deprived eye with an accompanying decrease in the extent of the deprived-eye column. This shift is explained by the ability of the non-deprived eye to favorably compete for territory that is initially shared by overlapping arbors. Hence, more of the non-deprived eye arbor is stabilized whereas the deprived eye atrophies due to the lack of stabilizing input. However, Friedlander et al. (1991) found that the deprived arbors showed a failure to grow rather than a retraction due to lack of stabilization. In this case, the nondeprived arbors were not simply able to stabilize over more territory in the absence of competing fibers from the other eye; rather, their growth was either accelerated or extended, resulting in a larger final terminal field. Friedlander et al. (1991) also found a number of cellular differences between the two groups of arbors that suggest that the shift seen in autoradiographic studies might be exaggerated. Hence, depending on the pattern of incorporation of a radiolabelled tracer, it is possible that the non-deprived arbor took up more tracer and, because of the relatively low resolution of autoradiography, obscured deprived arbors (see Friedlander et al. 1991, p.3285).

2.5 Conclusions

The inadequacy of synaptic numbers to accurately measure circuit complexity has shifted the emphasis to the geometrical arrangement of pre- and post-synaptic elements. As we have just examined, while there is no doubt that regressive events occur in axonal development, their role as the dominant feature of cortical development has been overstated. We suggest that regressive events should be viewed as the more general consequence of any theory which rejects a strong view of neural specificity (Sperry 1963). Any such view must posit stochastic sampling mechanisms, which by their very nature introduce error into a system. A constructive view allows for these sampling mechanisms to be directed, though probabilistic in nature, and so will likewise require an error-correction process, such as structural elimination. But, this is not to suppose that these error-correction processes are the only operative developmental mechanisms or that development involves a selection among intrinsically-generated structures. Rather, regressive events are only one process in a *dynamic interaction* between environmentally-derived activity and the neural growth mechanisms that activity regulates.

3 Constructivist Development

In the previous sections, we have suggested that there is good evidence for the view that development involves a progressive expansion and that constructive and regressive events co-occur as directed sampling mechanisms, as appears to be emerging from new techniques that allow the same structures to be examined successively over time (O'Rourke et al. 1994; O'Rourke & Fraser 1986; 1990; Purves & Hadley 1985; Purves et al. 1986; 1987). In this section, we turn to dendritic development. Dendritic development has typically been overlooked in theoretical models of development, where axonal development tends to dominate. We shall suggest that dendrites also play an important role in determining the representational properties of cortex and that their development supports a constructivist model. First, we consider some general reasons why dendritic development is important for the representational properties of cortex.

Constructivist models incorporate the principle of nonstationarity, a principle that in the theory of automata refers to a system's ability to make changes to its underlying set of mechanisms. Although the inclusion of nonstationarity may increase the learning capacities of a system, it is problematic to implement in practice, however, as the addition of changes to the underlying architecture can be unwieldy. That is, even small changes to the architecture can potentially have large consequences for the system's overall performance. One way this may happen is if changes to an architecture undo previous learning by reconfiguring structural elements that represented already-acquired knowledge (a phenomenon sometimes referred to as catastrophic interference). Such a process could also have negative consequences for learning if it introduced large-scale changes rather than incremental ones. For example, large-scale changes could make learning too sensitive to small details of training (resulting in overfitting) and would undo the dependency on previous states that makes learning incremental, and thereby stable.

These concerns lead to the following two related conditions that must be satisfied by a nonstationary mechanism in order to be useful:

1 . *The Locality Condition* the addition of structure must be at the appropriately local scale and must not result in wholesale changes in representation with each new elemental change;

2 . *the Stability Condition* Under normal circumstances, local changes must not undo previous learning.

3.1 Dendrites as learning structures

Dendritic growth satisfies the locality condition in that, unlike the retraction of an axon which might alter the connectivity of thousands of cells, dendritic segments are local processing elements, whose elimination will not have large-scale changes in the overall pattern of connectivity. Hence, as we consider below, the constructivist learning scheme we characterize operates at the level of individual dendritic segments, rather than the whole cell, and thus builds in this locality condition.

The stability condition is satisfied by the fact that for the most part dendrites grow only at their distal tips and thus at any stage of its development only the outer regions of a dendrite will be susceptible to structural modifications (Uylings et al. 1978; Jacobs et al. 1993). Conversely, more proximal dendritic regions will have already made the transition from labile to stable structures, and will therefore be relatively insensitive to subsequent alterations. Below, this will be seen again in the case of cortical structures underlying higher cognitive functions in humans, in which distal dendritic segments display more growth and variability in response to cognitive demands. The basic idea behind this, then, is that every portion of a dendrite will have a limited temporal window in which activity-dependent processes can effect its structure. Although there are instances in which dendrites that have already undergone development can be altered (e.g., Mooney et al. 1992), the signal that initiates this large-scale modification is itself a large-scale event, such as deafferentation, that requires a change at that scale.

Another important fact is that dendrites grow much more slowly that do axons. An approximate value for axon growth is around 500 microns/day for axons versus 15-35 microns/day for dendrites (see Uylings et al. 1990; Katz et al. 1984). This suggests that these two structures are sensitive to statistical structure at different timescales and that dendrites are extracting this structure on a much longer temporal scale.

A more general reason for examining the growth of dendrites in relation to the construction of representations lies in the obvious fact that dendrites are the primary receptive surface of a neuron, and so the information-processing properties of cortex depend centrally on dendritic morphology. Further, we shall suggest that as a consequence of nonlinear processing and cellular signaling mechanisms, the spatial configuration of dendrites is a central determinant of their information processing properties. For these reasons, dendrites are a central element in the integrative signaling properties of cells, and to a large extent their structure determines their information processing properties. It is essential, then, to assess the developmental processes that lead to the determination of dendritic form and the role that such processes may play in the determination of the representational properties of neural circuits. Although we are emphasizing dendritic development, aspects of axonal development also satisfy these conditions; for example, growth at the distal portion of axons is also a local event. As it is in the interaction between dendrites and axons that the structure of the mature system emerges, we will incorporate both these mechanisms in the constructivist schemes we explore in section 4.

3.2 Dendritic arborization

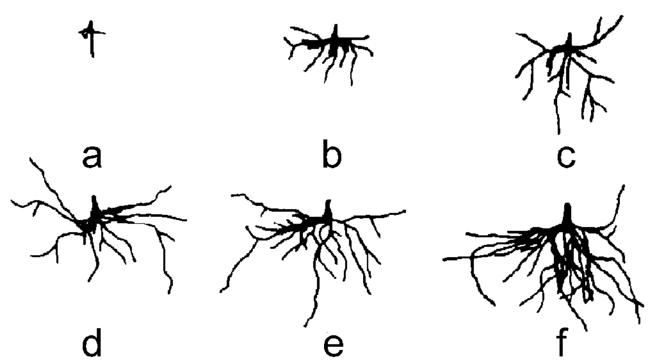

Figure 3. Camera lucida drawings of basal dendrites of layer V human pyramidal cells; a, newborn; b, 3 months; c, 6 months; d, 15 months; e, 24 months; f, adult (from Schade & van Groenigen 1961).

The extent to which constructivist principles operate during human development are illustrated by the following classical result from human neuroanatomy (Schade and van Groenigen 1961) as illustrated in Figure 3. The basal dendrites of a typical cortical cell of layer III in frontal cortex of a newborn have a total length of just 200 microns (1/5 mm). Between birth and six months of age, the dendrites of this cell will expand over ten times and come to have a total length of over 2000 microns. By two years of age, the dendrites of this cell will add another 1000 microns. Perhaps most startling, even by two years of age, the dendrites of this cell have not yet reached the halfway point in their development, but will more than double between two years and adulthood to have a mature total length of nearly 7000 microns. The dendrites of this cell, then, grow to over thirty times their length at birth and undergo the majority of this growth after two years of age. The average density of synapses on cortical pyramidal neurons is around 1 per micron.

Study	System	Remarks
Mathers, 1979	rabbit visual cortex, layer V pyramidal cells	postnatal expansion of dendritic arbor and length
Juraska and Fifkova, 1979	rat visual cortex	progressive dendritic expansion of pyramidal cells, layers II-III, V
Simonds and Scheibel, 1989	Broca's area	progressive dendritic expansion into the sixth year
Parnavelas and Uylings, 1980	rat visual cortex, layer IV stellate cells	progressive dendritic development lasting to postnatal day 20, followed by distal expansion to postnatal day 90
Buell and Coleman, 1981	human parahippocampal gyrus	increased branching and length (+35%) in normal aging, but not in senile dementia; first demonstration of plasticity in mature human brain
Becker et al., 1984	human visual cortex, layer III and V	layer III did not reach mature values until 2 years of age, followed by a non-significant decline to 7 years; layer V apical dendrites twofold progressive expansion; basal dendrites showed a progressive increase to 18 months with a slight decrease to 2 years; after 2 years, they expanded back to values at 18 months
Petit et al., 1988;	sensorimotor cortex	postnatal expansion into adulthood from about 300 μm total length to 1600.
Lubke and Albus, 1989	cat striate cortex; 150 intracellularly filled layer VI pyramidal cells	prolonged postnatal threefold expansion into adulthood of basal dendrites; from about 450 to 1300 μm total dendritic length

Table 3. Representative studies of progressive dendritic expansion. See text for details.

Table 3 summarizes some further studies of progressive dendritic expansion. While there is little doubt that regressive events also occur in dendritic development (e.g., Koester & O'Leary 1992; Vercelli et al. 1992) we take the above examples to be an important motivation for examining in more detail the processes involved in regulating dendritic growth. In the following sections, we examine the mode of dendritic development in some detail---the extent to which progressive processes may be seen to operate at the level of dendritic structure and their malleability by changes in afferent input. From these sources of evidence, we then formulate some principles underlying this development, consider their cellular basis, and relate these principles back to the learning and representational properties of cortex.

3.3 Environmental conditions and dendritic development

Unlike axonal growth, which in many cases begins during migration (Shoukimas and Hinds, 1978), dendrites typically do not begin to differentiate prior to the completion of migration and final placement within a cortical layer (Noback & Purpura 1961). This suggests that the cellular environment may be a particularly important factor in determining dendritic form, as studies of genetically altered animals have demonstrated (Rakic & Sidman 1973; Mariani et al. 1977; Caviness & Rakic 1978; Pinto-Lord & Caviness 1979).

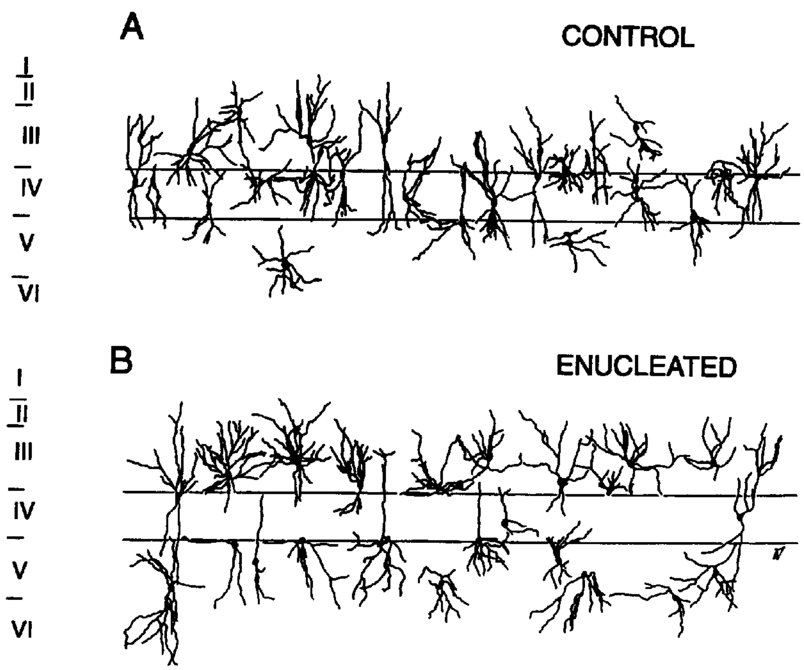

Figure 4. Dendritic organization in visual cortex of normal mice (A) and enucleated mice (B). The degenerative afferent termination in evident in B, where layer IV is sparsely covered with dendrites, whereas adjacent layers are more heavily covered, suggesting that these dendrites have reorganized according to remaining patterns of afferents (from Valverde 1968

