

Below is the unedited preprint (not a quotable final draft) of:

Quartz, S. & Sejnowski, T.J. (1997). The neural basis of cognitive development: A constructivist manifesto. *Behavioral and Brain Sciences* 20 (4): 537-596.

The final published draft of the target article, commentaries and Author's Response are currently available only in paper.

€ For information about subscribing or purchasing offprints of the published version, with commentaries and author's response, write to: journals_subscriptions@cup.org (North America) or journals_marketing@cup.cam.ac.uk (All other countries). €

THE NEURAL BASIS OF COGNITIVE DEVELOPMENT: A CONSTRUCTIVIST MANIFESTO

Steven R. Quartz and Terrence J. Sejnowski

Computational Neurobiology Laboratory,
and The Sloan Center for Theoretical Neurobiology,
The Salk Institute for Biological Studies,
10010 North Torrey Pines Rd.
La Jolla, CA 92037
steve@salk.edu

Howard Hughes Medical Institute,
The Salk Institute for Biological Studies,
and Department of Biology,
University of California, San Diego,
La Jolla, CA 92037.
terry@salk.edu

Keywords

neural development; cognitive development; constructivism; selectionism; mathematical learning theory; evolution; learnability.

Abstract

How do minds emerge from developing brains? According to "neural constructivism," the representational features of cortex are built from the dynamic interaction between neural growth mechanisms and environmentally derived neural activity. Contrary to popular selectionist models that emphasize regressive mechanisms, the neurobiological evidence suggests that this growth is a progressive increase in the

representational properties of cortex. The interaction between the environment and neural growth results in a flexible type of learning: "constructive learning" minimizes the need for prespecification in accordance with recent neurobiological evidence that the developing cerebral cortex is largely free of domain-specific structure. Instead, the representational properties of cortex are built by the nature of the problem domain confronting it. This uniquely powerful and general learning strategy undermines the central assumption of classical learnability theory, that the learning properties of a system can be deduced from a fixed computational architecture. Neural constructivism suggests that the evolutionary emergence of neocortex in mammals is a progression toward more flexible representational structures, in contrast to the popular view of cortical evolution as an increase in innate, specialized circuits. Human cortical postnatal development is also more extensive and protracted than generally supposed, suggesting that cortex has evolved so as to maximize the capacity of environmental structure to shape its structure and function through constructive learning.

1 Introduction

How do neural mechanisms participate in, or underlie, cognitive development? In what ways do cognitive and neural processes interact during development, and what are the consequences of this interaction for theories of learning? In short, how is the mind built from the developing brain?

Although these questions are at the heart of cognitive science, development still resists our attempts to understand it. To develop is to change, and therein lies the challenge. For the structures underlying development are not fixed; they undergo large-scale changes during cognitive skill acquisition. What is more, these changes may not be intrinsically determined; they may depend on interacting with the environment. If so, then the distinction between biological maturation---the brain unfolding according to its intrinsic schedule---and learning breaks down. Descriptions at both levels of explanation, the biological and the cognitive, would then be required in an integrated level of explanation.

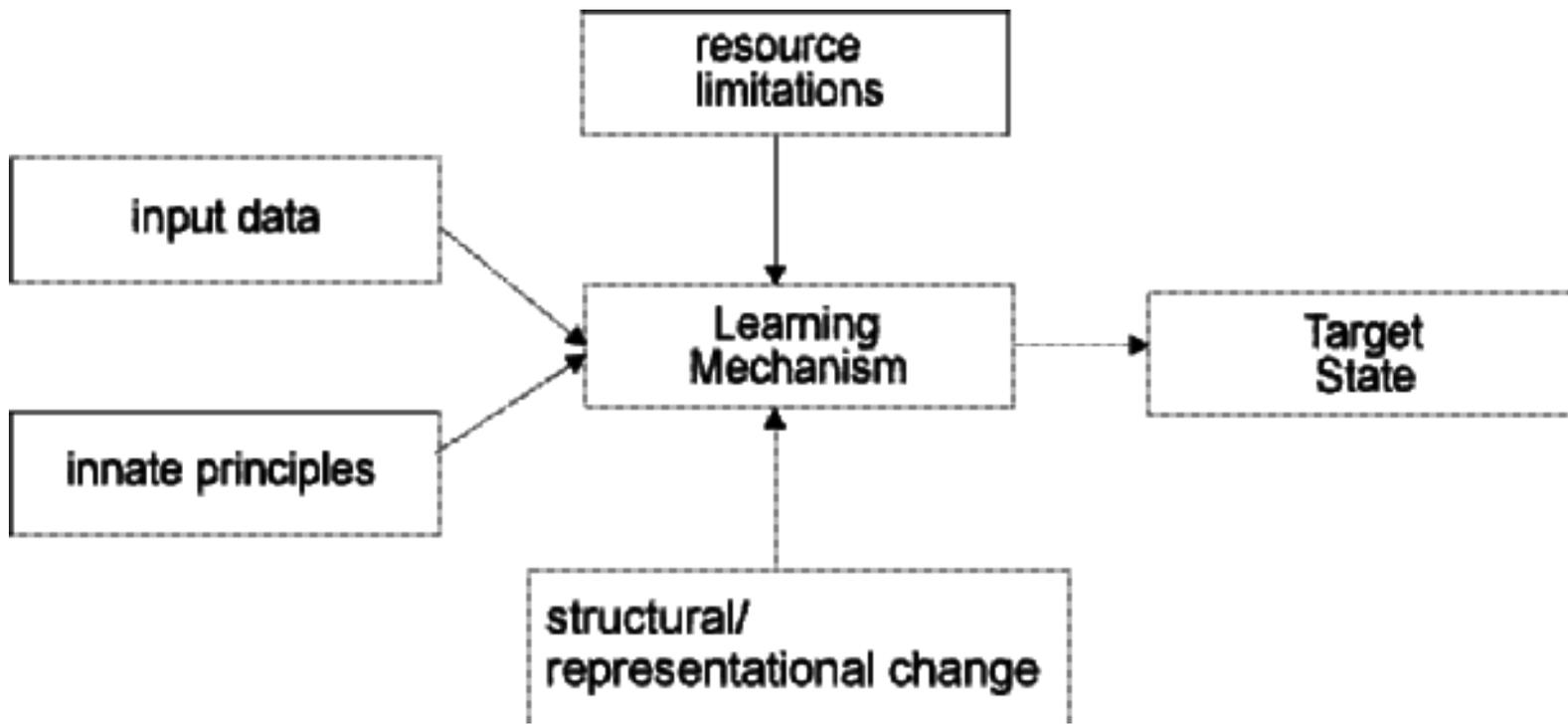


Figure 1: The logic of the problem of development. At its most abstract, the problem is defined as one of characterizing the mapping from some set of input data into the target state (the adult competence; see Wexler & Culicover (1980) for such a treatment). This problem is complicated by two elements that are often dismissed from such a characterization: changing resource limitations in terms of working memory and structural/representational change.

If cognitive and neural processes really do interact, then an added degree of complexity arises in analyzing development, for learning may induce large changes in the very structures involved in learning. This complicates matters, because now learning can alter what can be subsequently learned. To borrow a phrase from physics, systems with time-dependent properties are said to be "nonstationary" (e.g., Pinker 1979). This term applies to the "learning mechanism" or "acquisition device," roughly defined as the set of processes and structures that transform input data into the steady state representing the mature competence. The nonstationary learner is thus one in which learning causes large scale changes to its basic mechanisms. Pursuing a popular though perhaps misleading metaphor, learning causes major changes to the underlying hardware. Developmental theorists are accordingly confronted with the challenge of characterizing a nonstationary learning mechanism (Figure 1).

Because a nonstationary learning mechanism is difficult to explain, a typical response is to wish it away by invoking methodological principles minimizing change during development. Chomsky (1980), for example, idealized development as an instantaneous process, supposing that removing all temporal elements would have no effect on a developing system's acquisition profile. And, Pinker (1984), following Macnamara (1982), formulated the "continuity hypothesis," that children and adults should be viewed as qualitatively identical until we are forced to abandon this principle by some explanatory failure.

The possibility that learning guides brain development was also excluded from much of developmental psychology, although there are some important exceptions (e.g., Hebb 1949; Gottlieb 1976; Greenough et al.

1987). In his classic work on biology and language, for example, Lenneberg (1967) viewed the biological contribution as unfolding according to an intrinsic schedule. Interest in the neurobiological underpinnings of cognitive development thus waned. More, recently, however, a renewed excitement over the prospect of reuniting brain and cognitive development has begun to emerge. At its center is a vibrant developmental cognitive neuroscience (e.g., Johnson, 1990; Neville 1991; Karmiloff-Smith 1992 [see also BBS multiple book review: BBS (17) 1994]; Bates & Elman 1993; Plunkett & Sinha 1992; Elman et al. 1996). It is complemented by a resurgence of neural network research placing learning and the creation of internal representations once again in the fore of cognitive science (Rumelhart et al. 1986). Combined, these advances have led to the central assumptions of cognitive and computational neuroscience: that (1) meaningful and far-ranging connections can be made between cognitive and neural levels of description, and that (2) it is only through the mutual constraints that both levels impose on each other that a satisfactory theory can emerge (Churchland & Sejnowski, 1988).

1.1 A Neurocomputational Approach to Nonstationarity

In this target article, we examine how nonstationarity in development can be characterized from a neurocomputational perspective and we explore its implications for development. A difficult methodological problem must first be addressed, however. As already mentioned, nonstationarity is troublesome because it introduces time-dependent changes to the structures underlying acquisition. Since the currency of cognition is representations, the developmental problem is to characterize representational change in terms that correspond to structural changes in the learning mechanism. As we explore, the strength of cognitive neuroscience lies in its power to characterize just this sort of change.

Our method for studying nonstationarity in development is the following:

1. characterize the changes to the neural substrates underlying acquisition;
2. characterize the processes regulating these changes;
3. examine their correspondence to representational change;
4. explore their implications for the developing system's learning properties.

This methodology focuses on the neural basis of cognitive development. It has long been claimed that the dearth of neural constraints makes such an approach hopeless, but recent advances in developmental and computational neuroscience make it possible to characterize the learning mechanism structurally. This approach may provide a basis for understanding change in development with constraints that other developmental accounts have largely lacked.

Candidate Complextixy Measure

Cellular Component

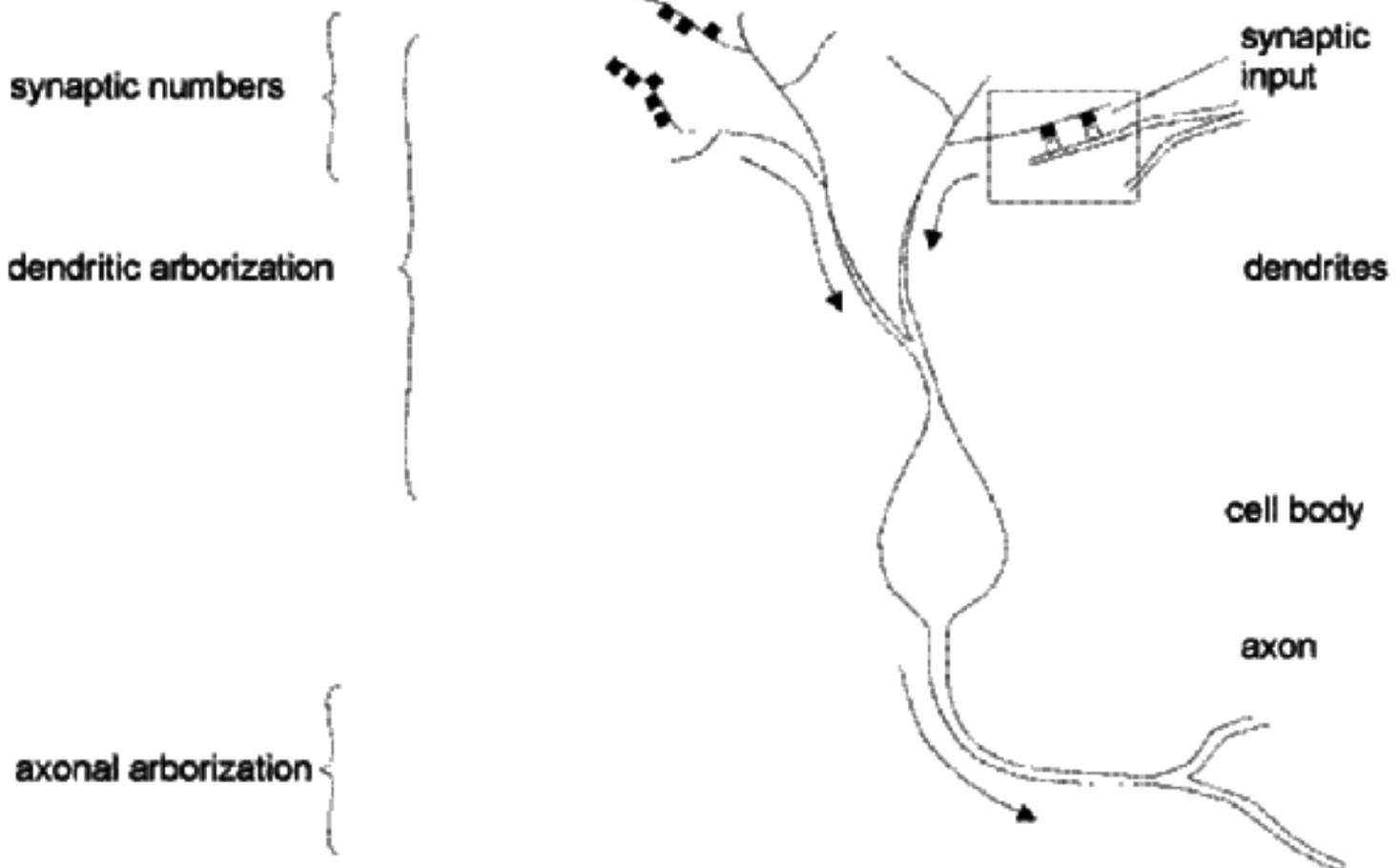


Figure 2: Candidate measures of structural/representational complexity. There are three possible (non-exclusive) measures: synaptic numbers, dendritic arborization, and axonal arborization. The figure also summarizes the basics of neural information processing elements: synaptic input travels along the dendrites to the cell body, where it is integrated and an action potential fires down the axon if some threshold is exceeded.

The first step is to provide an appropriate structural way to measure representational change. This is one of the primary aims of this target article. We will explore three possible measures for representational change: synaptic numbers, axonal arborization, and dendritic arborization (Figure 2). Applying the above methodology, we will then examine the neurobiology of these changes during development, the involvement of intrinsic and extrinsic factors in regulating them, and their adequacy as indicators of representational complexity.

Two themes emerge from finding a structural measure of representational complexity: (1) development is a progressive increase in the structures underlying representational complexity, and (2) this increase depends on interaction with a structured environment to guide development. These form the basis of *neural constructivism*, the developmental theory we present. This name reflects the Piagetian view that there is an active interaction between the developing system and the environment in which it is embedded. Like Piaget's theory, ours too emphasizes the constructive nature of this interaction, with representational structures progressively added during development. Our primary concern in this target article, however, is to

examine the neural processes regulating structural change and their implications for representational change. In particular, dendritic development fulfills important requirements for a nonstationary learning mechanism, suggesting how dendritic development under the influence of environmentally derived activity conforms to cognitive schemes for the construction of mental representations. Although cortical development is often thought limited primarily to the first two years of life, this development is far more extensive, prolonged, and progressive. This indicates that the extent of human cortical postnatal development has been widely underestimated along with its potential role in building mental representations under the guidance of environmentally derived information.

This environmentally-guided neural circuit building is a form of learning, "constructive learning," a unique and powerful response to the learning pressures confronting a developing system undermining the central assumptions of classical formal learning theory. The central problem confronting a cognitive system is to find an appropriate class of representations for specific problem domains. Many views suppose that these representations have to be pre-existing, but constructive learning builds these under the influence of the environment, acting alongside the general constraints that are imposed by the neural architecture. As a result, it offers powerful learning abilities while minimizing the need for domain-specific prespecification and so avoiding the heavy burden that nativism places on genetic mechanisms.

Ours is not a return to *tabula rasa* learning, however; learning is a dynamic interaction between a changing, structured environment and neural mechanisms. The neural machinery is extensively shaped by activity stemming from the environment, while its intrinsic properties also constrain this modulation and play an indispensable role in shaping the resulting structures. This interaction, however, is sufficient to determine the mature representational properties of cortex with no need for domain-specific predispositions somehow embedded *a priori* in the recipient cortex. As a consequence, this makes the relation between environmental changes---whether natural or cultural---and brain structure a direct one. This suggests an evolutionary perspective as a progression to more flexible representations, in contrast to evolutionary psychology (e.g., Barkow et al 1992; Pinker 1994).

The far-reaching interaction between environmental structure and neural growth mechanisms suggests that development has been in the grips of a misleading dichotomy. On the one hand, empiricists offer a general-purpose, computational architecture with generic learning procedures; on the other hand, rationalists offer domain-specific knowledge implanted in cognitive structures. It is striking how readily so much of the developmental literature falls into these two extremes. Neural constructivism rejects this dichotomy, positing general, intrinsic processes along with environmentally derived neural activity carrying domain specific-information. Together, through constructive learning, these two sources build the mature computational architecture and its representational properties. This interaction thus promises a rich account of development that integrates both cognitive and neural levels of description into a single framework, opening up new opportunities for collaboration between psychologists and neurobiologists.

2 Measures of representational complexity

The brain is above all a representational device (for a detailed discussion, see Pylyshyn 1984; Churchland & Sejnowski 1992). By "representation" we mean the neural encoding of environmentally-derived information and transformations resulting from the application of mental operations. The best-known account of mental

representation is in terms of language-like systems with a primitive lexicon and syntactic rules corresponding to mental operations (Chomsky 1980). Neural networks offer alternative representational encodings, particularly distributed representations (Churchland & Sejnowski 1992). Although representational complexity can be defined for both types of representations (Quartz 1993), neural network measures depend on structural properties, making the relationship between complexity and structure a direct one.

For development the first concern is the source of mental representations and second, the extent of representational change (Karmiloff-Smith 1992; Bates & Elman 1993). This latter concern brings us back to nonstationarity. While nonstationarity was minimized in cognitive theories such as Chomsky's, two neurobiologically-inspired views embrace nonstationarity: selectionism and neural constructivism. Neural constructivism belongs to the tradition starting with Hebb (1949) and taken up by Gottlieb (1976; 1991) and Greenough et al. (1987), who, rejecting a dichotomy between cognitive and neural, explored how learning guides the developing brain. A contrasting tradition began with Jerne (1967), who applied selectionist thinking to brain development, although the roots of this tradition go back to Wilhelm Roux's application of Darwinian principles to cellular interactions in 1881. Variants of selectionism have been defended by cognitive psychologists (e.g., Mehler 1985; Piatelli-Palmarini 1989), psycholinguists (e.g., Lightfoot 1989; 1991), and selectionist brain theorists (e.g., Jerne 1967; Changeux & Danchin 1976; Rakic et al. 1986; Edelman 1987).

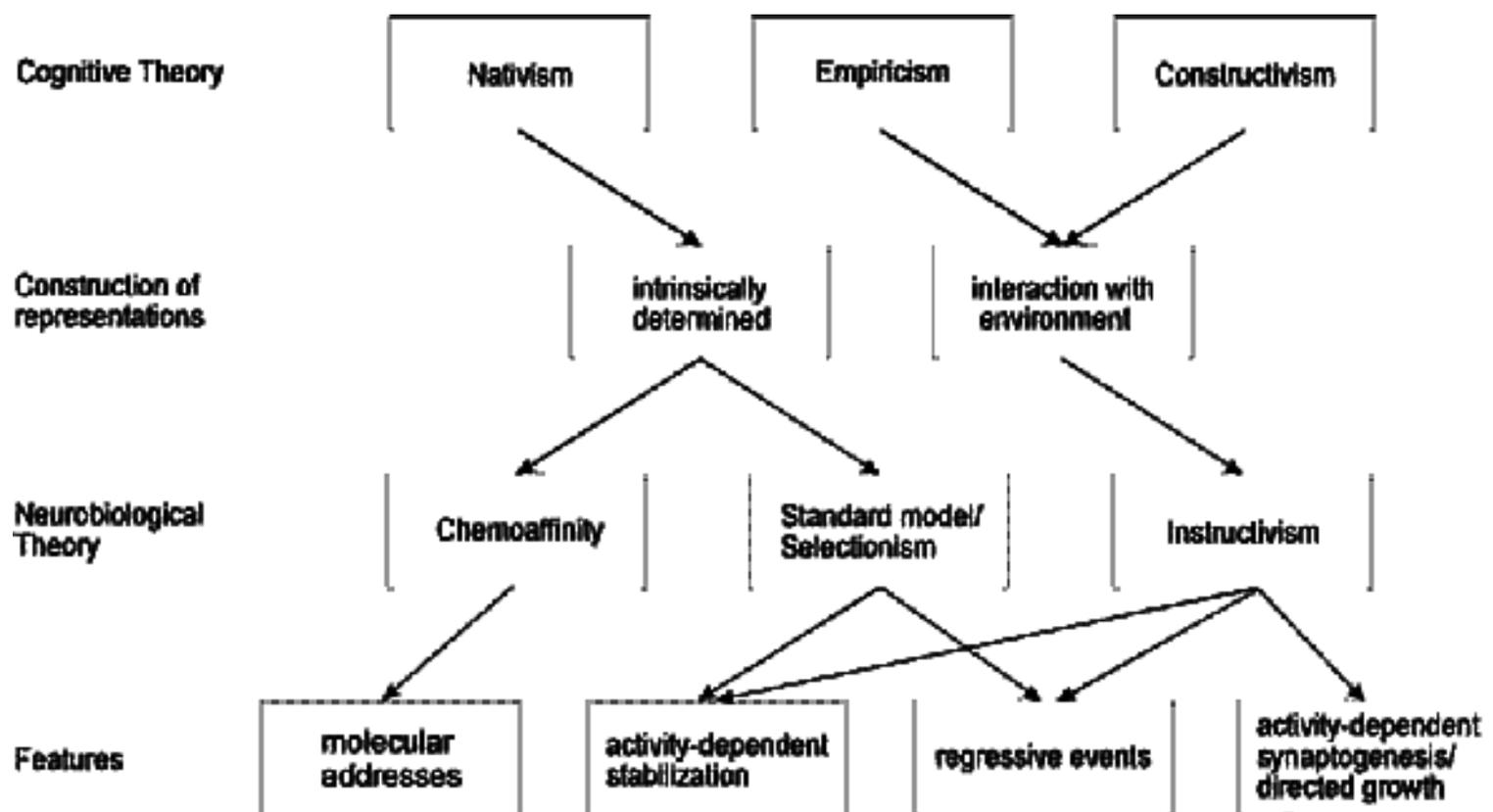


Figure 3: Summary of the relations among major cognitive and neural theories. Instructivism is often identified with Hebb's work, since on his view the environment guides or instructs brain growth. We refer to such theories as constructivist since the latter is a more widely

recognized class of theories.

Selectionism takes its inspiration from immunology and population biology, with subsets of populations being selected on the basis of fitness criteria. Selectionism divides development into 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. The goal of the latter stage is to pick out those representations with the highest "fitness" to underlie mature skills (Figure 3 presents a summary of developmental positions). Although not identified with selectionism, a popular view of neural development stemming from Hubel and Wiesel's work on the visual system also emphasizes development as the selective elimination of exuberant initial growth (Wiesel & Hubel 1962; 1963; 1965; Shatz & Stryker 1978; LeVay et al. 1980). To avoid a proliferation of terminology, we will refer to any view positing an initial exuberant growth followed by elimination as "selectionism," keeping in mind that it comes in different strengths.

Selectionism's learning mechanism operates by reducing an overly complex set of representations to a subset in the mature state. In an important selectionist statement, Changeux and Dehaene (1989, p.82) bluntly put it this way: "activity does not create novel connections, but, rather, contributes to the elimination of pre-existing ones." Indeed, completing the evolutionary analogy, allowing for directed growth in development would be akin to allowing for Lamarckian processes in evolution. For selectionism, then, development marks a reduction in representational complexity. In contrast, neural constructivism sees development as a progressive increase in representational complexity. Since these differ over changes in representational complexity, it is important to ask, what do these accounts use as a measure of representational complexity?