Table 4 summarizes some studies on the effect of manipulations of afferent termination on dendritic development. One of the earliest such studies is that of Valverde (1968), which examined the effects of enucleation on the dendrites of stellate cells in mouse visual cortex. As Figure 4a illustrates, in normal development cells outside of layer IV, the location of the primary thalamic projection, project dendrites throughout layers III, IV, and V. In contrast, as figure 4b illustrates, cells outside of layer IV in enucleated animals did not project to that layer, but directed their axons toward layers III and V as though they were looking for afferent input outside layer IV. Valverde also found a significant reduction in the number of dendritic spines in layer IV, again indicating the decrease in dendrites in that layer. Valverde (1968 p.290) concludes, ‘‘dendrites are not passive structures but actively growing neuronal formations

Study	System	Remarks
Valverde, 1968	mouse visual cortex, stellate cells	enucleation resulted in dendrites redirected toward remaining afferents
Ruiz-Marcos and Valverde, 1970	mouse visual cortex	enucleation resulted in decrease in dendritic complexity, most pronounced in layer III pyramidal cells
Valverde, 1971	mouse visual cortex	dark-reared subjects placed back into normal environment; new growth in apical dendrites seen by 2 days
Volkmar and Greenough, 1972	occipital cortex	enriched environmental rearing resulted in increased higher order branches
Greenough and Volkmar, 1973	occipital cortex	follow up of Volkmar and Greenough (1972); found most increase in basal dendrites of pyramidal cells
Borges and Berry, 1976, 1978	rat visual cortex, layer IV stellate cell	dark rearing reoriented dendrites toward residual input
Ulyings et al. 1978	adult rat, visual cortex	enriched environments increased dendritic complexity and length of layer II, III pyramidal cells
Fiala et al., 1978	dentate granule cells	significant differences between complex and impoverished environment dendritic branches, length and width of dendritic fields
Juraska et al. 1980	adult rat, visual cortex	enriched environments increased dendritic complexity and length
Camel et al., 1986	rat visual cortex	dendritic alterations induced by exposure to a complex environment persisted even after return to individual caging for 30 days
Harris and Woolsey, 1981	mouse somatosensory cortex	vibrissal removal results in reduced representation in corresponding barrel cortex with increase in spared vibrissae
Conlee and Parks, 1983	avian cochlear nucleus	monaural acoustic deprivation led to 38% reduction in dendritic

		length
Schilling et al., 1991	in vitro study of Purkinje cell development	branching of Purkinje cell dendrites depends on functional synaptic contacts
Wallace et al., 1992	rat visual cortex, layer III pyramidal cells	+6 % increased total dendritic length in basal dendrites within four days of exposure to a complex environment
Mooney et al., 1992	hamster superior colliculus	enucleation results in superior collicular neurons to redirect their dendrites and develop response properties appropriate for the spared modality

Table 4. Representative studies demonstrate the effects of differential rearing and manipulations to activity on dendritic development. See text for details.

which must accommodate to changing functional demands.'' While the study of Valverde and many of the other early studies summarized in Table 4 illustrate that dendritic form is dependent on functional afferent pathways, the recent study of Mooney et al. (1992) illustrates that changes in afferent pathways can lead to dramatic dendritic reorganization that results in changes in the response modality of the affected cell. Specifically, Mooney et al. (1992) examined the effects of neonatal enucleation on the dendritic morphology of superior collicular (SC) neurons. Analogous to the earlier findings of Valverde, they found that the dendrites of SC neurons were redirected toward sources of residual input, the deep layer of the SC, whose input is from somatosensory axons. When they examined the physiological response properties of these neurons, they found that a majority (65%) of SC cells that would normally be visually responsive instead displayed somatosensory receptive fields. This underlies an important result: in the enucleated animal, the dendrites redirected their growth to find active afferents and, where these were of a different modality, the cell's response modality could be changed to reflect the residual source of afferent axons. That is, cells which normally showed visual response properties could be induced to develop somatosensory receptive fields when somatosensory inputs were the dominant source of inputs.

3.3.1 Directed dendritic development and patterns of activity

The previous results have shown that dendrites play an active role in outgrowth. In this section, we examine the relation between dendritic growth and activity in more detail. One model system in which this question has been addressed is the development of ocular dominance columns (introduced in section 2.3.1), in which the relation between OD columns and dendritic form may be assessed (Katz et al. 1989). A particularly interesting question concerns what would happen to those cells near a border between two columns. Would the dendrites of such cells ignore the border, or would their organization respect these borders? A striking finding of Katz et al. (1989) was that the dendritic arbors of layer 4c cells whose soma was near the border of an ocular dominance column in fact respected that border and arborized preferentially within its column (Figure 5).

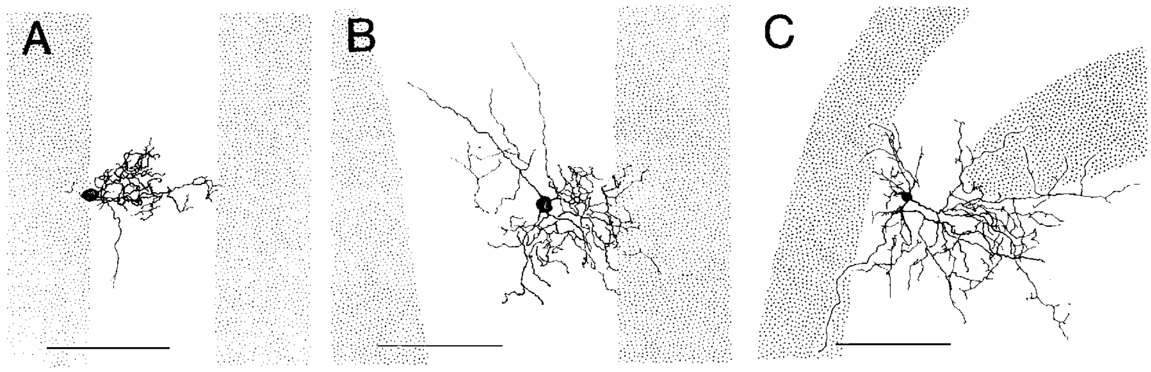

Figure 5. Cells in the optic tectum whose dendrites are confined to one stripe. A, Although the cell body is almost on the border, the dendrites are asymmetric and remain confined to one stripe; B, Cell with somata near the middle of stripe, whose dendrites grow up to and along a border but do not cross it; C, the cell's dendrites curve around a stripe from the other eye but remain in the home column (from Katz & Constantine-Paton 1988).

A natural question to ask regarding this development and the resulting dendritic configuration is, what is the signal regulating this bias away from neighboring columns? As Katz et al. (1989) note, one likely source of this signal derives from correlated activity within a column, since it originates from one eye, but is discontinuous at the borders between stripes from different eyes. This change in correlated activity might therefore underlie the bias away from the neighboring region if the postsynaptic cell maximized the amount of correlated input it receives. What would the role of such a developmental signal be? The most straightforward role would directly involve the development of the response properties of the cell. Since cells of layer 4c are almost exclusively monocular, that is they respond to stimulation from only one eye, by maximizing correlated input and avoiding uncorrelated input, a cell's dendrites would come to arborize within a single column and would thus help to maintain monocularity. Further, by taking advantage of a signal that was intrinsic to the afferents, this organization would come about without the need for pre-specifying it. Similar principles of dendritic development in the somatosensory cortex have also been observed (Greenough & Chang, 1988).

That the determination of dendritic form is dependent on patterned afferent activity is supported by analogous experiments by Katz and Constantine-Paton (1988) in the vertebrate optic tectum. While an optic tectum normally receives input from only one eye, it can be induced to receive input from two eyes by experimentally adding a third eye primordium during embryonic development (Constantine-Paton & Law, 1978). In these cases, afferents from the two eyes segregate into alternating stripes reminiscent of ocular dominance columns. A striking result of this study was that dendrites of tectal cells that do not normally receive input from more than one eye nonetheless are organized so as to respect the experimentally induced stripes. As in the Katz et al. (1989) study, it is the degree of correlation in the afferent activity, rather than simply the presence of activity that underlies dendritic organization.[1]

---

[1] Differences in the degree of rostral-caudal dendritic bias between normal and stripe-induced cells provides support for the view that this development involves progressive growth rather than elimination of exuberant structure (see Katz & Constantine-Paton (1988, p. 3178)). The conclusion according to Katz et al. (1989, p.1393) is that, ''the pattern of afferent segregation has played a significant role in shaping the structure of the postsynaptic dendritic field of cortical neurons''

An interpretation of these results is that dendritic segments function as detectors of correlated activity and grow preferentially into such regions. This is supported by the study of Tieman and Hirsch (1982), who demonstrated that exposure to lines of only one orientation during development had specific effects on dendritic development. When Tieman and Hirsch (1982) examined the dendritic field orientations of cells from cats raised with exposure to lines of a single orientation, they found that the dendritic orientations were significantly elongated in correspondence to this shift in the visual environment. An insight from this study is that a dendritic tree samples its input space actively in response to the structure of the environment. A similar result has been obtained for layer IV stellate cells by Coleman et al. (1981). They (1981, p.19) suggest, ''If an alteration of the spatio-temporal pattern of the afferent activity is sufficient to lead to dendritic alterations during development, this implies that dendritic trees may develop in a form that will optimize spatio-temporal summation for the postsynaptic neuron." Flood and Coleman (1979) also reported a similar finding using [^{14}C]2-deoxyglucose. This study thus supports the principle that the dendrites of these cells developed so as to maximize their ability to sample their input space, according to variable environmental conditions.

Recently, Kossel et al. (1995) used many of the experimental manipulations that led to the activity-dependent rules for axonal growth to examine dendritic growth. Specifically, they used double labeling techniques to simultaneously visualize OD columns and dendritic morphology under conditions of monocular deprivation and divergent squint (strabismus). Strabismus results in a decrease in between-eye correlations and therefore should enhance ocular segregation, as has been seen in the case of presynaptic arborizations (Shatz et al. 1977). Kossel et al. (1995) found this to be the case for the dendritic fields of layer IV stellate cells, the primary cell type that seems to reflect the afferent arborization. In contrast, in the case of monocular deprivation, the dendrites of cells in the non-deprived column were only weakly influenced by borders, reflecting the decrease in the amount of uncorrelated activity across that border. Interestingly, Kossel et al. (1995) also found that cells in the deprived column extended their dendrites into the nondeprived activity. This agrees with other results we discussed indicating that dendrites are not merely passive structures but actively seek out regions of correlated activity. As Kossel et al. (1995) conclude, their results support the notion that dendrites develop according to similar principles as do axonal arbors and that both structures develop according to patterns of correlated input activity.

3.3.2 The cellular basis of directed dendritic growth

The activity-dependent stabilization of co-active synapses has come to be a central developmental theme in both experimental work (rev. in Shatz 1990; Cline 1991) and in computational models of development (reviewed in Goodhill, 1992). The discovery of the NMDA type of glutamate receptor made an important contribution to this research. A primary reason for this is that the kinetic properties of the NMDA-receptor fit with the cooperative account of synaptic plasticity involved in both developmental findings (reviewed in Shatz 1990; Cline 1991) and in mature forms of plasticity (Kelso et al. 1986).

The view of learning and development stemming from the associative conditions of the NMDA-receptor fall under the rubric of Hebbian learning. In development, however, Hebbian learning is generally limited to governing the stabilization of existing synapses, as on the standard model. Hebb's (1949) original proposal, however, was made in the context of neural development and the activity-dependent construction of new synapses and neuronal cell assemblies. For example, Hebb includes a discussion of Kapper's neurobiotaxis theory, an early, extreme constructivist model of development, and defends a limited version of it. There is, in fact, suggestive evidence that the constructivist model is the most appropriate model for the properties that the NMDA-receptor possesses and that the Hebbian model is properly one which includes directed growth. For example, from their experimental observations, Katz et al.

(1989) suggest that such a broader action of the associative principles of the NMDA-receptor account may underlie the organization of dendritic structures. they state (p.3178):

> Our observations that single tectal dendrites can function as autonomous postsynaptic detectors of correlated afferents are consistent with the proposed role for the NMDA conductance. Depolarization of a single dendrite by activity in a subset of converging synapses would allow glutamate to activate the conductance within a restricted domain of the postsynaptic cell. This could, in turn, provide cues for stabilizing and enlarging a small portion of the dendritic arbor, independent of the behavior of other dendrites.

Studies in culture further support the role for NMDA-mediated constructive processes in dendritic development. For example, Brewer and Cotman (1987) found that NMDA-receptor mediated activity in hippocampal dentate granule cell cultures results in significant branching and outgrowth while NMDA blockade led to a significant decrease in these measures. Similar results in a variety of other systems have been reported (e.g., Pearce et al. 1987; Balazs et al. 1989; Bulloch & Hauser 1990).

4 Principles of constructivist growth

We are now in a position to extract some principles of constructivist development and their implications for the learning properties of a developing system. The evidence we reviewed in previous sections indicates that dendrites are not simply passive structures that develop according to an intrinsic schedule. Rather, they are dynamic structures whose mature morphology depends fundamentally on patterns of afferent termination.[2]. At the basis of the characterization we present is the notion that the informational structure of the environment plays a fundamental role in determining the structural properties of cortex which correspond to, or underlie, its representational properties. In order to extend this characterization, we shall consider two important and related consequences of activity-dependent development for the representational properties of cortex. These are:

1. Activity-dependent development may be seen as a directed construction of the representation or feature space that underlies mature cognitive skills.

2. Activity-dependent development underlies functional clustering, which define local regions of higher-order feature extraction.

One of the most important principles of cortical representation involves the use of geometric principles in information processing design (Mitchison & Durbin 1986; Mead 1989; Durbin & Mitchison 1990; reviewed in Churchland & Sejnowksi 1992); regarding this principle, Mead (1989, p.277) states:

---

[2] Axonal growth will fit into this account in the following sense. Local axonal growth may be sensitive to the development and stabilization of synapses, so that local axonal outgrowth may result from synapse formation. This would have the effect of putting more presynaptic structure into a local region in an activity-dependent manner, thereby essentially increasing the probability of subsequent synapse formation in that region. The outgrowth of axonal projections, such as the development of horizontal connections (Callaway and Katz 1991) suggest that the elaboration of axon terminals at this fine level may proceed according to this manner

> Computation is always done in the context of neighboring information. For a neighborhood to be meaningful, nearby areas in the neural structure must represent information that is more closely related than is that represented by areas further away. Visual areas in the cortex that begin the processing sequence are mapped retinotopically. Higher-level areas represent more abstract information, but areas that are close together still represent similar information. It is this map property that organizes the cortex such that most wires can be short and highly shared; it is perhaps the single most important architectural principle in the brain.

From this principle, the three-dimensional structure of a neural area can be conceived as defining a feature space in which geometric distance has an direct analogue in terms of representational properties. Since this structure is defined by the three-dimensional topology of neural connectivity, and constructivist principles play a central role in determining this topology, it should now be apparent how these principles allow for environmentally-derived activity to fundamentally shape the representational properties of cortex. The inclusion of geometric principles of information processing entail that locally regulated processes of outgrowth correspond to semantically coherent principles of representation construction. In section 5, we shall suggest that this feature confers important and unique learning-theoretic properties on developing cortex. First, though, we discuss the important constructivist principle that activity-dependent processes play a fundamental role not only in constructing this feature space but also in how it is sampled.

Whereas one may roughly think of the construction of a feature space to be a presynaptic feature, the postsynaptic cell also determines the properties of this space while its integrative properties determine how it is sampled. A popular view of the integrative properties of neurons, however, stems from the connectionist neuron. At its most basic, such a model neuron typically sums its input and sends a (perhaps graded) output if some threshold is exceeded. This makes the entire cell the basic computational unit. In contrast, biological neurons are thought to segregate into sub-regions that function as autonomous processors, where these sub-units comprise the basic computational units. In particular, and of relevance to our discussion, local regions of dendrites may function as electrically isolated subunits (Koch et al. 1982, 1983; Shepherd & Brayton 1987; Mel 1992a; 1992b; 1994; Jaslove 1992). A basis for this position is the finding that dendrites are not simple signal integrators with passive conductance properties as in classical cable models (Rall 1964). For example, imaging studies have found that some dendritic systems (e.g., CA1 hippocampal neurons) possess a heterogeneous distribution of voltage-gated Ca^{2+} channels, which suggest nonlinear membrane properties (Jones et al. 1989; Regehr et al. 1989). Intradendritic recordings in these cells likewise reveals strong nonlinearities in their electrical properties (Bernardo et al. 1982; Wong et al. 1979).

Theoretical work suggests that these properties would confer a number of additional computational properties on a cell, with the result that such a cell would be a substantially more powerful computational device (Feldman & Ballard 1982; Durbin & Rumelhart 1989; Mel & Koch 1990; Koch & Poggio 1992). A well-known abstract model of such a cell is the sigma-pi model neuron. The output of a sigma-pi neuron is the weighted sum of multiplicative groups of input, and thus each group is akin to a hidden unit in a standard PDP model neuron. Such a cell thus has the property that it can perform the nonlinear discrimination that requires a hidden-layer network of connectionist neurons. More biological versions have also been proposed (Jaslove 1992). According to this view, along with others positing nonlinear dendritic processing, the pattern of termination onto a dendritic tree is of central importance to the determination of the integrative properties of that cell. In particular, what is referred to as ''clustering'' of related inputs would result in a pattern of termination that would mirror the informational structure of the input to that cell: electrotonically close synapses would encode some common feature of the input space (Mel 1992a; 1992b; 1994) and effectively fire the cell.

The presence of cluster-encoded features would significantly alter both the representational properties of cortex and its computational power.

4.1 Developmental mechanisms underlying clustering

For clustering to serve an interesting information processing role, there must be a mechanism whereby temporal patterns of activity are somehow translated into appropriate spatial patterns of termination. As Mel states (1992b, p.43):

> the ordering of afferent connections onto an excitable dendritic arbor is a crucial determinant of the cell's responses to different patterns of synaptic input: It is this ordering, or permutation, that determines which input patterns will activate synapses that are spatially grouped into clusters, and which will not. If the nervous system is to take advantage of this capacity for pattern discrimination based on spatial ordering, then a learning mechanism capable of manipulating synaptic ordering must be available to these neurons.

A number of Hebbian schemes have been proposed to underlie the formation of these clusters, whereby a cell could effectively tune itself to its input space (Brown et al. 1991; Mel 1992a; 1992b). For example, Mel (1992a; 1992b) constructs what he calls a ‘‘clusteron", a simple model in which a Hebb-style rule reorganizes input so that co-active terminals will tend to cluster. Mel's model, though, operates by numerous random rearrangements of synapses on a fixed 1-dimensional postsynaptic structure and tests these various patterns of termination by the stabilization rules of the standard model. This suffers from a number of problems. Primary among these is the lack of biological plausibility of establishing numerous contacts from an axon to the same postsynaptic target over time. There appear to be too many confounding factors present to suppose that an axon can sample the same postsynaptic target over time. In addition, these postsynaptic structures themselves develop over time and thus are not the static structures such a view supposes. Below, we examine a more robust mechanism that may underlie this clustering.

This sampling problem has been considered in a more general context in Montague at al. (1991) and Gally et al. (1990). Through a consideration of the sparse connectivity of the developing nervous system, Gally et al. (1990) proposed the action of a spatially diffusible substance that could communicate a cell's activity to all cells within its effective diffusible neighborhood, whether or not that cell shared a synaptic contact with it (Figure 6). Such a volume signal has a number of attractive properties from a developmental and computational perspective (Montague et al. 1991) and has been proposed to underlie a form of learning (defined below) referred to as volume learning (reviewed in Montague & Sejnowski, 1994). Gally et al. (1990) proposed that nitric oxide, a membrane permeable gas, possesses a number of characteristics that make it a leading candidate for such a role. Subsequent research has found that nitric oxide plays a key role in synaptic plasticity (Bohme et al. 1991; Haley et al. 1992; O'Dell et al. 1991; Schuman and Madison 1991) and transmission (O'Dell et al. 1991; Manzoni et al. 1992; Montague et al. 1994).

The familiar Hebbian learning rule can be modified to become a volume learning rule where the postsynaptic term is replaced by one that includes reference to the production of a diffusible substance as in (see Montague et al. 1991; Montague and Sejnowksi, 1994):

$$\Delta w(t) = \eta\left[x(t) - \theta\text{pre}\right]\left[\mu(r,t) - T\text{pre}\right]$$

where $\Delta w(t)$ is the change in the 'weight' or synaptic efficacy of a connection, η is a constant controlling rate of change of synaptic efficacy, $x(t)$ is a measure of presynaptic activity and θpre is a threshold that determines whether a terminal is active at time t. $T\text{pre}$ is a threshold, dependent on the activity of the presynaptic terminal, which determines the direction of synaptic change. The postsynaptic factor of typical Hebbian rules has been replaced by a term for substance concentration, $\mu(r,t)$ at time t located at position r.

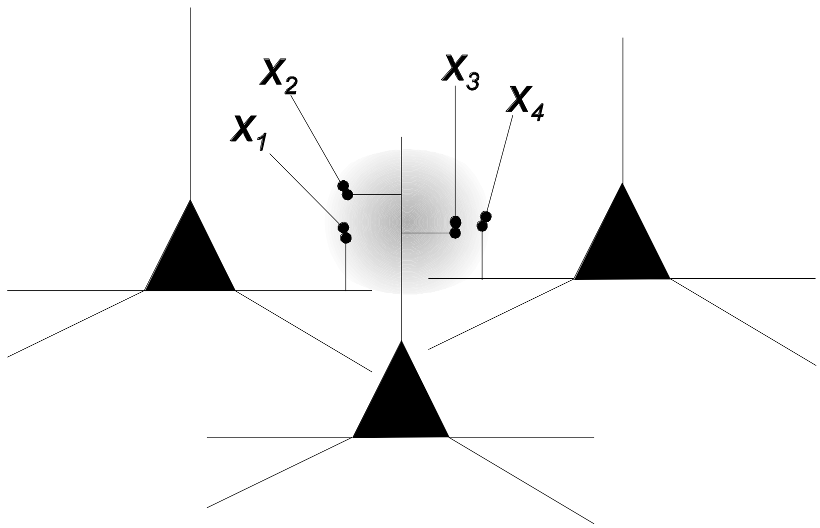

Figure 6. A diffusible substance allows synapses in a local volume of tissue to communicate whether or not they share a connection. Using such a signal, it is possible for synapse X_1 and X_4 to modify their weights according to an associative learning rule. (From Montague & Sejnowski 1994).

A learning rule such as (1) takes coincidence ''off the synapse'' and into a local volume of neural tissue, thereby allowing the volume to hold associations. The reason for this is that the postsynaptic factor that is required for an associative condition is no longer confined to the cellular compartment but is now broadcasted throughout a local volume to all cells that send processes into that volume.[3] It should now be apparent how this results in robust sampling mechanisms and clustering. Specifically, rather than having to sample identical postsynaptic structures as on a model that is confined to the postsynaptic compartment to hold coincidence, a volume rule allows cells to sample these diffusion-defined volumes. This has the additional advantage of allowing informationally-related features to be encoded across a group of cells synapsing within that volume even if there are no direct synaptic contacts between the cells.

As we are exploring in large-scale computer simulations, this scheme may be easily modified to include activity-dependent branching by making the probability of branching/retraction at some

[3]Specificity is maintained by requiring that the presynaptic terminal be coincidentally active (See Montague & Sejnowski (1994) for discussion).

terminal segment proportional to the weight of its associated synapse over some time, so that branching/retraction would be slow relative to rates of weight change. The effect of making the probability of branching dependent on synaptic weight is that this automatically transfers the associative conditions necessary for weight changes to the conditions for branching/retraction. The value of directed growth into these volumes is that it augments the processes leading to what we refer to as spatial clustering, that is, functional clustering of statistically correlated afferent axons into spatial domains which define higher-order features of the input space. This, then, corresponds to a locally-regulated growth that allows differential sampling as determined by the correlation structure of input patterns to determine spatial clusters. Further, since the production of the diffusible substance is postsynaptic, this indicates that postsynaptic structures play an important role in determining the properties of this feature space. We suggest that the establishment of spatial domains as regions of higher-order features will be central to the information-processing capacities of neuronal populations.