There are three main candidates for this measure: synaptic numbers, axonal arborization, and dendritic arborization (Figure 2). Although these are all related, it is worthwhile to examine them separately, as a change in one does not necessarily imply a change in the others. In the remainder of this section, we consider the support for changes in these measures along with their adequacy as measures of representational power.

2.1 Synaptic numbers

2.1.1 Synaptic numbers over development

Synapses, the anatomically defined connections between brain cells, are probably the most widely used measure of brain complexity. Studies of structural plasticity after learning, for example, typically hold that "more is better" (e.g., Turner & Greenough 1985). Developmental neurobiology has made a major effort to examine changes in synaptic numbers quantitatively. This effort was made despite the technical difficulties such studies present, both in obtaining unbiased counts (Coggeshall 1992; Coggeshall & Lekan 1996) 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 quantitative studies, that of Rakic et al. (1986) has been the most influential. Taking a brute force approach, they measured synaptic density in motor, somatosensory, prefrontal, visual, and limbic

cortices of the rhesus monkey at a number of developmental ages. A striking finding was that synaptic density levels changed uniformly across all regions they examined, with a peak between two and four months of age, followed immediately by a dramatic decline. Their interpretation of these findings was that a single, intrinsic signal regulated cortical synaptogenesis.

This conclusion was eagerly adopted by selectionists, who took it as strong evidence for their view of an initial creation of prerepresentations followed by selective loss. As Rakic et al. (1986, p.234) state, "if experience alters synaptic number during development it does so by causing selective survival of certain synapses, not by regulating their initial formation." A limitation of this study, however, was that synaptic density was not related to particular cell types. If particular cell types matured at different rates, only a direct examination of their development would make these differences apparent. Indeed, the study of Lund et al. (1990a), for example, found that the synapses of cells in adjacent sublamina of visual cortex (layer 4c alpha and beta stellate cells) reach various landmarks almost 4 weeks apart. This indicates that even nearby cells do not necessarily follow a similar, intrinsic developmental timetable.

What could make these neighboring cells mature at different rates? Although near one another, they receive input from distinct pathways, the magno- and parvo-cellular, which carry two different types of information about the visual world from the retina to the cortex, with different spatial resolutions and sensitivities to contrast and motion (see Churchland & Sejnowski 1992 for a review). Under various deprivation conditions minimizing differences in the activity of these pathways, the development of these adjacent cells was more synchronous (Lund et al. 1990b). This suggests that differences in input activity, and not just intrinsic mechanisms, play an important role in regulating cellular maturation, a theme we explore in detail in section 2.1.2.

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. With a larger sample, 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). It is hard to reconcile this finding with the claim that selective synapse elimination underlies cognitive development since no such process appears to operate in the prefrontal cortex during the most formative years of cognitive development. Indeed, an additional complication comes from studies showing that brain volume increases during this period, particularly in prefrontal cortical areas (see Dekaban & Sadowsky 1978; Jernigan et al. 1991 and references therein). Even if synaptic density remains constant, these volume increases imply synapse addition.

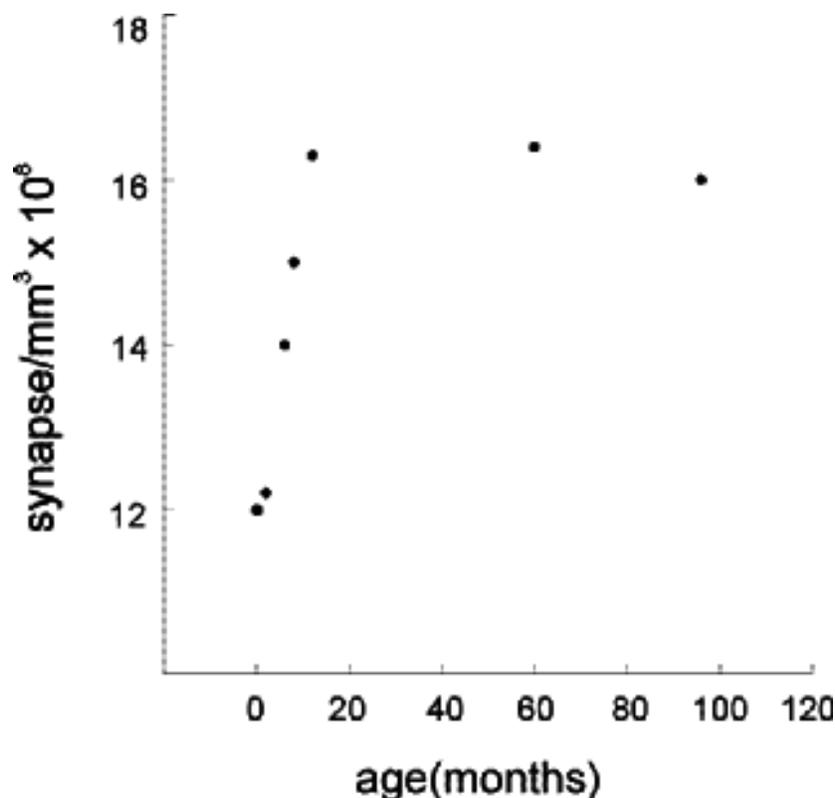


Figure 4: Synaptic density in layer III, human middle frontal gyrus from birth to 8 years. Data from Huttenlocher (1979). Note the absence of data points between 1 and 5 years of age.

Human studies analogous to Rakic et al.'s have been influential in forming a selectionist view of human development, one that has been particularly popular in psychology (e.g., Siegler 1989). The actual data on human cortical development, however, are scarce. In fact, the only published quantitative study of synaptic measures in developing human non-visual cortex is Huttenlocher's (1979) widely cited one (Figure 4). This work is important and more such basic human neuroanatomical work is certainly needed. It is worthwhile, however, to point out some of this study's limitations. Huttenlocher examined a total of 21 brains (with data only for layer III of middle frontal gyrus) from subjects ranging in age from newborn to 91 years of age (without gender distinctions). The major limitation, though, in attempting to infer general timetables of human prefrontal development, is the lack of samples between the ages of 1-5 and 8-15. This alone severely weakens a selectionist interpretation, as the ages of interest are not represented in the data sample (Huttenlocher (1990) himself raises many of these caveats, but they do not seem to have been heeded). From what data there are, however, no significant differences in synaptic density measures emerge between one year and eight years of age. This study, therefore, provides no evidence for a selectionism.

Some of the confusion surrounding claims of selective human cortical development can be clarified by the study of Bourgeois et al. (1994; see also Rakic et al. 1994). Matching their data with conceptual ages in the human data, Bourgeois and colleagues suggest that human prefrontal synaptic development does not undergo any significant reduction before puberty. This, then, renders the claim that selective elimination underlies cognitive skill acquisition highly problematic. Instead, to the degree that there is a reduction in synaptic density, it takes place well after the formative years of cognitive development.

We conclude that although these results do not imply that synaptic overproduction is not real (see O'Leary

1992 for a review of pruning in development), its generality has been overstated. The link between cognitive development and synaptic elimination in cerebral cortex is questionable. And, as we consider below, its intrinsic source of structure is likewise overstated.

2.1.2 Environmental conditions and synaptic numbers

The study of Lund et al. (1990b) suggests that differences in input activity might underlie differences in the growth rates of two nearby cell types. This raises the question of whether activity-dependent mechanisms are involved in establishing neuronal structure. A major source of evidence for this role comes from differential rearing/environment studies. These are important for distinguishing between competing models of development as they can distinguish between activity-dependent and independent modes of growth throughout the lifespan. Being able to assess environmental effects across the lifespan is important because it rules out the possibility that these changes reflect only stabilization of already-present structure. In some cases, this research also correlates specific structural changes with a single manipulated environmental parameter (e.g., Tieman & Hirsch 1982).

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.

Early work examined the effects of differential rearing conditions on synaptic numbers and density (reviewed in Boothe et al. 1986). Systematic structural changes were found to mirror changes in the animal's environment (summarized in Table 1). Of particular interest was Valverde's (1971) finding that these effects were largely reversible.

This paradigm has also been used to examine whether the number of synapses changes in mature forms of plasticity (Table 1). Black et al. (1990) have shown that the formation of new synapses accompanies motor learning tasks in behaving animals. To distinguish between the possibility that motor activity and not motor learning caused the increases in synaptic number, a control group underwent exercise on tasks that were easily mastered and required little learning. In these cases, there were no significant increases in cortical synapses. There was, however, angiogenesis (increased capillary support), as one would expect from increased demands. In contrast, cerebellar Purkinje cells in rats that learned complex locomotor tasks showed a 25% increase in synaptic numbers.

This line of work indicates that an important activity-dependent component in synaptic development remains as a lifetime capacity. Biological systems often conserve useful mechanisms; this appears to be another example of a mechanism that had an important developmental role that was then utilized in mature forms of plasticity (learning).

2.1.3 Synapse number as a measure of circuit complexity

Although there is good evidence for activity-dependent synapse formation, there are a number of reasons why synaptic numbers alone are an inadequate measure of representational complexity. What is the relation between synaptic numbers and a cell's computational or representational power? For this relation to be direct, many functional assumptions must be made about a cell's signal integration properties. In connectionist models, for example, there is a direct relation between the number of connections and a network's complexity. Each input and weight effectively serves as a parameter or degree of freedom because the connectionist unit is a "point neuron". This idealization abstracts away the spatial properties of integration and possible nonlinear conductance properties---the spatial arrangement of the units plays no role in processing.

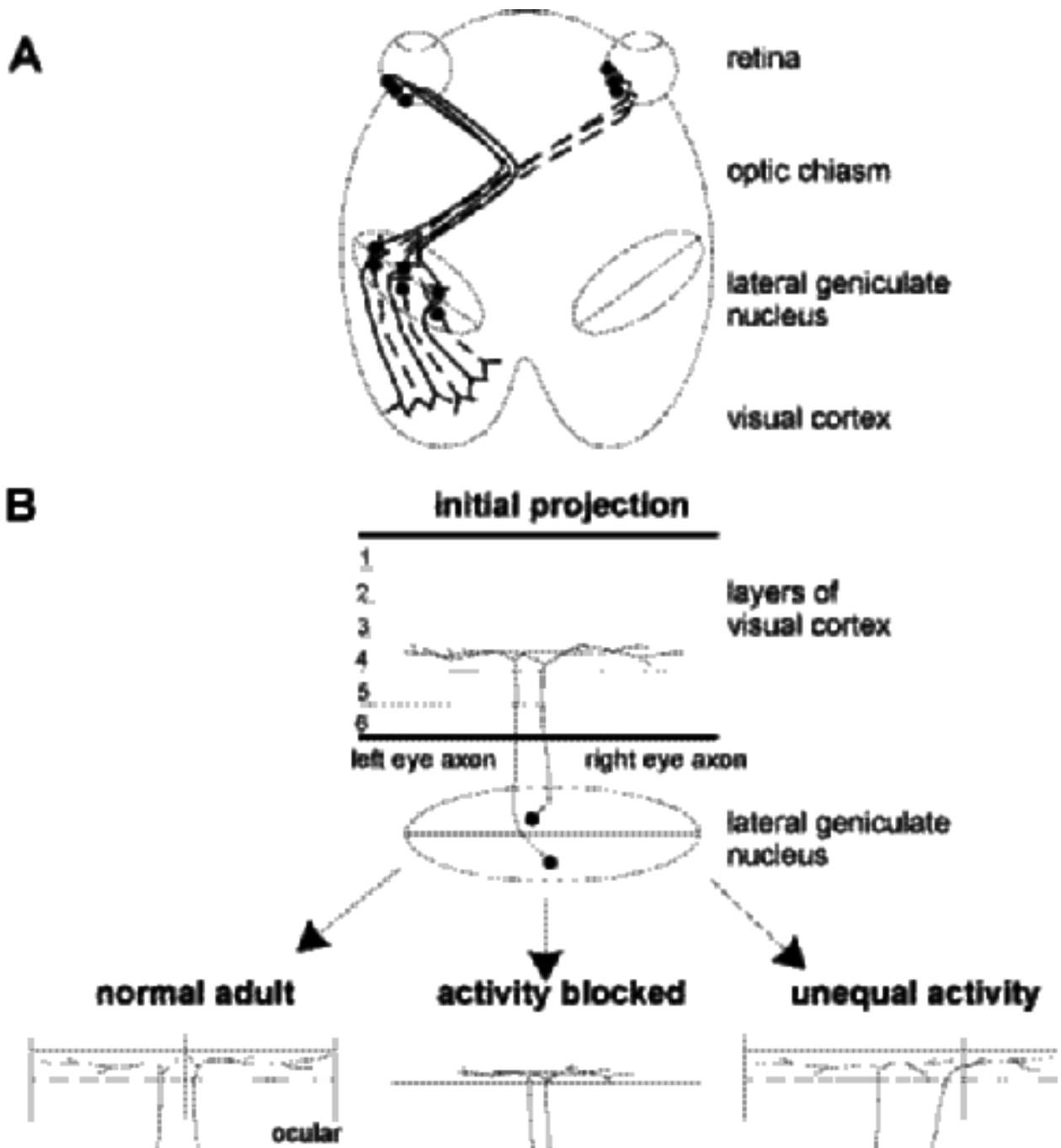
In many real neurons the spatial arrangement of pre- and postsynaptic elements is thought to be crucial to

their response properties. One reason for this is the presence of active conductance properties in the cell's membrane; these amplify or otherwise change the incoming signal in nonlinear ways. Nonlinear dendritic conductance properties, now well established (Stuart & Sakmann 1994; Mainen et al. 1995), 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. We consider these two measures below.