5 A learning-theoretic approach to development

The principles we examined in the last section suggested that directed growth plays a central role in constructing representations that mirrored the informational structure of the learner's environment. In this section, we extend our treatment of constructivism to examine its learning-theoretic properties. We suggest that the constructivist strategy of allowing the informational structure of the environment to play a role in determining the representational properties of cortex has a number of unique learning properties that undermine many of the central assumptions regarding acquisition in cognitive science.

5.1 Development and learnability

The dominant view of development in cognitive science, stemming from language acquisition, is often regarded as based on a rigorous model, formal learning theory. The formal learning theory that has been used most extensively in cognitive science derives from Mark Gold's work on language identification (Gold 1967). Gold established upper bounds or worse-case scenario results by asking what a general learner was capable of learning when presented with input strings belonging to some language. For each string, or input sample, the learner conjectures a hypothesis regarding the grammar that generates that language. The learner was said to *identify the language in the limit* if it eventually chose a grammar that was consistent with every string. A natural question to ask regarding the learner in Gold's paradigm is, where do the grammars that the learner conjectures come from? The answer is that Gold's learner adopts a maximally general strategy and first simply enumerates every grammar in the class. Starting with the first grammar, the learner then rejects each grammar in turn if it is inconsistent with the inputs so far and conjectures the next grammar in the enumeration. It follows that this learning procedure will eventually converge on the grammar if the correct grammar has some finite position in the enumeration. A ranking of grammars known as *The Chomsky hierarchy* results in the class of primitive recursive grammars being the highest or most powerful that are learnable by such a system since they are the highest decidable language (which just ensures that the right grammar has a finite place in the enumeration). As Pinker (1979) notes, this learner may have to test on the order of 10^{100} possible grammars even in an extremely simplified case---a computation that could never actually be performed. The reason the learner was so slow, then, was because it adopted such a general strategy; although this guarantees convergence, it makes learning in general prohibitive because of the vast search required.

While these results appear prohibitive for learning, they are limited by their concern only for convergence in the limit and their requirement that the learner precisely identify the target concept (no mistakes allowed). In 1984, Valiant introduced a probabilistic model of learning that remedied these two limitations, thus becoming the standard model of inductive inference in the field (see Dietterich 1990 and Natarajan, 1991 in the case of machine learning). Rather than require the learner to never make mistakes, Valiant's learner was allowed to make a hypothesis that was with high probability a good approximation. This framework was thus dubbed the ''probably approximately correct'' (PAC) model of learning (Angluin 1988). The model also addressed the question of convergence time, as it distinguished between feasible and infeasible learning by classifying problems according to whether or not they were learnable in polynomial time. Valiant's model thus shifted the main emphasis of the learning problem from what is in principle learnable to what is learnable from some representation class in feasible time.

As we mentioned, the key result that came out of the Gold paradigm was that the child must come equipped with a highly restricted set of hypotheses regarding the target grammar--in the case of language, a universal grammar. This conclusion falls out of the view of learning as essentially a search problem through a hypothesis space to the target concept. In order to make this a feasible search, the space must be restricted by building in an *inductive bia*s, roughly the system's background knowledge. One of the key virtues of the Valiant model was that it allowed for a quantification of the relation between inductive bias and learning performance from within a complexity-based learning theory (e.g., Haussler 1989). These results, then, were able to provide bounds on the feasibility of learning various problems relative to the complexity of the representation class or background knowledge, and so were able to demonstrate what could not be fully characterized in the earlier limit-based FLT: learning systems face severe learning-theoretic pressures and can be successful in some domain only if they have solved this difficult prior problem involving representation. That is, from the perspective of the PAC model of learning, the fundamental problems of learning are not those involving statistical inference, but center around how to find appropriate representations to underlie efficient learning (reviewed in Geman et al. 1992). It is this problem that precedes the treatment of learning as statistical inference, as the choice of representation class that a learner possesses will largely determine the success of learning as statistical inference. These results apply directly to the problem of development, as the question of how a system determines the representations that underlie efficient learning in some domain is really a question of what developmental strategies a system has come upon that are an adequate response to these learning-theoretic pressures. This, then, implies that this fundamental problem facing natural cognitive systems is a developmental one.

What makes the problem of learning from this perspective so difficult? In statistical studies, estimation error is decomposed into two components: bias and variance. Very roughly, bias concerns the distance between the target function and the closest function in the learner's hypothesis space, whereas variance refers to the distance between the target function and the estimated function. In the context of neural networks, a small network will be highly biased in the sense that the class of functions allowed by weight adjustments will be very small. If the target function is poorly approximated by this class of functions, then the bias will contribute to error. By making a network large, and so flexible in terms of what it can represent (by decreasing bias) the contribution of variance to error typically increases. High variance can only be overcome by prohibitively large training sets. As Geman et al. (1992) state it, this seems to result in a dilemma: highly biased learners will work only if they have been carefully chosen for the particular problem at hand whereas flexible learners seem to place to high a demand on training time and resources. This is essentially the impasse that leads to nativism. Indeed, from an entirely different perspective, Geman et al. (1992) likewise suggest that purposefully-introduced bias may be the best way out of this dilemma.

Here we raise a neurobiological constraint. That is, the suggestion that significant elements of the problem domain must be built in is so far motivated purely from a learning-theoretic perspective. The degree to which this strategy is neurobiologically plausible, however, also needs to be assessed. For natural systems, the constraint that a learning theory be consistent with underlying neural mechanisms has been severely underestimated. Indeed, in our opinion this biological constraint has equal footing with the learning-theoretic constraint and both must be viewed as two complementary constraints that developmental systems must satisfy. Although the evolutionary psychology approach prides itself on being naturalistic, typically it does not attempt to identify neural mechanisms to underlie its developmental claims. As we consider next, such views become increasingly problematic when the issue of mechanism is raised.

As we suggested in Quartz and Sejnowski (1994), we think the view that strong, domain-specific knowledge is built into cortical structures runs into severe difficulties when assessed from the perspective of developmental neurobiology. Although we will not review that material in detail here, recent experiments involving heterotopic transplant studies (Stanfield & O'Leary 1985; Schlaggar & O'Leary 1991; reviewed in O'Leary et al. 1992), cross modal rewiring experiments (Frost 1982; Sur et al. 1988; Roe et al. 1990, 1992; Pallas et al. 1990; reviewed in Sur et al. 1990; Shatz 1992) and clonal analysis of cell migration (Walsh & Cepko 1988; 1992) have established that, while the cortex is not a *Tabula Rasa*, since there may be a common basic circuitry and repetitive arrays of cell types (see O'Leary et al. 1992), the regional characteristics of mature cortex depend fundamentally on interaction with afferent input and is largely equipotential at early stages (in agreement with studies on cortical plasticity and early brain damage). On this view, O'Leary (1990) refers to the immature cortex as protocortex, which shares a common laminated structure, cell types, and basic circuitry but which diminishes the need for prespecification. It is the differing pattern of afferent activity, reflective of different sensory modalities, that confers area-specific properties onto the cortex--not predispositions that are somehow embedded in the recipient cortical structure. Further, the fact that many of these processes operate before birth, as in the case of spontaneous visual activity (Maffei & Galli-Resta 1990; Meister et al. 1991), suggests that cortical specification could begin by the very mechanisms that will be used postnatally through interaction with an environment. Hence, the fact that various regions of cortex receive different patterns of afferent termination and activity seems to be the prime determinant of specialized cortical functions. Such a view, in which the cortex is ''enslaved by the periphery", has a number of clear advantages in terms of flexibly responding to varying environmental pressures, plasticity, and changing body size (see Walsh & Cepko 1992; 1993).

When the neurobiological constraint is added to the learning-theoretic one, it appears that another impasse arises. From the perspective of learning theory, it appeared that the only response to the learnability problem was to build in much of the problem domain *a priori* in the form of highly specialized structures. And yet, from the perspective of the biological constraint it appeared that cortical structures do not build in this knowledge, but rather allow both pre- and post-natal activity to determine the areal features of cortex. In the following section, we suggest that the constructivist strategy offers a powerful means of escaping this impasse.

5.2 The constructivist strategy

Recall that the paradigm of learning in the formal accounts we discussed is one known as *selective induction*, that is, one in which the concept to be learned is arrived at through the search through some hypothesis space. Most studies of learning assume that the learning mechanism's resources are fixed, or time-invariant, and so view learning as a process of selective induction. This is true for Gold's learner and most neural network modeling. Simple counting arguments show that the probability of a learner searching through a fixed hypothesis space to successfully learn a concept chosen at random is exponentially small (reviewed in Dietterich 1990) is because successful learning in this paradigm demands that the hypothesis space the learner considers be an exponentially small subset of possible concepts (see Blumer et al. 1988). Hence, the failure to learn really may just indicate that the learner has chosen the wrong hypothesis space rather than reveal anything significant about the learnability of some target concept. As Baum (1989, p.203) states, ''a pragmatic learner should be willing to use any class of representations necessary to solve his problem. He should not be limited by *a priori* prejudices.'' This, then, suggests that the more central issue is one of model selection, or, more generally, one of determining appropriate representations to underlie some task, and so we turn to the constructivist strategy in these terms.

In contrast to learning as selective induction, the central component of the constructivist model is that it does not involve a search through an *a priori* defined hypothesis space, and so is not an instance of model-based estimation, or parametric regression. Instead, the constructivist learner builds this hypothesis space as it learns, and so is characterized as a process of activity-dependent construction of the representations that are to underlie mature skills. This has shifted the problem from one of parametric estimation to a nonparametric regime. We, therefore, need to ask, what is the effect of allowing a system to add structure as it learns?

We begin a description of this form of learning with an important asymptotic result that provides some useful notions for what follows. From a nonparametric framework, White (1990) has demonstrated that a network with the capacity to add units at an appropriate rate relative to its experience is a consistent nonparametric estimator and so may learn any arbitrary mapping. White's demonstration depends on a general nonparametric strategy: slowly increase representational capacity by reducing bias at a rate that also reduces variance. Bias in a network is dependent on the number of units, and so as a network grows its approximation capacities increase. Indeed, a number of general results of the analysis of neural networks has been in terms of their approximation properties (e.g., Cybenko 1989 Hornik et al. 1989; Girosi & Poggio 1990). For example, Hornik et al. (1989) demonstrated that multilayer feedforward networks using squashing functions are universal approximators, capable of arbitrarily accurate approximations to arbitrary mappings. In the context of establishing consistency results, the important element of this growth is that it occurs at a rate that reduces variance. Encouraging bounds on the rate of convergence have recently been obtained (Barron 1994).

Moving from these results to complexity-based ones, it has been demonstrated that constructivist networks possess significantly more powerful learning properties than fixed networks. Indeed, Judd's (1988) demonstration that fixed networks are NP-complete has been extended to architectures of just three nodes (Blum & Rivest 1988). These results stem from considering the complexity of what is called ''the loading problem'', that is, given a set of training examples (pairs of input and output) is there a set of weights that produce the right output for each input? This is really just a memorization problem, but one that must be solved before questions of generalization to novel inputs can be assessed. As Blum and Rivest (1988) note, these results stem from the fixed architecture property of the networks under consideration. In contrast, the loading problem becomes polynomial if the network is allowed to add hidden units, and so PAC learning is possible since the VC-dimension for such a net is polynomial. For example, Redding et al. (1993) have presented a constructivist algorithm that trains in time polynomial in the number of examples and also finds a compact representation (see also Baum 1988; 1989).

Together, these results indicate a fundamental distinction between the constructivist strategy and models of selective induction. In order for the latter to have any chance of learning, the network must build in much of the problem domain *a priori*. Besides this strategy being implausible from a neurobiological perspective, there are other more general reasons why employing highly-biased networks is not a sound strategy in the biological case. Primary among these is that the highly-biased network will only work for the specified application, and if the nature of the application is not properly understood then the network will in fact be a poor performer. Hence, the attempt to tailor-make network architectures to suit the particular demands of some problem domain may be a useful heuristic strategy for artificial networks whose problem space is defined, or at least delimited, in advance by the designer. But, biological learners face an additional problem: not only is the problem space not *a priori* defined, but it is itself changing. That is, the environment is highly nonstationary as well. Systems that are highly specialized for the anticipation of a particular problem domain will of course fail in the event of significant changes to that domain. The upshot is that specialization may bring efficiency, but it comes at the expense of flexibility.

The nonstationarity of the environment is a severe challenge to finding robust and flexible systems. While most natural systems are only confronted with ecological changes, human cognition is confronted by both cultural and technological innovations that require highly flexible and adaptive representations. It is questionable whether the pace of such change could be matched by representations that required an intrinsic specification. The constructivist learner, on the other hand, need not be committed to an *a priori* defined class of representations, but may allow the structure of the environment to guide the construction of the representations it uses. Recent work suggests that this strategy is as general purpose as possible, and remains to some extent a form of representational plasticity throughout the lifespan. By allowing a network to build its representations as it learns in response to the informational structure of its environment, the question of learnability becomes relaxed from the question of what is learnable from some particular representation class to the question of what is learnable from any representation class. For example, Baum (1989; 1990) has shown that networks with the power to add structure as a function of learning are complete representations, capable of learning in polynomial time any learning problem that can be solved in polynomial time by any algorithm. As Baum notes (1989, p.201) this makes the learner a general or universal learner. This is in contrast to systems which utilize incomplete representations, as in a fixed a priori defined hypothesis space. As we consider below, most negative learnability results depend on a system using incomplete representations.

The schemes we examined in section 4 indicate how the addition of structure was not simply the result of an intrinsic maturational timetable but depended on the informational structure of the environment and particular learning episodes to construct neural structure in an experience-dependent fashion. Further, the schemes we presented have the effect of adding structure in those regions where additional parameters are required in order to represent or approximate some function. In artificial neural networks a number of algorithms have been proposed which add structure in a similar manner (e.g., Fahlman & Lebiere 1990; Frean 1990; Hirose et al. 1991; Platt 1991; Azimi-Sadjadi et al. 1993; Wynne-Jones 1993; Kadirkamanathan & Niranjan 1993; Shin & Ghosh 1995). These constructivist algorithms have become increasingly sophisticated and now regularly beat fixed architectures on benchmarks and also find compact representations. The general strategy of constructivist learning is this: rather than start with a large network as essentially a guess of the class of target concepts, the constructivist model avoids the difficulties associated with overparameterized networks by starting with a small network. The learning algorithm then adds appropriate structure according to some performance criterion and where it is required until a desired error rate is achieved. This allows the particular structure of some problem domain to determine the connectivity and complexity of the network. Since the network has the capacity to respond to the structure of the environment in this way, the original high bias is reduced through increases in network complexity, which allows the network to represent more complex functions. Hence, the need to find a priori a good representation is replaced by the flexibility of a system that can respond to the structure of some task by building its representation class as it samples that structure to learn any polynomial learnable class of concepts.

The core feature of the constructivist model is that it uses a strategy in which it adds parameters as a function of exposure to a problem domain. This is related to another important property of nonstationary learning mechanisms: incremental learning. Empirical studies of nonstationary learning have demonstrated a further important property of their learning which refutes a central claim of traditional acquisition (Chomsky 1965; Wexler & Culicover 1980). One of the motivations for Chomsky's instantaneous learning idealization was the assumption that resource limitations of the learning mechanism would have only negative effects on learning. Hence, idealizing to the mature learning mechanism was taken to be an innocuous assumption since this would, if anything, result in a more powerful learning mechanism. Recently, this assumption has been contradicted by a number of studies demonstrating that an incremental learner---one in which some resource is initially limited---actually may allow a learner to learn some problem which would not be possible with the mature mechanism. One of the best examples stems from Elman (1990; 1991), who demonstrated an important property of incremental learning in applications to language acquisition. Specifically, Elman (1991) found that networks that start small by limiting some resource, such as working memory, could learn the structure of embedded sentences with long distance dependencies, whereas larger networks could not. Briefly, an initially restricted network was more sensitive to the low-order statistics in the input data than was a larger network, as these resource limitations effectively acted as a filter on the input to simplify the initial input to the system. However, as these resources increased, the prior learning served as a constraint on subsequent learning and the network could effectively use the lower-order statistics it had learned as a basis to learn the higher-order statistics (see Fahlman & Lebiere (1990) for a discussion of incremental learning in the context of constructivist networks). In contrast, a network that started with the full complement of resources could not learn the task. This suggests that starting small and increasing resources not only is an important property of learning for approximation and representational capacities, but also corresponds to an important heuristic for staging the complexity of learning problems by allowing resource limitations to effectively act as a filter over the incoming signal.

5.3 Constructivism and language acquisition

It is sometimes assumed that the classical learnability arguments of cognitive science proved that human languages are not learnable. For example, Tooby and Cosmides (1992) cite negative results on language learnability as supporting the view of cognitive structures as intrinsically specified and selected for by evolutionary pressures. However, it is not the case that classical learnability arguments showed that human languages are unlearnable. Indeed, as we have discussed, the negative learnability results of cognitive science depend on a particular model of learning as selective induction, a worst case analysis, and extremely rigid performance conditions. Hence, the negative learnability results do not indicate anything about the learnability of human language as much as they do about the insufficiency of the particular learning model.

How should negative learnability results be interpreted? Generally, negative results on learnability can stem from two main sources. For example, if C is not learnable from H, then either C is too large or H is too restricted (see Baum 1989; 1990; Valiant 1991). Most negative results, such as the loading problem results (Judd 1988; Blum & Rivest 1989) are of the second sort. Indeed, the constructivist network escapes these negative learnability results by constructing more powerful representations than the fixed architecture can represent. It is important to note, however, that the common claim regarding the unlearnability of natural languages is not taken to be of this form, that is, is thought not to be one concerning unlearnability relative to some class of representations. Rather, it is thought to be a representation-independent result concerning the unlearnability of some class relative to any class of representations. Is this claim justified? The answer is no---the only representation-independent results concern complicated cryptographic functions, such as polyrandom functions (Goldreich et al. 1984). Since natural systems are not thought to be confronted with the general decryption problem, these results are of little relevance to the learnability of natural functions. Hence, there is no justification for the claim that functions of interest to natural systems are representation-independent unlearnable. And, to the extent that the problems of interest to natural systems are learnable in this latter sense, which appears largely to be the case, the constructivist learner is a general and powerful approach to these problems.

The negative results stemming from Gold's model follow from a very weak conception of the learner and evidence, very rigid performance requirements, a worst-case analysis, and the assumption that the learner undergoes no qualitative changes. The view of learning we presented in this paper suggests that the appropriate response to these results is to question this model of learning rather than embrace its results as the cornerstone of a research program. The dynamic interaction between the structure of the environment and neural mechanisms that are responsive to that structure, which we have placed at the center of the constructivist strategy, suggests that this is a far richer view of learning than the traditional view allows, and one whose properties are only beginning to be understood. For example, neural networks have found much more implicit structure in environmentally-derived activity than was previously thought to be available to a learner. How these cues are integrated into this view of learning as structural development, and how the system's actions on its environment and their consequences shape a theory of the development of cognitive representations remain as outstanding problems in the study of cognition. Some of the principles we have explored in this paper are very simple steps in the direction of characterizing the complex interactions governing neuronal development and the corresponding construction of the representations that underlie mature cognitive skills. There is little question that this interaction will be far richer than the principles we have considered can capture.

6 Conclusions

Understanding how cognitive and neural processes interact to construct the representations that underlie mature cognitive skills is perhaps the central problem of cognition. We have suggested that the nonstationarity that is the hallmark of this process can be approached from its neural basis. From this approach, we have suggested that development is best seen as the progressive construction of representations in which environmentally-derived activity plays a fundamental role in determining the structure and function of the mature brain. This constructivist strategy uniquely solves two fundamental problems: 1) developmental strategies must solve the severe learning-theoretic pressures to find adaptive representations that underlie efficient learning, and 2) the developmental mechanisms that underlie successful developmental strategies must themselves be robust and biologically feasible.

From an evolutionary view, this strategy is in contrast to the popular view of the evolution of cortex as the progressive increase in specialized cortical structures. At the expense of increased vulnerability during a protracted developmental period, the constructivist strategy possess a number of important advantages: alterations in the learner's environment through either natural or human innovation can be accommodated by flexible cortical representations without the requirement that such changes be mediated by intrinsic mechanisms of specification. Cortical evolution should thus be viewed as the progression to more flexible and adaptive representations and away from rigid, pre-specified representations that are the hallmark of lower systems. According to such a view, the protracted postnatal development that is so characteristic of human development represents a stage of representation construction that, while coming at the cost of increased vulnerability and parental investment, also permits a prolonged flexibility and striking capacity of the environment to shape and structure the brain. For this reason, we suggest that human development is characterized by two process, first a prolonged period in which neural structures respond to the informational structure of the environment to construct representations, and, second, efficient learning, which is characterized by a period of rapid learning made possible by the first.

It is sometimes argued that bipedalism placed such severe constraints on the dimensions of the birth canal that it resulted in a particularly immature human brain at birth. In fact, there is nothing unique about the degree of immaturity of the human brain at birth. As Gibson (1993) points out, the human brain at birth is about 25% of its mature weight, making it about average among mammals. In contrast, the kitten's brain at birth is only 10% of its mature weight. Depending on cortical layer, total dendritic length for cells in human primary visual cortex at birth are between 30 and 50% of adult values (Becker et al. 1984). For these layer V pyramidal cells adult values are reached by around four months compared to 18 months for layer III cells. What is striking about human cortical development, however, is the protractedness of non-visual cortical development. Cells in frontal cortex are much slower to develop and undergo the majority of their growth after two years of age (Schade and van Groenigen, 1961). In addition, the extent of their postnatal development is dramatic---these cells grow to over thirty times their dendritic length at birth. Scheibel (1993) likewise reports an extended protractedness of dendritic development in Broca's area, in which mature morphologies emerge only after 6-8 years. Why, then, is human non-visual cortical development so slow and extensive? The view we have suggested is that this reflects a powerful learning strategy by allowing the structure of these circuits to be critically dependent on the structure of the environment which they sample over this prolonged developmental period. In addition, the dendritic structures of these cells become markedly more complex than those in visual cortex, suggesting their information-processing role is more complex. In the case of Broca's area, the dendritic structure of these cells becomes more complex than the analogous site in the right hemisphere. In this case, activity-dependent rules of the sort we have considered for the development of dendritic structures would result in increasingly complex structures with correspondingly increasing complex inputs since their role is to build circuits that mirror the informational structure of their input.

We have outlined a framework for approaching the problem of nonstationarity in development. We suggest that a solution to this problem is necessary for a theory of acquisition as the presence of large-scale structural changes during development severely alters the learning properties of such a system. We suspect that the extent of these structural changes has been undervalued by many cognitive theorists, perhaps due to the widely-held assumption that little actual changes occur in development. Additionally, the tendency to view the biological basis of development as one not contributing to the learning properties of a system (e.g., Lenneberg 1967), may have also contributed to the lack of interest in the neural basis of cognitive development. In contrast, we have suggested that cognitive and neural processes interact fundamentally throughout development. On a more methodological note, this interaction strongly suggests that cognitive and neural levels of description must be integrated into a single explanatory framework in order to characterize this fundamental interaction. Finally, the constructivist framework provides a meeting ground for cognitive scientists to interact with neuroscientists and to benefit from the major advances that have occurred in our understanding of the central nervous system. Rather than being strictly reductive, which is often thought to follow from a biological account, grounding a theory of development in biological structures and mechanisms points to the fundamental importance of the informational structure of the environment. This suggests that characterizing this structure and examining its role in neuronal development will likewise play an indispensable role in understanding how mature cognitive skills emerge.

References

Anderson, J. C., Dehay, C., Friedlander, M. J., Martin, K. A. & Nelson, J.C. (1992) Synaptic connections of physiologically identified geniculocortical axons in kitten cortical area 17. *Proceedings of the Royal Society of London* 250:187-94.

Antonini, A. & Stryker, M.P. (1993) Development of individual geniculocortical arbors in cat striate cortex and effects of binocular impulse blockade. *Journal of Neuroscience* 13:3549-73.

Azimi-Sadjadi, M.R., Sheedvash, S. & Trujillo, F.O. (1993) Recursive dynamic node creation in multilayer neural networks. *IEEE Transactions on Neural Networks* 4:242-56.