2.2 Axonal arborization

Patterns of axonal arborization have been widely used as a measure of representational complexity. Indeed, studies of axonal arborization have largely contributed to selectionism.

2.2.1 Axonal arborization over development



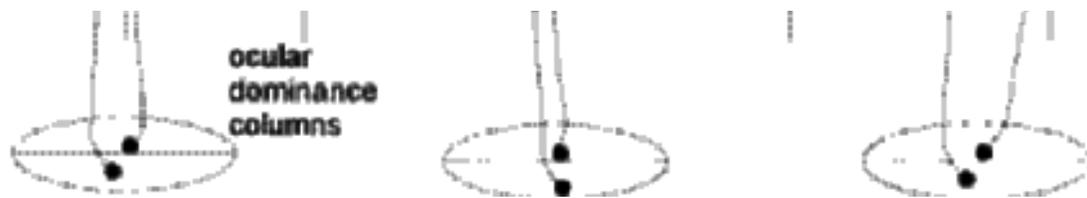


Figure 5: The generation of specificity according to the standard model. A, schematic representation of the connections between retinal ganglion cells and the lateral geniculate nucleus of the thalamus and connections to primary visual cortex. B, the two-phased process of development and perturbations 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).

The electrophysiological, developmental, and anatomical study of early vision is central to modern neuroscience. In particular, the retinogeniculocortical pathway, and especially the development of ocular dominance (OD) columns of layer IV, primary visual cortex have been important sources for selectionism (Wiesel & Hubel 1962; 1963; 1965; Shatz & Stryker 1978; LeVay et al. 1980). Ocular dominance columns are alternating, anatomically defined regions of input from primarily one eye (OD development is summarized in Figure 5). As Antonini and Stryker (1993, p.3549) note, two hypotheses regarding their development have been suggested. One, conforming to selectionism, emphasizes two phases in OD development: a period of exuberant axonal growth followed by selective axonal pruning. The other, more constructivist, hypothesis emphasizes the general expansion of axon collaterals alongside selective pruning.

The early studies on OD formation used a technique known as autoradiography (Hubel & Wiesel 1972; Shatz et al. 1977): a sugar or amino acid with an attached radioactive label is injected into the eye, whence it is transported into the visual system, thereby marking its path. Autoradiographs, pictures of the pattern of radioactivity, revealed an originally diffuse pattern that became segregated into periodic structures at later developmental stages.

While the evidence from autoradiography suggested a selectionist interpretation, as did evidence from other developing systems, such as the neuromuscular junction (Bennett & Pettigrew 1974), autoradiography has a poor spatial resolution, limiting the ability to identify cellular components. Because of these technical limitations, it would be desirable to have more direct anatomical evidence, including studies of identified arbors.

Because of the small diameter of geniculocortical axons during development, they are very difficult to intracellularly fill. For this reason, there have been few direct studies of individual cells. The lack of such studies, along with other systems appearing to support selectionism (e.g., the neuromuscular junction), has led to the common interpretation that OD development fits selectionism, an interpretation particularly popular among neural modelers and cognitive scientists (e.g., Swindale 1980; Linsker 1986; Miller et al. 1989). Selectionism brings increased tractability to modeling efforts since the initial state of a network may be set to correspond to the end of the period of exuberant growth. The dynamics of directed growth and the instability it brings can then be ignored. Instead, a Hebbian rule can operate on exuberant structures to pick

out the appropriate final structure. The simplest Hebbian rule (see Sejnowski & Tesauro (1989) for discussion) is a type of correlational learning rule in which the temporal coincidence of pre- and post-synaptic activity results in a strengthening of that synapse. Further conditions, supposing that anti-correlations weaken connections (or various normalization assumptions, which have the same effect), provide a sufficient set of mechanisms to drive progressive segregation of initially overlapping arbors. The discovery of the NMDA-receptor, which seemed to have the right kinetics to implement the Hebbian rule, also generated a great deal of excitement and offered the possibility that this economical rule could underlie important facets of neural development (reviewed in Shatz 1990).

The actual view that emerged from the experimental work, however, is quite different and supports neural constructivism. Although the important early study of LeVay and Stryker (1979) contains a pair of Golgi-stained cells at early and late developmental stages that are often cited as evidence for selectionism, the authors also make it clear that arbors increase in complexity throughout development, as does the entire retinogeniculate pathway (see Table 2). More recently, Antonini and Stryker (1993) confirm this more constructivist interpretation, and place even more emphasis on directed growth than the earlier studies (see p. 3572 for discussion). Using a technique known as bulk-filling and anterograde transport, they found some retraction of early axonal structure, but they also found substantial progressive increases in axonal measures throughout development. During the first week of segregation between left and right eye input, axonal branch point numbers increase strikingly from a median value of 34.5 to one of 96, while total arbor length increased from 7538 to 12848 microns (Antonini & Stryker 1993). Similarly, the area an arbor occupied also increased during this time, becoming significant by postnatal day 39. Even this is still about half the coverage area of the mature axonal arbor (Humphrey et al. 1985).

Anderson et al. (1992) have managed to fill intracellularly a small sample of physiologically characterized, geniculocortical cells in developing kitten. Although the youngest cells they could examine were near the end of the major events in OD formation, they found no evidence for initial exuberance of X-pathway axons; instead, axonal arbor area fell on the lower range of adult arbor areas, consistent with their more detailed studies in a higher visual area, area 18 (see Table 2).

Hence, although selectionism has dominated modeling work, the experimental work points to more balance between both selective elimination and selective growth, with a progressive increase in the complexity of the arbors underlying OD formation. As Antonini and Stryker conclude (1993, p.3572), "[N] development thus appears to involve both selective elimination of widely extended branches and considerable growth and elaboration."

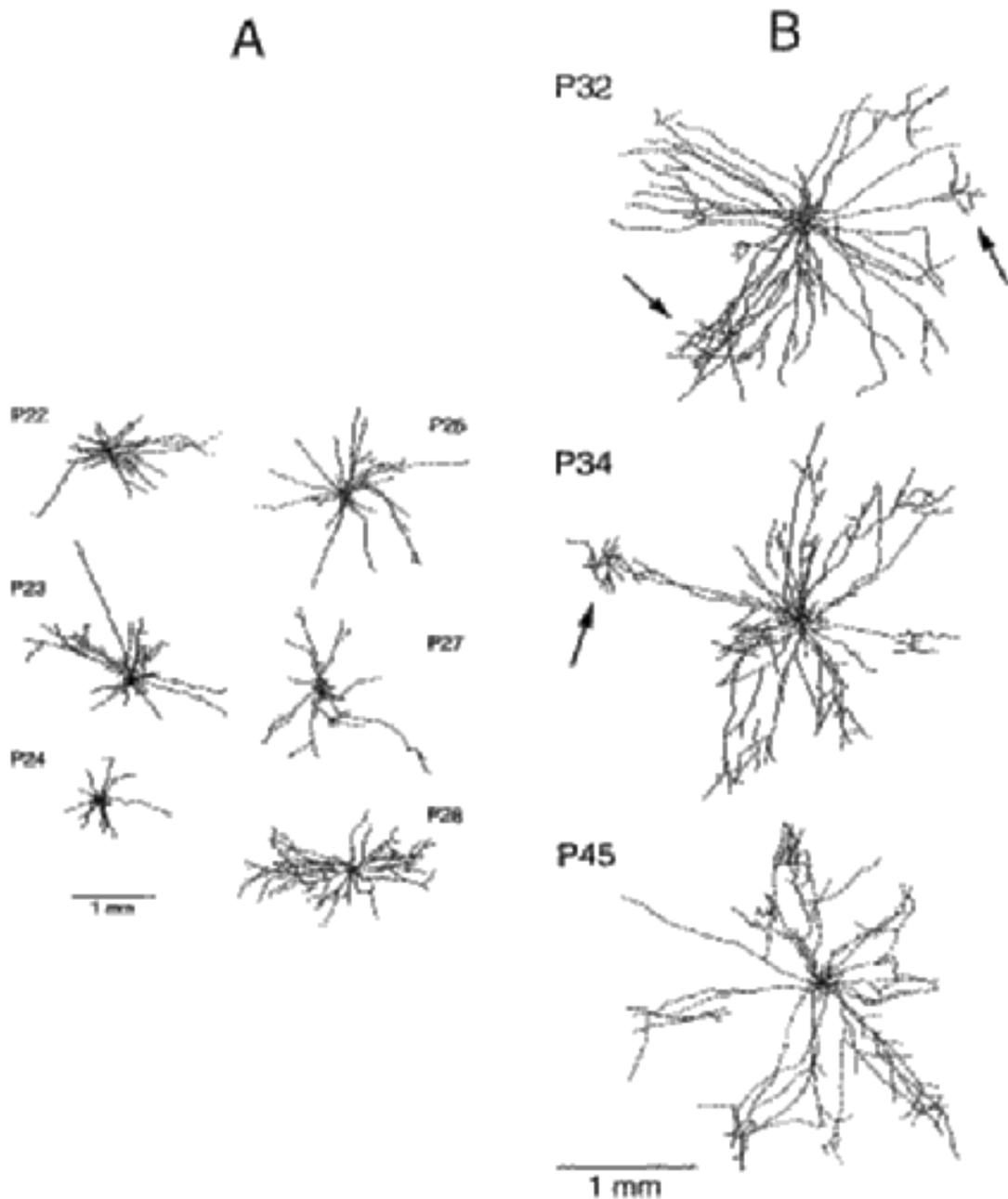


Figure 6: The development of horizontal connections in the ferret. A, At postnatal day 22 (P22) little clustering can be seen. As development proceeds, however, axon cluster formation become evident, particularly by P28. B, Subsequent development, seen here under higher magnification, reveals the refinement and elaboration of axon clusters. The development of these horizontal connections supports the hypothesis that axon development involves both selective retraction of inappropriately placed arbors and progressive elaboration of correctly placed arbors. Reproduced by permission from Durack & Katz (1996).

Similarly, in describing the development of intrinsic horizontal connections in area 17 (see Figure 6), Callaway and Katz (1990, p.1151) state that "after the basic axonal framework occupies only appropriate regions, axons then branch extensively and specifically within appropriate regions to attain their final

adultlike morphology." They also suggest that this is likely to be a general developmental theme. Indeed, as Purves and Lichtman (1985, pp.279-80) note in a classic text on developmental neurobiology, there is a general increase in the complexity of axons within the other model systems commonly interpreted as underlying selectionism: the development of the neuromuscular junction and the parasympathetic ganglion (Lichtman 1977).

The experimental evidence of selective growth of correctly positioned arbors further supports the role of activity in constructive processes of development. Even long before OD columns have begun to segregate, activity is required for the initial outgrowth of thalamic axons into layer IV, as Herrmann and Shatz (1995) demonstrated. They state (1995, p.11245):

activity is required far earlier in the development of thalamocortical connections than previously supposed: blockade of activity patterns prevents thalamic axons from elaborating their initial terminal branches in layer 4. Indeed, there is so little branching that it is as if the thalamic axons had failed to recognize this cortical layer as their appropriate target.

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.

Thus, activity-dependent outgrowth plays a central role in this study at the earliest stages of axonal growth in one of the best studied model systems.