Bailey, C.H. & Chen, M. (1988a) Long-term sensitization in Aplysia increases the number of presynaptic contacts onto the identified gill motor neuron L7. *Proceedings of the National Academy of Sciences of the United States of America* 85:9356-9.

Bailey C.H. & Chen, M. (1988b) Morphological basis of short-term habituation in Aplysia. *Journal of Neuroscience* 8:2452-9.

Balazs, R., Hack, N., Jorgensen, O. S. & Cotman, C. W. (1989) N-methyl-D-aspartate promotes the survival of cerebellar granule cells: pharmacological characterization. *Neuroscience Letters* 101:241-6.

Barkow, J.H., Cosmides, L. & Tooby, J. (eds). (1992) *The adapted mind: Evolutionary psychology and the generation of culture*. Oxford University Press.

Barron, A.R. (1994) Approximation and estimation bounds for artificial neural networks. *Machine Learning* 14:115-33.

Baum, E.B. (1988) Complete representations for learning from examples. In: *Complexity in Information Theory*, ed. Abu-Mostafa. Springer-Verlag.

Baum, E.B. (1989) A proposal for more powerful learning algorithms. *Neural Computation* 1:201-207.

Becker, L.E., Armstrong, D.L. Chan, F. & Wood, M.M. (1984) Dendritic development in human occipital cortical neurons. *Developmental Brain Research* 13:117-124.

Bernardo, L.S., Masukawa, L.M. & Prince, D. A. (1982) Electrophysiology of isolated hippocampal pyramidal dendrites. *Journal of Neuroscience* 2:1614-1622.

Black, J. E., Isaacs, K.R., Anderson B.J., Alcantara, A.A., & Greenough, W.T. (1990) Learning causes synaptogenesis, whereas motor activity causes angiogenesis, in cerebellar cortex of adult rats. *Proceedings of the National Academy of Sciences of the United States of America* 87:5568-72.

Blum, A. & Rivest, R.L. (1998) Training a 3-node neural network is NP-complete. In*: Advances in Neural Information Processing Systems*, ed. D.S. Touretzky. Morgan Kaufmann.

Blumer, A., Ehrenfeucht, A., Haussler, D. & Warmuth, M. (1988) Learnability and the Vapnik-Chervonenkis dimension. *UCSC-CRL-87-20.*

Bohme, G.A., Bon, C., Stutzmann, J.M., Doble, A. & Blanchard, J.C. (1991) Possible involvement of nitric oxide in long-term potentiation. *European Journal of Pharmacology* 199:379-81.

Boothe, R. G., Greenough, W. T., Lund, J. S., & Wrege, K. (1979) A quantitative investigation of spine and dendrite development of neurons in visual cortex (area 17) of Macaca nemestrina monkeys. *Journal of Comparative Neurology* 186:473-89.

Borges, S. & Berry, M. (1976) Preferential orientation of stellate cell dendrites in the visual cortex of the dark-reared rat. *Brain Research* 112:141-7.

Borges, S. & Berry, M. (1978) The effects of dark-rearing on the development of the visual cortex of the cat. *Brain Research* 180:277-300.

Bourgeois, J.P., Goldman-Rakic, P.S. & Rakic, P. (1994) Synaptogenesis in the prefrontal cortex of rhesus monkeys. Cerebral Cortex 4:78-96.

Bourgeois, J.P., Jastreboff, P.J. & Rakic, P. (1989) Synaptogenesis in visual cortex of normal and preterm monkeys: evidence for intrinsic regulation of synaptic overproduction. *Proceedings of the National Academy of Sciences of the United States of America* 86:4297-301.

Brewer, G.J., & Cotman, C.W. (1989) NMDA receptor regulation of neuronal morphology in cultured hippocampal neurons. *Neuroscience Letters* 99:268-73.

Brown, R. (1973) *A first language: the early stages*. Harvard Univeristy Press.

Buell, S.J. & Coleman, P.D. (1981) Quantitative evidence for selective dendritic growth in normal human aging but not in senile dementia. *Brain Research* 214:23-41.

Bulloch, A.G. & Hauser, G.C. (1990) Sprouting by isolated Helisoma neurons: enhancement by glutamate. *International Journal of Developmental Neuroscience* 8:391-8.

Callaway, E.M. & Katz, L.C. (1991) Effects of binocular deprivation on the development of clustered horizontal connections in cat striate cortex. *Proceedings of the National Academy of Sciences of the United States of America* 88:745-749.

Callaway, E.M. & Katz, L.C. (1992) Development of axonal arbors of layer 4 spiny neurons in cat striate cortex. *Journal of Neuroscience* 12:570-82.

Camel, J.E., Withers, G.S. & Greenough, W.T. (1986) Persistence of visual cortex dendritic alterations induced by postweaning exposure to a "superenriched" environment in rats. *Behavioral Neuroscience*. 100:810-3.

Caviness, V.S. & Rakic, P. (1978) Mechanisms of cortical development: a view from mutations in mice. In: *Annual Review of Neuroscience*, ed. W.M. Cowan, Z.W. Hall, & E.R. Kandel.

Chang, F.L. & Greenough, W.T. (1984) Transient and enduring morphological correlates of synaptic efficacy change in the rat hippocampal slice. *Brain Research* 309:35-46.

Chang P.L., Isaacs, K.R. & Greenough, W.T. (1991) Synapse formation occurs in association with the induction of long-term potentiation in two-year-old rat hippocampus in vitro. *Neurobiology of Aging* 12:517-22.

Changeux, J.P., & Danchin, A. (1976) Selective stabilisation of developing synapses as a mechanism for the specification of neuronal networks. *Nature*, 264, 705-712.

Chomsky, N. (1965) *Aspects of the theory of syntax*. MIT Press.

Chomsky, N. (1980) Rules and representations. *Behavioral and Brain Sciences* 3:1-61.

Churchland, P. S. & Sejnowski, T. J. (1992) *The Computational Brain*. MIT Press.

Cline, H.T. (1991) Activity-dependent plasticity in the visual systems of frogs and fish. *Trends in Neurosciences* 14:104-111.

Coggeshall, R.E. (1992) A consideration of neural counting methods. *Trends in Neurosciences* 15:9-13.

Coleman, P.D., Flood, D.G., Whitehead, M.C. & Emerson, R.C. (1981) Spatial sampling by dendritic trees in visual cortex. *Brain Research* 214:1-21.

Conlee, J. W. & Parks, T.N. (1983) Late appearance and deprivation-sensitive growth of permanent dendrites in the avian cochlear nucleus (Nuc. Magnocellularis). *Journal of Comparative Neurology* 217:216-226.

Constantine-Paton, M. & Law, M.I. (1978) Eye-specific termination bands in tecta of three-eyed frogs. *Science* 202:639-41.

Cragg, B.G. (1975) The development of synapses in kitten visual cortex during visual deprivation. *Experimental Neurology* 46:445-451.

Cybenko, G. (1989) Approximation by superpositions of a sigmoid function. *Mathematics of Control, Signals, and Systems* 2:303-314.

Dekaban, A.S., & Sadowsky, D. (1978) Changes in brain weights during the span of human life: relation of brain weights to body heights and body weights. *Annals of Neurology* 4:345-356.

Dietterich, T.G. (1990) Machine learning. *Annual Review of Computer Science* 4:255-306.

Durbin, R. & Mitchison, G.J. (1990) A dimension reduction framework for understanding cortical maps. *Nature* 343:644-7.

Durbin, R. & Rumelhart, D.E. (1989) Product units: A computationally powerful and biologically powerful extension to backpropagation networks. *Neural Computation* 1:133-142.

Edelman, G. (1987) *Neural Darwinism: The theory of neuronal group selection*. Basic Books.

Elman, J. (1990) Finding structure in time. *Cognitive science* 14:179-211.

Elman, J. (1991) Incremental learning, or the importance of starting small. In*: Proceeding of the Thirteenth Annual Conference of the Cognitive Science Society*.

Fahlman, S.E. & Lebiere, C. (1990) The cascade-correlation architecture. In: *Advances in Neural Information Processing Systems*, ed. D.S. Touretzky. Morgan Kaufmann.

Feldman, J.A. & Ballard, D.H. (1982) Connectionist models and their properties. Cognitive Science 6:205-254. Fiala, B.A., Joyce, J.N. & Greenough, W.T. (1978) Environmental complexity modulates growth of granule cell dendrites in developing but not adult hippocampus of rats. *Experimental Neurology* 59:372-83.

Flood, D.G. & Coleman, P.D. (1979) Demonstration of orientation columns with [^{14}C] 2-deoxyglucose in a cat reared in a striped environment. *Brain Research* 173:538-542.

Fodor, J. (1983) The Modularity of Mind. Bradford Books.

Frean, M. (1990) The upstart algorithm: a method for constructing and training feedforward neural networks. *Neural Computation* 2:198-209.

Friedlander, M.J., Martin, K.A. & Vahle, H.C. (1985) The structure of the terminal arborizations of physiologically identified retinal ganglion cell Y axons in the kitten. *Journal of Physiology (Lond*) 359:293-313.

Friedlander, M.J. & Martin, K.A.C. (1989) Development of Y axon innervation of cortical area 18 in the cat. *Journal of Physiology (Lond)* 416:183-213.

Friedlander, M.J., Martin, K.A.C. & Wassenhove-McCarthy, D. (1991) Effects of monocular visual deprivation on geniculocortical innervaton of area 18 in cat. *Journal of Neuroscience* 11:3268-3288.

Frost, D.O.(1982) Anomalous visual connections to somatosensory and auditory systems following brain lesions in early life. *Brain Research* 255:627-35.

Gally J,A., Montague, P.R., Reeke, G.N. Jr & 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. *Proceedings of the National Academy of Sciences of the United States of America* 87:3547-51.

Gazzaniga, M.S. (1995) On neural circuits and cognition. *Neural Computation* 7:1-13.

Gazzaniga, M.S. (1992) *Nature's mind*. Basic Books.

Geman, S., Bienenstock, E., & Doursat, R. (1992) Neural networks and the bias/variance dilemma, *Neural Computation* 4:1-58.

Gibson, K R. (1990) New perspectives on instincts and intelligence: Brain size and the emergence of hierarchical mental constructional skills. In: *"Language" and intelligence in monkeys and apes: Comparative developmental perspectives*. ed. S. T. Parker, K.R. Gibson. Cambridge University Press.

Girosi, F. & Poggio, T. (1990) Networks and the best approximation property. *Biological Cybernetics* 63:169-76.

Globus, A., & Scheibel, A.B. (1967) The effect of visual deprivation on cortical neurons: a golgi study. Experimental neurology 19:331-245. Globus, A., Rosenzweig, M.R., Bennett, E.L. & Diamond, M.C. (1973) Effects of differential experience on dendritic spine counts in rat cerebral cortex. *Journal of Comparative and Physiological Psychology* 82:175-81.

Gold, E.M. (1967) Language identification in the limit. *Information and Control* 10:447-474.

Goldreich, O., Goldwasser, S. & Micali, S. (1984) How to construct random functions. *Journal for the Association of Computing Machinery* 33:792-807.

Greenough, W. T. & Chang, F.L. (1988) Dendritic pattern formation involves both oriented regression and oriented growth in the barrels of mouse somatosensory cortex. *Brain Research* 471:148-152.

Greenough, W,T. & Volkmar, F.R. (1973) Pattern of dendritic branching in occipital cortex of rats reared in complex environments. *Experimental Neurology* 40:491-504.

Haley, J.E., Wilcox, G.L. & Chapman, P.F. (1992) The role of nitric oxide in hippocampal long-term potentiation. *Neuron* 8:211-6.

Harris, R.M & Woolsey. T.A. (1981) Dendritic plasticity in mouse barrel cortex following postnatal vibrissa follicle damage. *Journal of Comparative Neurology* 196:357-37.

Haussler, D. (1989) Quantifying inductive bias: AI learning algorithms and Valiant's learning framework. *Artificial Intelligence* 36:177-222.

Hebb, D.O. (1949) *The Organization of Behavior: A Neuropsychological Theory*. John Wiley and Sons.

Hirose, Y., Yamashita, K. & Hijiya, S. (1991) Back-propagation algorithm which varies the number of hidden units. *Neural Networks* 4:61-6.

Hirschfeld, L. A. & Gelman, S.A. (eds). (1994*) Mapping the mind: Domain specificity in cognition and culture*. Cambridge University Press.