The experimental work in OD formation suggests an important avenue of research that needs to be explored: activity-dependent rules that add structure during development. What support for selectionism comes from other areas of development?

2.2.2 Environmental conditions and axonal development

Many of the studies summarized in Table 2 also point to an important activity-dependent component in axonal development. Among these, the Friedlander et al. (1991) study is an important one, as it examined the effects of monocular deprivation on Y-type geniculocortical axons found in cats. According to selectionism, monocular deprivation causes an expansion of columns of the non deprived eye. This expansion is caused by relatively more correlated activity from the non deprived eye, and so its arbors compete favorably for territory that is initially shared by overlapping arbors. The deprived eye columns, in contrast, shrink (see Figure 5). Friedlander et al. (1991), however, found that the deprived arbors did not shrink due to a lack of stabilizing input, but instead failed to grow in the first place. The nondeprived arbors did not simply stabilize over more territory in the absence of competing fibers. Rather, their growth was accelerated and extended.[\[1\]](#)

These studies, then, point to the central role of activity in the progressive growth of these systems. We next examine the third candidate measure, dendritic arborization.

2.3 Dendritic development

Dendritic development has typically been overlooked in theoretical models of development, which focus instead on axonal development. We suggest that dendrites play an important role in determining the representational properties of cortex and that their development supports neural constructivism. First, we will consider some general reasons why dendritic development is important for the representational properties of cortex.

As we noted in section 1, nonstationarity, while perhaps increasing the learning capacities of a developing system, introduces a number of complications. The main reason for this is that changes to the underlying architecture can be unwieldy. Even small architectural changes can have severe 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 a nonstationary mechanism must satisfy:

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.

2.3.1 *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 cause large-scale changes in the overall pattern of connectivity. Hence, the constructivist learning scheme operates at the level of individual dendritic segments, rather than the whole cell, thereby building in these conditions.

It is also important that dendrites grow much more slowly than do axons. Axons grow at approximately 500 microns/day compared to 15-35 microns/day for dendrites (see Uylings et al. 1990; Katz et al. 1984). This suggests that the two are sensitive to statistical structure at different timescales and that dendrites are extracting this structure on a much longer temporal scale.

A more important and general reason for examining the growth of dendrites in relation to the construction of representations is that dendrites are the primary receptive surface of a neuron. Moreover, since dendrites do not just conduct passively, but amplify the incoming signal in nonlinear ways, their processing properties make them central to how information is processed by neural systems. It is essential, then, to assess the developmental processes that shape dendritic form and the role of these processes in determining the representational properties of neural circuits. [\[2\]](#)

2.3.2 *Dendritic arborization*

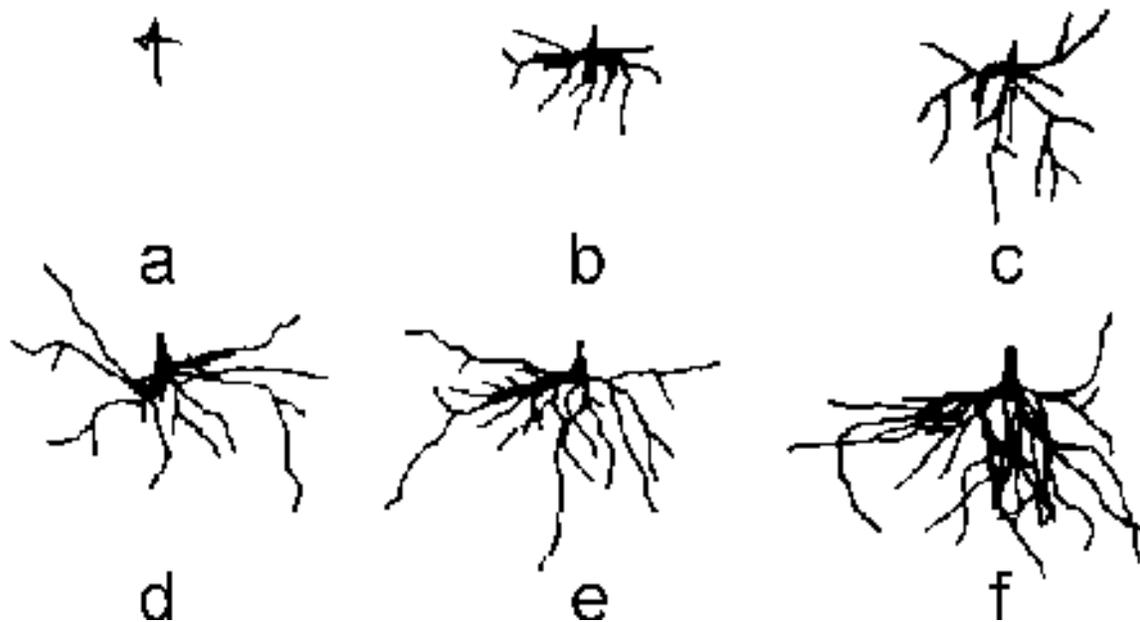


Figure 7: Camera lucida drawings of basal dendrites of layer V human pyramidal cells; a, newborn; b, 3 months; c, 6 months; d, 15 months; e, 24 months; f, adult (from Schade & van Groenigen 1961).

The dendrites of cortical neurons are highly varied, even among cells within a single class. This variety was already apparent to Golgi and Cajal in the 19th century, but recent techniques for staining the arbors of electrophysiologically characterized neurons at different stages of cortical development have revealed that dendrites are dynamic structures whose growth can be affected by many factors. As dendrites grow, the integration of synaptic activity is altered in ways that depend on the geometry of the branches and the placement of synapses (Segev & Davis 1995 contains an excellent collection of classic and contemporary views on dendritic function).

Figure 7 shows a striking example of neural constructivism in human development. 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, this cell's dendrites expand over ten times to a total length of over 2000 microns, or 2 mm. By two years of age, this cell's dendrites add another 1000 microns. Perhaps most startling, even by two years of age, these dendrites have not yet reached their halfway point, but will more than double between two years and adulthood to a mature total length of nearly 7000 microns (Schade & van Groenigen 1961). 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. It should be borne in mind that although there is a reduction in synapses per micron of dendritic length, the increasing dendritic surface area of these cells implies an overall increase in the number of synapses.

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 m 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 m 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. Although there is little doubt that regressive events also occur in dendritic development (e.g., Koester & O'Leary 1992; Vercelli et al. 1992), the above examples motivate the search for the processes regulating dendritic development.[\[3\]](#) In the following sections, we accordingly examine the mode of dendritic development in some detail---the extent of progressive processes at the level of dendritic structure and their malleability by changes in activity. From this, we go on to formulate some features of dendritic development, considering their cellular basis, and relating these to the learning and representational properties of cortex.

2.3.3 Environmental conditions and dendritic development

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.

Unlike axons, which in many cases begin to grow during migration (Shoukimas & Hinds, 1978), dendrites typically do not begin to differentiate until they complete their migration and their 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).

Table 4 summarizes some studies on the effects of manipulating input pathways on dendritic development. One of the earliest is Valverde (1968), which examined the effects of enucleation on stellate cell dendrites in mouse visual cortex. As Figure 8a illustrates, in normal development cells outside layer IV, the location of the primary thalamic projection, extend dendrites throughout layers III, IV, and V. In contrast, as Figure 8b illustrates, in enucleated animals, cells outside layer IV did not project their dendrites into that layer. Instead, they directed their dendrites toward layers III and V, as though they were looking for afferent input outside layer IV. Valverde (1968 p.290) concludes, "dendrites are not passive structures but actively growing neuronal formations which must accommodate to changing functional demands."

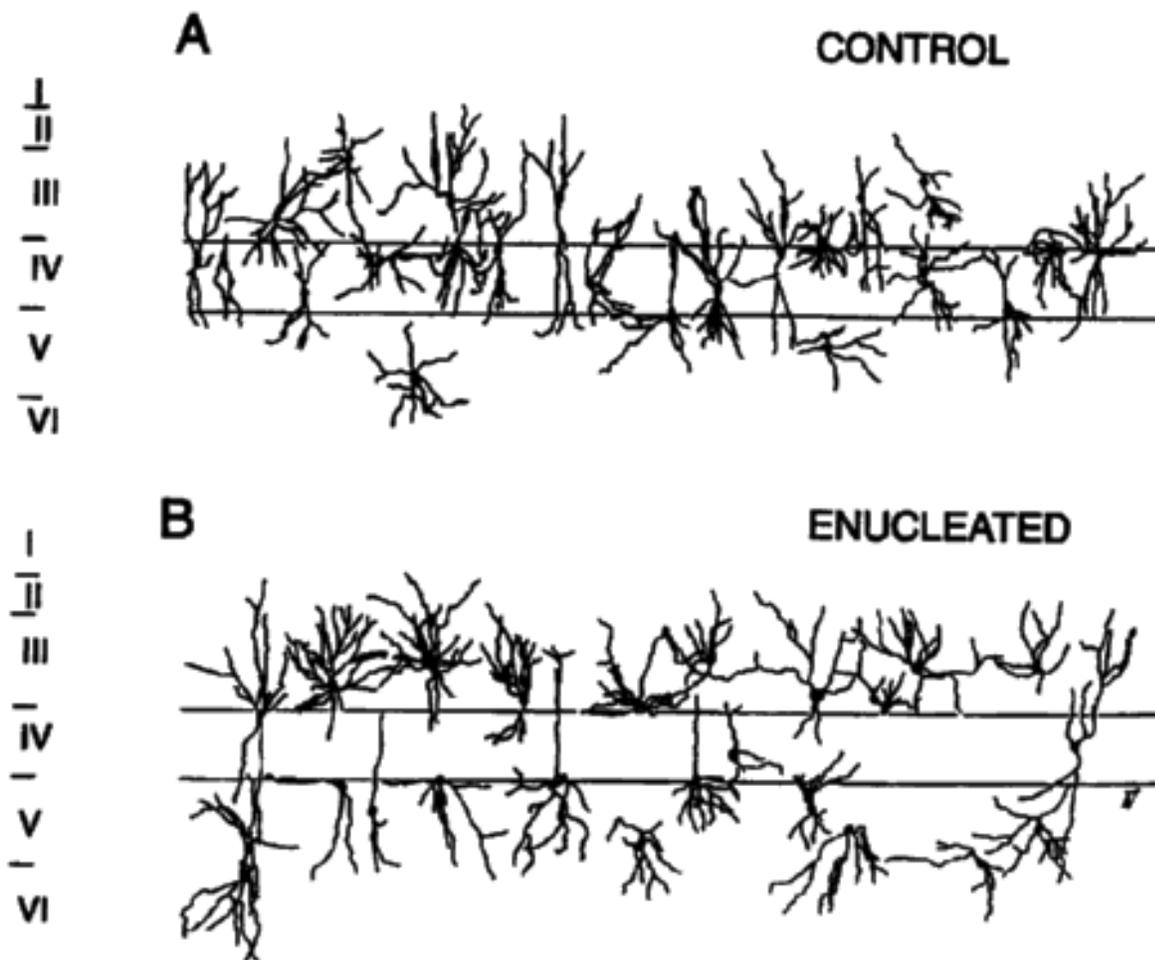


Figure 8: Dendritic organization in visual cortex of normal mice (A) and enucleated mice (B). The degenerative afferent termination is evident in B, where layer IV is sparsely covered with dendrites, whereas adjacent layers are more heavily covered, suggesting that these dendrites have reorganized according to remaining patterns of afferents (from Valverde 1968).

While studies such as Valverde's illustrate the dependence of dendritic form on afferent pathways, the study of Mooney et al. (1992) illustrates the striking malleability of developing dendrites. Mooney et al. (1992) examined the effects of neonatal enucleation on the dendritic morphology of superior collicular (SC) neurons. Like 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. But when they examined these cells' physiological response properties they found that a majority of them were no longer visually responsive, as in the normal case, but now had somatosensory response properties.

This leads to an important result: in the enucleated animal, the dendrites redirected their growth to find active afferents; where these were of a different modality, the cells changed their response properties to reflect this residual source. So, these response properties corresponding to the cell's function are not predetermined, but depend on interacting with the information modality latent in the pattern of incoming activity.

2.3.4 Directed dendritic development and patterns of activity

The finding that dendrites actively seek out incoming activity and shape their responses to mirror that activity calls for a closer look. A particularly attractive model system is the development of ocular dominance columns (introduced in section 2.2.1). As alternating columns represent activity from different eyes, with their border representing changes in correlated activity, what happens to developing dendrites at the borders of these columns is particularly revealing. Would the dendrites of cells near a border ignore it, or would their organization respect borders? A striking finding of Katz et al. (1989) was that the dendrites of these layer IV cells did indeed respect that border and preferred to keep within its column.