Hornik, K., Stinchcombe, M. & White, H. (1989) Multilayer feedforward networks are universal approximations. *Neural Networks* 2:359-366.

Hubel, D., & Wiesel, T. (1962) Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *Journal of Physiology (London)* 160:106-154 .

Hubel, D. & Wiesel, T. (1963) Receptive Fields of Cells in Striate Cortex of Very Young, Visually inexperienced Kittens. *Journal of Neurophysiology* 26:994-1002.

Hubel, D. & Wiesel, T. (1965) Binocular Interactions in striate cortex of kittens reared with artificial squint. *Journal of Neurophysiology*, 28, 1041-1059.

Hubel D. & Wiesel, T. (1972) Laminar and columnar distribution of geniculo-cortical fibers in the macaque monkey. *Journal of Comparative Neurology* 146:421-50.

Humphrey, A.L., Sur, M., Uhlrich, D.J. & Sherman, S.M. (1985) Projection patterns of individual X- and Y-cell axons from the lateral geniculate nucleus to cortical area 17 in the cat. *Journal of Comparative Neurology* 233:159-89.

Huttenlocher, P. R. (1979) Synaptic density in human frontal cortex-developmental changes and effects of aging. *Brain Research* 163:195-205.

Huttenlocher, P.R. (1990) Morphometric Study of human cerebral cortex Development. Neuropsychologia 28:517-527. Huttenlocher, P. R., & de Courten, C. (1987) The development of synapses in striate cortex of man. *Human Neurobiology* 6:1-9.

Huttenlocher P.R., de Courten, C., Garey, L.J. & Van der Loos H. (1982) Synaptogenesis in human visual cortex: evidence for synapse elimination during normal development. *Neuroscience Letters* 13:247-52.

Jacobs, B., Schall, M. & Scheibel, A.B. (1993) A quantitative dendritic analysis of Wernicke's area in humans. II. Gender, hemispheric, and environmental factors. *Journal of Comparative Neurology* 327:97-111.

Jaslove, S.W. (1992) The integrative properties of spiny distal dendrites. *Neuroscience* 47:495-519

Jerne, N. (1967) Antibodies and learning: Selection versus instruction. In: *The Neurosciences: A Study Program*, ed. G.C. Quarton, T. Melnechuk, F.O. Schmitt. Rockefeller University Press.

Jernigan, T.L., Archibald, S.L., Berhow, M.T., Sowell, E.R., Foster, D.S., & Hesselink, J.R. (1991) Cerebral structure on MRI, Part I: Localization of age-related changes. *Biological Psychiatry* 29:55-67.

Jones, O.T., Kunze, D.L. & Angelides, K.J. (1989) Localization and mobility of omega-conotoxin-sensitive Ca$^{2+$ channels in hippocampal CA1 neurons. *Science* 244:1189-93.

Judd, S. (1988) On the complexity of loading shallow neural networks. *Journal of Complexity* 4:177-192.

Juraska, J.M. & Fifkova, E. (1979) A Golgi study of the early postnatal development of the visual cortex of the hooded rat. *Journal of Comparative Neurology* 183:247-56.

Juraska, J.M., Greenough, W.T., Elliott, C., Mack, K.J. & Berkowitz, R. (1980) Plasticity in adult rat visual cortex: an examination of several cell populations after differential rearing. *Behavioral and Neural Biology* 29:157-67.

Kadirkamanathan, V. & Niranjan, M. (1993) A function estimation approach to sequential learning with neural networks. *Neural Computation* 5:954-75.

Kalil, R.E., Dubin, M.W., Scott, G. & Stark, L.A. (1986) Elimination of action potentials blocks the structural development of retinogeniculate synapses. *Nature* 323:156-8.

Katz, L.C. & Constantine-Paton, M. (1988) Relationships between segregated afferents and postsynaptic neurones in the optic tectum of three-eyed frogs. *Journal of Neuroscience* 8:3160-80.

Katz, L.C., Gilbert, C. D. & Wiesel, T.N. (1989) Local circuits and ocular dominance columns in monkey striate cortex. *Journal of Neuroscience* 9:1389-1399.

Katz, M.J., George, E.B., & Gilbert, L.J. (1984) Axonal elongation as a stochastic walk. *Cell Motility* 4:351-370.

Kelso, S.R., Ganong, A.H. & Brown, T.H. (1986) Hebbian synapses in hippocampus. *Proceedings of the National Academy of Sciences of the United States of America* 83:5326-30.

Koch, C. & Poggio, T. (1992) Multiplying with synapses and neurons. In: *Single Neuron Computation*, ed. T. McKenna, J. Davis, S. Zornetzer. Academic Press.

Koch, C., Poggio, T. & Torre, V. (1982) Retinal ganglion cells: a functional interpretation of dendritic morphology. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* 298:227-63.

Koch, C., Poggio, T. & Torre, V. (1983) Nonlinear interactions in a dendritic tree: localization, timing, and role in information processing. *Proceedings of the National Academy of Sciences of the United States of America* 80:2799-802.

Koester, S.E., & O'Leary, D.D. (1992) Functional classes of cortical projection neurons develop dendritic distinctions by class-specific sculpting of an early common pattern. *Journal of Neuroscience* 12:1382-93.

Kossel, A., Lowel, S. & Bolz, J. (1995) Relationships between dendritic fields and functional architecture in striate cortex of normal and visually deprived cats. *Journal of Neuroscience* 15:3913-3926.

Lee, K.S., Schottler, F. Oliver, F. & Lynch, G. (1980) Brief bursts of high-frequency stimulation produce two types of structural change in rat hippocampus. *Journal of Neurophysiology* 44:247-258.

LeVay, S., Wiesel, T.N. & Hubel, D.H. (1980) The development of ocular dominance columns in normal and visually deprived monkeys. *Journal of Comparative Neurology* 191:1-51.

Macnamara, J. (1982) *Names for things: a study of child language*. Cambridge, MA.: MIT Press.

Maffei, L. & Galli-Resta, L. (1990) Correlation in the discharges of neighboring rat retial ganglion cells during prenatal life. *Proceedings of the National Academy of Sciences* 87:2861-2864.

Manzoni, O., Prezeau, L., Marin, P., Deshager, S., Bockaert, J. & Fagni L. (1992) Nitric oxide-induced blockade of NMDA receptors. *Neuron* 8:653-62.

Mariani, J., Crepel, F., Mikoshiba, K., Changeux, J.P. & Sotelo, C. (1977) Anatomical, physiological and biochemical studies of the cerebellum from Reeler mutant mouse. *Philosophical Transactions of the Royal Society of London, Series B* 281::1-28.

Mathers, L.J. (1979) Postnatal dendritic development in the rabbit visual cortex. *Brain Research* 168:21-9.

McCasland, J.S., Bernardo, K.L., Probst, K. & Woolsey, T.A. (1992) Cortical local circuit axons do not mature after early deafferentation. *Proceedings of the National Academy of Sciences of the United States of America* 89:1832-1836.

Mead, C. (1989) *Analog VLSI and Neural Systems*. Addison-Wesley.

Mehler, J. (1985) Language related dispositions in early infancy. In: *Neonate cognition : beyond the blooming buzzing confusion*, ed. J. Mehler & R. Fox.

Meister, M., Wong, R., Baylor, D. & Shatz, C.J. (1991) Synchronous bursts of action potentials in ganglion cells of the developing mammalian retina. *Science* 252:939-943.

Mel, B.W. (1992a) NMDA-based pattern discrimination in a modeled cortical neuron. *Neural Computation* 4:502-17.

Mel, B.W. (1992b) Information processing in an excitable dendritic tree. *CNS Memo 17*, Computational and Neural Systems Program, California Institute of Technology.

Mel, B.W. (1994) Information processing in dendritic trees. *Neural Computation* 6:1031-85.

Mel, B.W. & Koch, C. (1990) Sigma-Pi learning: On radial basis functions and cortical associative learning. In: *Advances in Neural Information Processing Systems*, ed. D.S. Touretzky. Morgan Kaufmann.

Miller, K.D., Keller, J.B. & Stryker, M.P. (1989) Ocular dominance column development: Analysis and simulation. *Science* 245:605-615.

Mitchison, G.J. & Durbin, R. (1986) Optimal numberings of an N X N array. *S.I.A.M. Journal on Algebraic and Discrete Methods* 7:571-582.

Montague, P.R., Gally, J.A. & Edelman, G.M. (1991) Spatial signaling in the development and function of neural connections. *Cerebral Cortex* 1:199-220.

Montague, P.R., Gancayco, C.D., Winn, M.J., Marchase, R.B. & Friedlander MJ. (1994) Role of NO production in NMDA receptor-mediated neurotransmitter release in cerebral cortex. *Science* 263:973-7.

Montague, P.R. & Sejnowski, T.J. (1994) The predictive brain: temporal coincidence and temporal orden in synaptic learning mechanisms. *Learning and Memory* 1:1-33.

Mooney, R.D., Nikoletseas, M.M., King, T.D., Savage, S.V., Weaver, M.T. & Rhoades, R. W. (1992) Structural and functional consequences of neonatal deafferentation in the superficial layers of the hamster's superior colliculus. *Journal of Comparative Neurology* 315:398-412.

Moore, R.Y., & Bernstein, M.E. (1989) Synaptogenesis in the rat suprachiasmatic nucleus demonstrated by electron microscopy and synapsin I immunoreactivity. Journal of Neuroscience 9:2151-2162. Natarajan, B. (1991) *Machine Learning: A theoretical approach*. Morgan Kaufmann.

O'Dell T.J., Hawkins, R.D., Kandel, E.R. & 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. *Proceedings of the National Academy of Sciences of the United States of America* 88:11285-9.

O'Kusky, J., & Colonnier, M. (1982a) Postnatal changes in the number of neurons and synapses in the visual cortex (area 17) of the macague monkey: a stereological analysis in normal and monocularly deprived animals. *Journal of Comparative Neurology* 210:307-315

O'Kusky, J.,& Colonnier, M. (1982b) Postnatal changes in the number of astrocytes, oligodendrocytes, and microglia in the visual cortex (area 17) of the macague monkey: a stereological analysis in normal and monocularly deprived animals. *Journal of Comparative Neurology* 210:307-315.

O'Leary, D.D.M. (1990) Do cortical areas emerge from a protocortex? *Trends in Neurosciences* 12:400-406

O'Leary, D.D.M., Schlaggar, B.L. & Stanfield, B.B. (1992) The specification of sensory cortex: lessons from cortical transplantation. *Expermental Neurology* 115:121-126.

Lenneberg, E.H. (1967) *Biological Foundations of Language*. Wiley.

Lightfoot, D. (1989) The child's trigger experience: Degree-0 learnability. *Behavioral & Brain Sciences* 12:321-375.

Lightfoot, D. (1991) *How to set parameters: arguments from language change*. Cambridge, MA.: MIT Press.

Lubke,J. & Albus, K. (1989) The postnatal development of layer VI pyramidal neurons in the cat's striate cortex, as visualized by intracellular Lucifer yellow injections in aldehyde-fixed tissue. *Brain Research. Developmental Brain Research* 45:29-38.

Lund, J.S., & Holbach, S.M. (1990a) Postnatal development of thalamic recipient neurons in the monkey striate cortex: I. Comparison of spine acquisition and dendritic growth of layer 4C alpha and beta spiny stellate neurons. *Journal of Comparative Neurology* 309:115-28.

Lund, J.S., Holbach, S.M., & Chung, W.W. (1990b) Postnatal development of thalamic recipient neurons in the monkey striate cortex: II. Influence of afferent driving on spine acquisition and dendritic growth of layer 4C spiny stellate neurons. *Journal of Comparative Neurology* 309:129-40.

Noback, C.R., & Purpura, D.P. (1961) Postnatal ontogenesis of neurons in cat neocortex. *Journal of Comparative Neurology* 171:291-308.

O'Rourke, N.A., Cline, H.T. & Fraser, S.E. (1994) Rapid remodeling of retinal arbors in the tectum with and without blockade of synaptic transmission. *Neuron* 12:921-34.

O'Rourke, N.A., & Fraser, S.E. (1986) Dynamic aspects of retinotectal map formation revealed by a vital-dye fiber-tracing technique. *Developmental Biology* 114:265-76.

O'Rourke, N.A. & Fraser, S.E. (1990) Dynamic changes in optic fiber terminal arbors lead to retinotopic map formation: an in vivo confocal microscopic study. *Neuron* 5:159-171.

Osherson, D.N., Stob, M., & Weinstein, S. (1986) *Systems that Learn*. MIT Press.

Pallas, S.L., Roe, A.W. & Sur, M. (1990) Visual projections induced into the auditory patheay of ferrets. . novel inputs to prmary audtory cortex (AI) from the LP/pulvinar complex and the topography of the MGN-AI projection. *Journal of Comparative Neurology* 298:50-68

Parnavelas, J.G. & Uylings, H.B. (1980) The growth of non-pyramidal neurons in the visual cortex of the rat: a morphometric study. *Brain Research* 193:373-82.

Pearce, I.A., Cambray-Deakin, M.A. & Burgoyne, R.D. (1987) Glutamate acting on NMDA receptors stimulates neurite outgrowth from cerebellar granule cells. *Febs Letters* 223:143-7.

Petit, T.L., LeBoutillier, J.C., Gregorio, A. & Libstug, H. (1988) The pattern of dendritic development in the cerebral cortex of the rat. *Brain Research*, 469, 209-219.

Pinker, S. (1979) Formal models of language learning. *Cognition* 1:217-283.

Pinker, S. (1994) *The language instinct*. W. Morrow and Co.

Piattelli-Palmarini, M. (1989) Evolution, selection and cognition: from "learning" to parameter setting in biology and in the study of language. *Cognition* 31:1-44.

Pinker, S. (1984) *Language learnability and language development*. Harvard University Press.

Pinker, S. (1989) *Language acquisition*. In Foundations of cognitive science. ed. M. Posner. MIT Press.

Pinto Lord, M.C. & Caviness, V.S. Jr. (1979) Determinants of cell shape and orientation: a comparative Golgi analysis of cell-axon interrelationships in the developing neocortex of normal and reeler mice. *Journal of Comparative Neurology* 187:49-69.

Platt, J.C. (1991) A resource-allocating network for function interpolation. *Neural Computation* 3:213-25.

Pomeroy, S. L., LaMantia, A.S., & Purves, D. (1990) Postnatal construction of neural circuitry in the mouse olfactory bulb. *Journal of Neuroscience* 10:1952-66.

Purves, D. & Hadley, R.D. (1985) Changes in the dendritic branching of adult mammalian neurones revealed by repeated imging in situ. *Nature* 315:404-406

Purves, D., Hadley, R.D. & Voyvodic, J.T. (1986) Dynamic changes in the dendritic geometry of individual neurons visualized over periods of up to three months in the superior cervical ganglion of living mice. *Journal of Neuroscience* 6:1051-1060.

Purves, D., Voyvodic, J., Magrassi, L. & Yawo, H. (1987) Nerve terminal remodeling visualized in living mice by repeated examination of the same neuron. *Science* 238:1122-1126.

Quartz, S.R. (1993) Nativism, neural networks, and the plausibility of constructivism. *Cognition* 48:123-144.

Quartz, S.R. & Sejnowski, T.J. (1994) Beyond modularity: neural evidence for constructivist principles in development. *Behavioral and Brain Sciences* 17:725-726.

Rakic, P., Bourgeois, J. P., Eckenhoff, M. F., Zecevic, N., & Goldman-Rakic, P. S. (1986) Concurrent overproduction of Synapses in diverse regions of the primate cerebral cortex. *Science* 232:232-235.

Rakic, P., Bourgeois, J.P., & Goldman-Rakic, P.S. (1994) Synaptic development of the cerebral cortex: implications for learning, memory, and mental illness. *Progress in Brain Research* 102:227-43.

Rakic, P. & Sidman, R.L. (1973) Weaver mutant mouse cerebellum: defective neuronal migration secondary to abnormality of Bergmann glia. *Proceedings of the National Academy of Sciences of the United States of America* 70:240-4.

Rall, W. (1964) Theoretical significance of dendritic trees for neuronal input-output relations. In: *Neural theory of modelling*, ed. R.F. Reiss. Stanford University Press.

Redding, N.J.,Kowalczyk, A. & Downs, T. (1993) Constructive higher-order network algorithm that is polynomial time. *Neural Networks* 6:997-1010.

Regehr, W.G., Connor, J.A. & Tank, D.W. (1989) Optical imaging of calcium accumulation in hippocampal pyramidal cells during synaptic activation. *Nature* 341:533-6.

Roe, A.W., Pallas, S.L., Hahm, J. & Sur,M. (1990) A map of visual space induced in primary auditory cortex. *Science* 250:818-820.

Roe, A. W., Pallas, S. L., Kwon, Y. H. & Sur, M. (1992) Visual projections routed to the auditory pathway in ferrets: receptive fields of visual neurons in primary auditory cortex. *Journal of Neuroscience* 12:3651-64.

Ruiz-Marcos. A. & Valverde, F. (1970) Dynamic architecture of the visual cortex. *Brain Research* 19:25-39.

Schade, J.P. & van Groenigen, W.B. (1961) Structural organization of the human cerebral cortex. I. Maturation of the middle frontal gyrus. *Acta Anatomica* 47:72-111.

Scheibel, A.B. (1993) Dendritic structure and language development. In*: Developmental neurocogniton: Speech and face processing in the first year of life*, e.d. B. de Boysson-Bardies. Kluwer Academic Publishers.

Schilling, K., Dickinson, M.H., Connor, J.A. & Morgan, J.I. (1991) Electrical activity in cerebellar cultures determines Purkinje cell dendritic growth patterns. *Neuron* 7:891-902.

Schlaggar, B.L. & O'Leary, D.D.M. (1991) Potential of visual cortex to develop an array of functional units unque to somatosensory cortex. *Science* 252:1556-1560

Schuman EM; Madison DV. (1991) A requirement for the intercellular messenger nitric oxide in long-term potentiation. *Science* 254:1503-6.

Shatz, C.J. (1990) Impulse activity and the patterning of connections during CNS development. *Neuron* 5:745-756.

Shatz, C.J. (1992) How are specifc connectons formed between thalamus and cortex. *Current Opinion in Neuorbiology* 2:78-82

Shatz C.J., Lindstrom, S. & Wiesel, T.N. (1977) The distribution of afferents representing the right and left eyes in the cat's visual cortex. *Brain Research* 131:103-16.

Shatz, C.J. & Stryker, M.P. (1978) Ocular dominance in layer IV of the cat's visual cortex and the effects of monocular deprivation. *Journal of Physiology* 281:267-83.

Shepherd, G.M. & Brayton, R.K. (1987) Logic operations are properties of computer-simulated interactions between excitable dendritic spines. *Neuroscience* 21:151-65.

Shin, Y. & Ghosh, J. (1995) Ridge polynomial networks. *IEEE Transactions on Neural Networks* 6:610-22.

Shoukimas, G.M. & Hinds, J.W. (1978) The development of the cerebral cortex in the embryonic mouse: an electron microscopic serial section analysis. *Journal of Comparative Neurology* 179:795-830.

Siegler, R.S. (1989) Mechanisms of Cognitive Development. *Annual Review of Psychology*, 40, 353-379.

Simonds, R.J. & Scheibel, A.B. (1989) The postnatal development of the motor speech area: a preliminary study. *Brain and Language* 37:42-58.

Sperry, R. (1963) Chemoaffinity in the orderly growth of nerve fiber patterns and connections. *Proceedings of the National Academy of Science* 50:703-710.

Stanfield, B.B. & O'Leary, D.D. (1985) Fetal occipital cortical neurones transplanted to the rostral cortex can extend and maintain a pyramidal tract axon. *Nature* 313:135-7.

Stryker, M. (1991) Activty-dependent reorganization of afferents in the developing mammalian visual system. In: *Development of the visual system*, ed. D. Lam & C. Shatz. MIT Press.

Stuart, G.J. & Sakmann, B. (1994) Active propagation of somatic action potentials into neocortical pyramidal cell dendrites. *Nature* 367:69-72.

Sur, M., Humphrey, A.L. & Sherman, S.M. (1982) Monocular deprivation affects X- and Y-cell retinogeniculate terminations in cats. *Nature* 300:183-5.

Sur, M., Garraghty, P.E. & Roe, A.W. (1988) Expermentially induced visual projections into auditory thalamus and cortex. *Science* 242:1437-1441.

Sur, M., Pallas, S. L. & Roe, A. W. (1990) Cross-modal plasticity in cortical development: differentiation and specification of sensory neocortex. *Trends in Neuroscience* 13:227-33.

Tieman, S.B., & Hirsch S. (1982) Exposure to lines of only one orientation modifies dendritic morphology of cells in the visual cortex of the cat. *Journal of Comparative Neurology* 211:353-362

Tooby, J. & Cosmides, L. (1992) The psychological foundations of culture. In *The adapted mind: Evolutionary psychology and the generation of culture*, ed. J.H. Barkow, L. Cosmides & J. Tooby. Oxford University Press.

Turner, A.M. & Greenough, W.T. (1985) Differential rearing effects on rat visual cortex synapses. I. synaptic and neuronal density and synapses per neuron. *Brain Research* 329:195-203.

Uylings, H.B.M., Kuypers, K., Diamond, M.C. & Veltman, W.A.M. (1978) Effects of differential environments on plasticity of dendrites of cortical pyramidal neurons in adult rats. *Experimental Neurology* 62:658-677.

Uylings, H.B.M., Van Eden, C.G., Parnavelas, J.G., & Kalsbeek, A. (1990) The prenatal and postnatal development of the rat cerebral cortex. In: *The cerebral cortex of the rat*, ed. B. Kolb & R.C. Tees. MIT Press.

Valiant, L.G. (1984) A theory of the learnable. *Communications of the ACM* 27:1134-1142.

Valiant, L.G. (1991) A view of computational learning theory. In*: Computation and Cognition: Proceedings of the First NEC Research Symposium*, ed. C.W. Gear. SIAM.

Valverde, F. (1967) Apical Dendritic Spines of the visual cortex and light deprivation in the mouse. *Experimental Brain Research*: 3:337-352.

Valverde, F. (1968) Structural changes in the area striata of the mouse after enucleation. *Experimental Brain Research* 5:274-92.

Valverde, F. (1971) Rate and extent of recovery from dark rearing in the visual cortex of the mouse. *Brain Research* 33:1-11.

Vercelli, A., Assal, F. & Innocenti, G. M. (1992) Emergence of callosally projecting neurons with stellate morphology in the visual cortex of the kitten. *Experimental Brain Research* 90:346-358.

Volkmar, F.R. & Greenough, W.T. (1972) Rearing complexity affects branching of dendrites in the visual cortex of the rat. Science 176:1445-1447. Wallace, C.S., Kilman, V.L., Withers, G.S. & Greenough, W.T. (1992) Increases in dendritic length in occipital cortex after 4 days of differential housing in weanling rats. *Behavioral and Neural Biology* 58:64-8.

Walsh, C. & Cepko, C.L. (1988) Clonally related cortical cells show several migration patterns. *Science* 241:1342-1345

Walsh, C. & Cepko, C.L. (1992) Widespread dispersion of neuronal clones across functional regions of the cerebral cortex. *Science* 255:434-440.

Walsh, C. & Cepko, C.L. (1993) Clonal dispersion in proliferative layers of developing cerebral cortex. *Nature* 362:632-5

Wexler, K, & Culcover, P. (1980) *Formal Principles of Language Acquisition*. MIT Press.

White, H. (1990) Connectionist nonparametric regression: multilayer feedforward networks can learn arbitrary mappings. *Neural Networks* 3:535-549.

Winfield, D.A. (1981) The postnatal development of synapses in the visual cortex of the cat and the effects of eyelid closure. *Brain Research* 206:166-171.

Wong, R.K., Prince, D.A. & Basbaum, A.I. (1979) Intradendritic recordings from hippocampal neurons. *Proceedings of the National Academy of Sciences of the United States of America* 76:986-90.

Wynne-Jones, M. (1993) Node splitting: a constructive algorithm for feed-forward neural networks. *Neural Computing and Applications* 1:17-22.

Zecevic, N., Bourgeois, J.P. & Rakic, P. (1989) Changes in synaptic density in motor cortex of rhesus monkey during fetal and postnatal life. *Brain Research. Developmental Brain Research* 50:11-32.