What is the signal that regulates this development? 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 received. What would the role of such a developmental signal be? The most direct role would be in the development of the response properties of the cell. Cells of layer 4c are almost exclusively monocular, that is, they respond to stimulation from only one eye. So, 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. In addition, 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 themes of dendritic development in the somatosensory cortex have also been observed (Greenough & Chang, 1988).

The dependence of dendritic form on patterned activity is supported by analogous experiments in the vertebrate optic tectum (Katz and Constantine-Paton 1988). Although the 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, which are not present in the normal frog. A striking result of the Katz and Constantine-Paton (1988) study was that tectal cell dendrites not normally receiving input from more than one eye nonetheless become 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.[\[4\]](#)

An interpretation of these results is that dendritic segments function as detectors of correlated activity and grow preferentially in such regions. Support comes from Tieman and Hirsch's (1982) finding that exposure to lines of only one orientation during development has specific effects on dendritic development. The dendritic field orientations of cells from cats raised with exposure to lines of a single orientation were significantly elongated in conformity with 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 environmental structure. A similar result has been obtained for layer IV stellate cells by Coleman et al. (1981), who suggest (p.19): "[I]f 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."

Recently, Kossel et al. (1995) used many of the experimental manipulations that led to activity-dependent

rules for axonal growth to examine dendritic growth. They used double labeling techniques to visualize OD columns and dendritic morphology simultaneously under conditions of monocular deprivation and divergent squint (strabismus). Strabismus results in a decrease in between-eye correlations and should therefore 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 the case of monocular deprivation, however, dendrites in the non deprived column were only weakly influenced by borders, reflecting the decrease in uncorrelated activity across that border.

Kossel et al. (1995) also found that cells in the deprived column extended their dendrites into the nondeprived activity. This agrees with other evidence we have reviewed that dendrites are not merely passive structures but actively seek out regions of correlated activity. As Kossel et al. (1995) conclude, their results suggest that dendrites develop according to the same sorts of rules that have been suggested for axonal arbors and that both structures develop according to patterns of correlated input activity.

2.3.5 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 computational models of development (reviewed in Goodhill, 1992). The discovery of the NMDA type of glutamate receptor made an important contribution to this research because the kinetic properties of the NMDA-receptor fit with the cooperative model of synaptic plasticity in both developmental (reviewed in Shatz 1990; Cline 1991) and mature forms of plasticity (Kelso et al. 1986).

Neurobiologists refer to the cooperative model of synaptic plasticity as "Hebbian learning," after Donald Hebb's (1949) proposal for a neurally plausible associative learning rule. In development, however, Hebbian learning is generally given a selectionist interpretation as a rule governing the stabilization of existing synapses. Hebb, though, made his original proposal in the context of neural development and the activity-dependent construction of new synapses in collections of neurons he called "cell assemblies." Hebb even discusses Kapper's neurobiotaxis theory, an early, extreme constructivism, and defends a limited version of it. Ironically, Hebb was reluctant to embrace a stronger version of constructivism because of Sperry's (1943) influential work. Sperry's elegant work on the regeneration of the retinotectal pathway led to his "Chemoaffinity hypothesis," that neurons bear unique molecular addresses making their connections precise, a hypothesis that would dominate neurobiological thinking for three decades.

There is suggestive evidence that neural constructivism is the most appropriate one for the NMDA-receptor's properties and that the Hebbian model should include directed growth. For example, from their experimental observations, Katz and Constantine-Paton et al. (1988) suggest that such a broader action of the NMDA-receptor's associative principles 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.

Cell culture studies further support the role of 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 whereas NMDA blockade leads to a significant decrease in these measures. Similar results have been reported in a variety of other systems (e.g., Pearce et al. 1987; Balazs et al. 1989; Bulloch & Hauser 1990).[\[5\]](#)

Recently, Williams et al. (1995) have shown that local stimulation along developing neuronal processes results in branching. These new branches are stabilized if the appropriate targets or signals are present. This branching is highly regulated and is calcium-dependent, as are the mechanisms involved in Hebbian learning. This again suggests that dendritic structure is added to those areas of activity to support more input from sources localized to that region.

3 Directed dendritic development and representational change

Now that directed dendritic growth appears to be an important component of brain development, we consider how it might underlie the development of the brain's representational properties. This is the third step in the methodology we outlined in section 1. Our aim is to first extract some general features of directed dendritic growth that conform to representation construction. Then, in section 4, we will suggest that this is a form of learning, "constructive learning," that makes the developing cortex a more powerful learner than usually supposed.

What sort of representations does the brain use? One of the most important principles of cortical representation involves "geometric principles of information processing design" (Mitchison & Durbin 1986; Mead 1989; Durbin & Mitchison 1990; reviewed in Churchland & Sejnowski 1992). Regarding this principle, Mead (1989, p.277) states:

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 physical structure of a neural area corresponds to a representational space. In this representational space, nearby things are more closely related semantically than things that are far apart. This map property is extremely powerful as a representational system. When brain areas can communicate, increasingly rich representations can be successively built up. Each area is a layer in an increasingly abstract feature space. Just as information in a map is held by such spatial properties as physical distance, the physical structure of cortex encodes information. With geometric principles of information processing the

information is held in the three-dimensional pattern of neural connectivity. As constructive factors play a central role in building this physical structure, they also shape the representational properties of cortex. Building neural circuits with directed growth thereby builds the brain's representational properties.

These spatial properties of representation are largely lost in the traditional connectionist network because of the way the connectionist neuron integrates information, typically summing its input and sending 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. Local dendritic segments might be the brain's basic computational units (see also Koch et al. 1982, 1983; Shepherd & Brayton 1987; Mel 1992a; 1992b; 1994; Jaslove 1992; Segev & Davis 1995). Dendrites are not simple signal integrators with passive conductance properties, as in classical cable models (Rall 1964). Imaging studies have found that some dendritic systems (e.g., CA1 hippocampal neurons) have a heterogeneous distribution of voltage-gated Ca^{2+} channels, suggesting nonlinear membrane properties (Jones et al. 1989; Regehr et al. 1989). Intradendritic recordings in these cells likewise reveal strong nonlinearities in their electrical properties (Wong et al. 1979; Bernardo et al. 1982;). In some instances, these properties make a dendritic segment act more like an axon, sending an amplified signal to the cell body (Stuart & Sakmann 1994).

Nonlinear properties give a cell more computational power than traditionally thought (Feldman & Ballard 1982; Durbin & Rumelhart 1989; Mel & Koch 1990; Koch & Poggio 1992). A cell with active dendritic segments can perform the nonlinear discrimination that requires a hidden-layer network of connectionist neurons. The spatial properties of a cell may also determine many of its functional properties. To connecting this back with our earlier discussion of geometric principles of information processing, when such a cell is embedded in a representational space, its spatial structure takes on additional significance. A phenomenon referred to as the "clustering" of related inputs onto dendritic segments results in a pattern of termination mirroring the informational structure of input: electrotonically close synapses encode common features of the input space and effectively fire the cell (Mel 1992a; 1992b; 1994). The presence of cluster-encoded features significantly alters both the representational properties of cortex and its computational power.

3.1 Developmental mechanisms underlying clustering

For clustering to serve an interesting information processing role, there must be an appropriate developmental scheme underlying its formation. For this, temporal patterns of activity must somehow be 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 subservise the formation of these clusters, with a cell able to tune itself to its input space (Mel 1992a; 1992b). Many of these schemes are biologically implausible, however, because of what is known as "the problem of sampling."

The sampling problem is the needle in a haystack problem: clusters depend on forming contacts from axons carrying similar information onto a single dendritic segment. Rearranging contacts involves the problem of finding the right dendritic segment. The sampling problem has been considered in a more general context by Montague et al. (1991) and Gally et al. (1990). In view of the developing nervous system's sparse connectivity, Gally et al. suggested that a spatially diffusible substance was acting (see Figure 9). Not confined to the anatomically defined synapse, a spatial signal is free to diffuse into a local volume, thereby potentially affecting all cells synapsing in that volume, whether or not a given cell shares a synaptic contact with it. In particular, Gally et al. proposed that nitric oxide, a membrane permeable gas, has a number of characteristics that make it a leading candidate for such a role. Subsequent research has confirmed that nitric oxide plays a key role in synaptic plasticity (Bohme et al. 1991; Haley et al. 1992; O'Dell et al. 1991; Schuman & Madison 1991) and transmission (O'Dell et al. 1991; Manzoni et al. 1992; Montague et al. 1994).

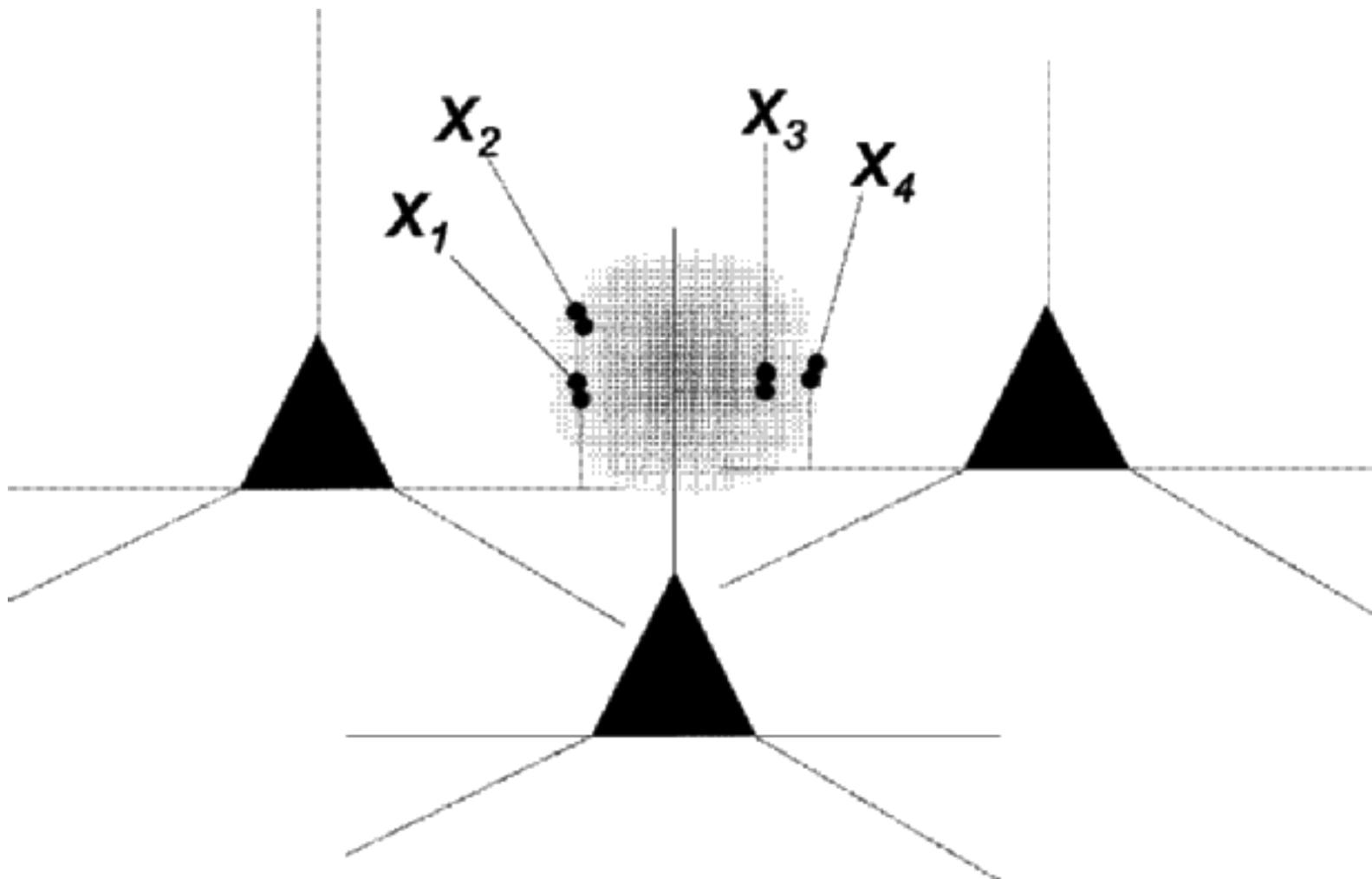


Figure 9: A diffusible substance allows synapses in a local volume of tissue to communicate whether or not they share a connection. Using such a signal, it is possible for synapse X1 and X4 to modify their weights according to an associative learning rule. (From Montague & Sejnowski 1994). One such rule is: $\Delta w_{ij} = \eta \cdot x_i \cdot y_j$, where Δw_{ij} is the change in the 'weight' or synaptic efficacy of a connection, η is a constant controlling rate of change of synaptic efficacy, x_i is a measure of presynaptic activity and y_j is a threshold that determines whether a terminal is active at time t . y_j 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, at time t located at position r .

Such a spatial 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 referred to as volume learning (reviewed in Montague & Sejnowski, 1994). This sort of learning rule takes associations "off the synapse" and into a local volume of neural tissue, thereby allowing the volume to hold associations. [6] This sort of mechanism could also play a central role in providing the robust sampling mechanisms that clustering requires. Instead of having to sample identical postsynaptic structures, 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 where two cells make no direct contact with each other.

In large-scale computer simulations in collaboration with P.R. Montague we are exploring how this scheme may be readily modified to include activity-dependent branching. The probability of branching/retraction at a terminal segment can be made proportional to the weight of nearby synapses over time. Making the probability of branching depend on synaptic weight automatically transfers the associative conditions necessary for weight changes to those 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 defining higher-order features of the input space. This, then, corresponds to locally regulated growth, allowing differential sampling as a function of the correlational structure of input patterns to form spatial clusters. In addition, since the production of the diffusible substance is postsynaptic, the postsynaptic structures play an important role in determining the properties of this feature space. Other mechanisms, such as the distribution of membrane channels and localized inhibitory input, will also participate in defining these clusters. We suggest that the establishment of spatial domains as regions of higher-order features will be central to the information-processing properties of neuronal populations.

3.2 Hierarchical Representation Construction

A particularly striking feature of the human brain is the protracted development of some of its parts. It is sometimes suggested that the severe constraints placed by bipedalism on the dimensions of the birth canal forced the human brain to be particularly immature at birth. There is nothing unique about the human brain's degree of immaturity at birth, however. As Gibson (1990) points out, at birth the human brain 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, dendrites of human primary visual cortex cells are between 30 and 50% of total adult lengths at birth. For these layer V pyramidal cells, adult lengths are reached by around four months compared to 18 months for layer III cells.

Much of non visual cortical development, in contrast, displays an extensive and more protracted development. Cells in frontal cortex are far 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---they grow to over thirty times their dendritic length at birth. Scheibel (1993) likewise reports a

long period of dendritic development in Broca's area in which mature forms emerge only after 6-8 years. Why, then, is human non-visual cortical development so slow to develop and so extensive?

Our view is that the human brain's development is a prolonged period in which environmental structure shapes the brain activity that in turn builds the circuits underlying thought. In place of pre-wired modules, patterned activity builds up increasingly complex circuits, with areas staging their development. Cortical areas further away from the sensory periphery wait in anticipation of increasingly complex patterns of activity resulting from development in lower areas. As this development proceeds, areas of the brain become increasingly specialized for particular functions, reflecting a cascade of environmental shaping. Some brain circuits close to the sensory periphery, such as in our early visual system, are in place by six months of age; but those in language areas, further away from the sensory periphery, do not begin to complete their development until the eighth year of life.

3.3 What is the role of regressive events in development?

The evidence we have examined demonstrates that the popular view of development as largely a regressive event must be reconsidered. We suggest that regressive events are simply the consequence of reduced neural specificity, as indicated by the counterevidence to Sperry's chemoaffinity hypothesis. Any theory, whether selectionist or constructivist, that rejects a strong view of neural specificity will thus need to posit regressive events. If cells do not bear nearly unique molecular addresses, then stochastic sampling mechanisms must be posited. These will by their very nature introduce some structure into a system that will later be eliminated. Neural constructivism allows these sampling mechanisms to be directed, but they are still stochastic. Structural elimination, or error-correction, are likewise required, but this does not mean that error-correcting processes are the only developmental mechanisms, or that developmental selection occurs only among intrinsically generated structures. [7] Rather, selection is only one kind of process in a *dynamic interaction* between environmentally derived activity and the neural growth mechanisms that activity regulates. This changes the emphasis from synaptic elimination to synaptic turnover. 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; Dailey & Smith 1996) reveal that constructive and selective events co-occur as required for directed sampling mechanisms.

Where does this leave the selectionism? We see no way for its strong interpretation to include mechanisms for directed growth without trivializing its driving analogy from population biology. Development that is directed is not selectionist---if environmental structure builds neural circuits, instead of simply selecting among variation created by intrinsic mechanisms, then the strict selectionist position is untenable.

4 A learning-theoretic approach to development

The neurobiological evidence we have examined suggests that the rigid distinction between learning and maturation can no longer be maintained. Instead, learning guides brain development in very specific ways. This question brings us to the fourth step of the method we outlined in section 1, to examine neural constructivism's learning properties. Does the interaction between learning and structural growth give a developing system any special learning properties? We now turn to our answer: this interaction gives a

developing system unique learning properties that undermine central assumptions about skill acquisition in cognitive science.

4.1 Development and learnability

The best known characterization of a developing system's learning properties comes from language acquisition---what syntactic properties 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 have provided mainly negative results: the environment is impoverished and could provide only limited information, so few syntactic properties could be learned. In the end, theorists concluded that the child must bring most of its syntactic knowledge, in the form of a *universal grammar*, to the problem in advance. Learning could only serve to set a few free parameters in that universal grammar so as to particularize it to the language confronting the child.

Although it is also based on empirical studies of linguistic input (e.g., Brown 1973), the perception that this striking view of syntax acquisition is based primarily on rigorous results in formal learning theory makes it especially compelling. Indeed, above all it is this formal feature that has prompted its generalization from syntax to the view of the entire mind as a collection of innately specified, specialized modules (e.g., Fodor 1983; Barkow et al. 1992; Gazzaniga 1992; Hirschfeld & Gelman 1994). Although Piaget's legacy remains undeniable in developmental psychology (e.g., Fischer 1980; Bates and MacWhinney 1987; Karmiloff-Smith 1992) it is probably no overstatement to suggest that much of cognitive science is still dominated by Chomsky's nativist view of the mind.

According to formal learning theory, development is a learning problem and so is constrained by the learning-theoretic pressures confronting any learner (Gold 1967; Pinker 1979; Wexler & Culicover 1980; Osherson et al. 1986). This assumption allows for a very general characterization of the learner. The classic formulation 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 could learn when presented with example sentences of some language. Gold supposed that the learner's task was to conjecture a hypothesis regarding the grammar that might generate 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 good question to ask is, where does Gold's learner get the grammars that it conjectures? Gold's learner adopts a maximally general strategy and first simply enumerates every grammar belonging to some class of grammars. Starting with the first grammar, the learner then rejects each one in turn if it is inconsistent with what it has seen so far and tries out the next grammar in the enumeration.

Such a learner will eventually find the right grammar if it has some finite position in the enumeration. The formal definition of a language from mathematical logic lends itself to formulating the languages that can be learned in this scenario. Primitive recursive languages emerge from a ranking of grammars known as *The Chomsky hierarchy* as the most powerful that can be learned by Gold's learner. They are the most powerful decidable language, which means that the right grammar will indeed have a finite place in the enumeration.

Some immediate troubles arise from Gold's model. As Pinker (1979) notes, this learner may have to test on

the order of 10¹⁰⁰ possible grammars even in an extremely simplified case---a computation that could never actually be performed. The learner is so slow because of the general strategy it adopts. Although this guarantees convergence, learning becomes in general impossible because of the vast search it requires. These prohibitive results may seem to indicate that language learning is impossible, but their consequences are ambiguous because of some major limitations. Even ignoring their dubious assumptions regarding the psychology of learning, there are two internal limitations: their concern only for convergence in the limit and their requirement that the learner precisely identify the target concept (no mistakes allowed).

In 1984, Les Valiant introduced a probabilistic model of learning that remedied these two limitations; his accordingly becoming the standard model of inductive inference in the field (see Dietterich 1990 and Natarajan, 1991 in the case of machine learning). Rather than disallowing any mistakes, Valiant's learner could make a hypothesis that was only a good approximation with high probability. This framework was dubbed the "probably approximately correct" (PAC) model of learning. It 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 in a hypothesis space (e.g., searching through the grammars) to the target concept. To make this a feasible search, the space must be restricted by building in an *inductive bias*, roughly the system's background knowledge. One of the Valiant model's key virtues was that it quantified the relation between inductive bias and learning performance from within a complexity-based account (e.g., Haussler 1989). Results with Valiant's model thus showed how difficult some problem was to learn with various inductive biases or background knowledge.

The Valiant model thus demonstrated what could not be fully characterized in the earlier limit-based formal learning theory: 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; they instead center around how to find appropriate representations to underlie efficient learning (reviewed in Geman et al. 1992). This problem precedes the treatment of learning as statistical inference, as a learner's choice of representation class (background knowledge) largely determines the success of learning as statistical inference.

What makes learning so difficult? In statistical studies, estimation error is decomposed into two components: bias and variance. Very roughly, bias is a measure of how close the learner's best concept in its representation space approximates the target function (the thing to be learned). Variance refers to the actual distance between what the learner has learned so far and the target function. To make this a bit more concrete, a small neural network will be highly biased in that the class of functions allowed by weight adjustments is 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, hence flexible in terms of what it can represent (by

decreasing bias) variance's contribution to error typically increases. That is, the network has many more possible states, and so is likely to be far away from the function of interest. This means that very large training sets will be required to learn because many examples will be required to rule out all the possible functions.

As Geman et al. (1992) state it, this results 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 too high a demand on training time and resources. This is essentially the same impasse that leads to nativism. Learning is too hard without severely restricting what can be learned. Indeed, from an entirely different perspective, Geman et al. (1992) likewise suggest that deliberately introduced bias (the nativist route) may be the best way out of this dilemma.

What makes these results interesting for the present discussion is that this basic problem of representation choice is a developmental one for natural systems. This, then, implies that the fundamental problem facing natural cognitive systems is a developmental one. How have natural systems chosen a developmental strategy to get out of this impasse?

Once we are talking about natural systems, it is worthwhile to raise a neurobiological constraint. So far, this discussion has proceeded as though the only significant factors were learning-theoretic pressures, but it is particularly important to consider whether the views coming out of learning theory are consistent with neurobiological constraints on development. For natural systems, the constraint that a learning theory should be consistent with underlying neural mechanisms has been severely underestimated. Indeed, in our opinion this biological constraint has equal footing with the learning-theoretic one and both must be viewed as complementary constraints that developmental systems must satisfy.

As suggested by Quartz and Sejnowski (1994), the view that strong, domain-specific knowledge is built into cortical structures runs into severe difficulties from developmental neurobiological evidence. Although we will not review that material in detail here, recent experiments on heterotopic transplants (Stanfield & O'Leary 1985; Schlaggar & O'Leary 1991; reviewed in O'Leary et al. 1992), cross modal rewiring (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, 1993) all establish that the regional characteristics of mature cortex depend fundamentally on interaction with afferent input. While the cortex is not a *tabula rasa*, as there may be a common basic circuitry and repetitive arrays of cell types (see O'Leary et al. 1992), it is largely equipotential at early stages (in agreement with studies on cortical plasticity and early brain damage (e.g., Neville 1991).

Consistent with 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. In addition, 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. A system in which the cortex is "enslaved by the periphery" has a number of clear advantages in terms of responding flexibly to varying environmental pressures, plasticity, and changing body size (see Walsh & Cepko 1992; 1993). In the previous section, we tried to suggest how this interaction between developing cortex and environmentally derived activity builds up the neural circuits underlying cognition.

Adding the neurobiological constraint to the learning-theoretic one results in yet another impasse. 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. Yet, from the perspective of biological constraints it appeared that cortical structures do not build in this knowledge, but rather allow both pre- and post-natal activity to determine features of the cortex. In the following section, we suggest that the neural constructivism offers a powerful means of escaping this impasse.

4.2 Constructive learning

The learning paradigm that is used in both Gold's and typically in Valiant's model is known as *selective induction*, with learning amounting to the search through some representation or hypothesis space. Most studies of learning assume that the learning mechanism's resources are stationary, or time-invariant, with learning a process of selective induction. 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). For this reason, the hypothesis space must be an exponentially small subset of possible concepts (see Blumer et al. 1988). This restriction in a learner's expressive power has far-reaching consequences, particularly in terms of limiting its flexibility.

To see these consequences, the first question to ask is, what does failure signify on such an account? Since the hypothesis space must be a very restricted subset of possible concepts, failure to learn may simply indicate that the learner chose the wrong hypothesis space; this may say nothing about the learnability of some class of concepts. 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." Is there a way for a learner to be more flexible, to avoid having to make such a restrictive initial choice of representations?

The constructivist learner shows its advantages here. It does not involve a search through an pre-defined hypothesis space, and so it is not one of selective induction (also known as model-based estimation, or parametric regression). Instead, the constructivist learner builds its hypothesis space as it learns. This has shifted the problem from one of parameter estimation to a nonparametric regime. We must accordingly ask, what is the effect of allowing a system to add structure---to build representations---as it learns?

Here again nonstationarity poses a challenge since we are asking about the effects of building representations according to the features of the learning problem. Neural network research has been particularly helpful in characterizing this sort of nonstationarity because the close relation between a network's architecture and its representational properties provides a constrained framework for looking at representational change during learning.

An increasingly sophisticated literature on the formal properties of neural networks has emerged. For example, a number of general results on the approximation properties of neural networks have been established (e.g., Cybenko 1989; Hornik et al. 1989; Girosi & Poggio 1990). From a nonparametric framework, White (1990) has demonstrated that a network that adds units at an appropriate rate relative to its experience is what statisticians call a consistent nonparametric estimator. This asymptotic property means that it can learn essentially any arbitrary mapping. The intuition behind this result, which will play a central role in characterizing constructive learning, follows a general nonparametric strategy: slowly increase representational capacity by reducing bias at a rate that also reduces variance. Since network bias depends on the number of units, as a network grows, its approximation capacities increase. The secret is regulating the rate of growth so that variance's contribution to error does not increase. Encouraging bounds on the rate of convergence have recently been obtained (Barron 1994).

White's demonstration of the power of neural networks depends on allowing the network to grow as it learns. In fact, many of the limitations encountered by neural networks are due to a fixed architecture. Judd (1988) demonstrated that learning the weights in a neural network is an NP-complete problem, and therefore computationally intractable, a result that extended to architectures of just three nodes (Blum & Rivest 1988). These results suggest that severe problems may be lurking behind the early success of network learning. As Blum and Rivest (1988) note, however, these results stem from the fixed architecture property of the networks under consideration. In contrast, the learning problem becomes polynomial (feasible) if the network is allowed to add hidden units. This suggests fundamentally different learning properties for networks that can add structure during learning. This has been confirmed by studies such as that of Redding et al. (1993), who presented a constructivist neural network algorithm that can learn very general problems in polynomial time by building its architecture to suit the demands of the specific problem.

Underlying this sort of result is Baum's (1988; 1989) demonstration 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 whatsoever. As Baum notes (1989, p.201), this makes the learner a sort of general or universal one. This is in contrast to systems which utilize incomplete representations, as in a fixed hypothesis space. Most negative learnability results, such as those for syntax, depend on a system using incomplete representations (see below). If a network is allowed to build its representations as it learns in response to the informational structure of its environment, the question of learnability shifts from the question of what is learnable from some particular representation class to the question of what is learnable from any representation class.

The general strategy of constructivist learning is this. Rather than start with a large network as a guess about the class of target concepts, avoid 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. Since the construction of the learner's hypothesis space is sensitive to the problem domain facing the learner, this is a way of tailoring making the hypothesis space to suit the demands of the problem at hand. This allows the particular structure of the 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 good representation beforehand 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. Research on constructive algorithms has become increasingly sophisticated, and the results with constructive learners are impressive (e.g., Fahlman & Lebiere 1990; Frean 1990; Hirose et al. 1991; Platt 1991; Azimi-Sadjadi et al. 1993; Wynne-Jones 1993; Kadiramanathan & Niranjana 1993; Shultz et al 1994; Shin & Ghosh 1995).

The research we have just examined indicates a fundamental distinction between the constructivist strategy and models of selective induction. For the latter to have any chance of learning, the network must build in much of the problem domain *a priori*. Besides the neurobiological implausibility of this strategy, there are more general reasons why using 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, but if the nature of the application is not properly predicted, the network will be a poor performer. Hence, tailor-making 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. Biological learners, however, face an additional problem: not only is the problem space not defined beforehand, it is changing---the environment is highly nonstationary. Systems that are highly specialized for the anticipation of a particular problem domain will 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. Although most natural systems are only confronted with ecological change, human cognition requires highly flexible and adaptive representations to accommodate both cultural and technological innovations. We doubt that the pace of this change can be met by a representational scheme requiring a major intrinsic specification.

4.3 Neural constructivism and language acquisition

Because constructive learning shifts the learning problem from what is learnable from a fixed hypothesis space to that of what is learnable from any representation class, it suggests that the class of learnable concepts might be larger than traditionally supposed. It is, therefore, worthwhile to ask whether constructive learning has any bearing on the results coming from Gold's work.

We can approach this question by first asking what the results from Gold's work really show. Do they demonstrate that syntax is unlearnable? The shift in the meaning of learnability we just mentioned suggests that the unlearnability of syntax has two possible senses. It may mean that syntax is not learnable from some fixed hypothesis space H . Two possible causes underlie this sort of unlearnability: either the target function (encoding syntax) is too large or H is too restricted (see Baum 1988; 1989; Valiant 1991). Most negative results are of the second sort. As we showed, the constructivist learner escapes these sorts of negative results by constructing more powerful representations than those contained by the fixed architecture. So, in this case a negative result just indicates that a poor hypothesis space was chosen---it is only a negative result for this specific hypothesis space and says nothing about the learnability of syntax itself.

Most cognitive scientists, however, do not view the unlearnability of syntax as this sort of result. Instead, they see it as a representation-independent result. This is a much stronger sort of result, claiming that syntax is unlearnable relative to any hypothesis space. In this case, there would be no reason to look for more effective representations or systems that can build representations as they learn because no representation at

all could possibly suffice. Is this justified? The answer is no---the only representation-independent results are for complicated cryptographic functions, such as those known as "polyrandom functions" (functions that cannot be distinguished from purely random ones in polynomial time; see Goldreich et al. 1984). This type of representation-independent result, however, is of little relevance to the learnability of syntax, or for the sorts of concepts natural systems must learn. Learning syntax is nothing like having to solve the general decryption problem. Hence, although the general perception is that Gold's work proved syntax to be representation-independent unlearnable, there is no justification for this strong claim.

The negative results regarding syntax are of the weaker sort: unlearnability relative to some fixed hypothesis space. It is also important to point out that there are some idiosyncratic features of Gold's learner that make learning appear to be hard: learning as selective induction, a stationary learner, extremely dubious assumptions regarding the psychology of learning, an impoverished account of linguistic input, a worst-case analysis, and extremely rigid performance conditions. Above all, because Gold's learner uses such a general strategy, simply enumerating an entire class of grammars, and then evaluating learning in the worst case, its results are limited to its own framework and have little applicability to the problem of learning in general. Indeed, to us the main lesson of learnability arguments in Gold's paradigm demonstrate is the insufficiency of its own model---the baby may have been thrown out with the mathematical bathwater.

Since syntax appears to belong to the class of concepts that are learnable by natural systems, as indicated by it not being a representation-independent unlearnable class, the appropriate response to results from Gold's framework is to reject this model of learning and begin to explore alternatives. In particular, nonstationary learners, long dismissed by Chomsky and others (e.g., Pinker, 1984), offer a more powerful response to the problem of learning. In particular, constructive learning is a maximally powerful approach, in that it forms complete representations, capable of learning any learnable concept.

The powerful learning properties of constructive learning are not its only advantages. We suggested that all candidate learners must satisfy both learning theoretic and neurobiological constraints. Constructive learning points to the dynamic interaction between a structured environment and the neural mechanisms that are responsive to that structure. As such, it minimizes the amount of built-in structure required, making it the only learner consistent with a largely equipotential cortex. Constructive learning is, therefore, the only learner consistent with both learning and neurobiological constraints.

The themes we have presented in this target article are very simple steps toward characterizing the complex interactions between developmental mechanisms and a structured environment. Already, however, we think they force extreme caution in formulating theories of acquisition in their absence. Although this interaction will be no doubt far richer than what we have captured, it raises some intriguing possibilities that have been discounted under the influence of nativist approaches, which we consider next.

No learner can be completely assumption free since pure *tabula rasa* learning is impossible---there must be some built-in assumptions. A future research direction will be to characterize the sorts of biases that are consistent with a largely equipotential cortex: those deriving from such features as generic initial cortical circuitry, conduction velocities, subcortical organization, learning rates, and hierarchical development. The way these constraints provide grounding for constructive learning to build the mature representational structures of cortex will likely be a very rich account, although the tools required to understand this complex

interplay are still rudimentary.

We also think it is important to turn attention back to examining environmental structure, as in earlier traditions of developmental psychology. Both nativism in psychology (e.g., Chomsky 1965; 1980) and selectionism in neurobiology (e.g., Edelman, 1987) have made much of poverty of the stimulus arguments. The upshot of these arguments has been a depreciation of the environmental structure's role in guiding acquisition. As neural network and neurological research are finding, however, there appears to be far more structure latent in the environment than the poverty of the stimulus arguments suppose. In addition, we think the details of human cortical development provides another clue to the richness of environmental structure. Because human cortical development is much more prolonged and extensive than what purely physical limits predict, we think this suggests that the human brain's evolution has maximized its capacity to interact and be shaped by environmental structure through progressively building the circuits underlying thinking, as we explore in more detail next.

4.4 Neural constructivism and evolution

Neural constructivism suggests an evolutionary view that contrasts with the view evolutionary psychologists propose. Evolutionary psychologists (e.g., Barkow et al. 1992; Pinker 1994) champion Chomsky's nativism. Viewing development as an intrinsic process, they sought a phylogenetic source for the innately specified skills comprising the human cognitive repertoire. Although Chomsky himself 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 coded responses to phylogenetic pressures acting over the course of evolutionary history.

Sometimes this view is inserted into a selectionist framework (e.g. Gazzaniga 1992). Selectionism, however, is strictly incompatible with the claim that evolutionary pressures have picked out specialized circuits. 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) is more consistent when he restates Roger Sperry's chemoaffinity hypothesis that address-encoding surface markers on individual cells underlie the connectivity of specialized circuits (see Figure 3). Unfortunately, neurobiologists abandoned this extreme view of neural specificity twenty-five years ago (see Easter et al. 1985 for a review). The recent comparative analysis of Finlay & Darlington (1995) lend further support to the view that the brain is not a hodgepodge of specialized circuits, each chosen by evolutionary pressures. A major challenge for evolutionary psychologists, then, is to show how their view satisfies neurobiological constraints.

According to neural constructivism, to see human cognitive evolution as the progressive increase in specialized structures is to misinterpret cortical evolution. The hallmark of cortical evolution is not the ever-increasing sophistication of dedicated or specialized cortical circuitry (Gazzaniga 1995) but an increasing representational flexibility that allows environmental factors to shape the human brain's structure and function. At the expense of increased vulnerability during a protracted developmental period, constructive learning allows alterations in the learner's environment either through natural or human innovation to be accommodated by flexible representations without such changes being mediated by intrinsic mechanisms of

specification. Human development accordingly consists of two processes, first a prolonged period of representation construction in which neural structures respond to the informational structure of the environment, and, second, rapid learning, made possible by the first.

5 Conclusions

Although psychologists and neurobiologists both study development, communication and collaboration between fields have been limited. Reasons for this vary. Until recently, there was a lack of pertinent neurobiological data. In addition, reductive works such as Lenneberg (1967) viewed advances in the biological basis of development as lessening the cognitive contribution. So, where connections were made, they reinforced the opposition of neural and cognitive descriptions of development, an opposition that was perhaps most strongly made in the functionalist contention that neural descriptions were irrelevant for cognitive explanations (the so-called arguments from "multiple instantiability").

The extent and duration of large-scale brain changes during development has also been underappreciated. Whereas many researchers believe that the major events in brain development end by two years of age, the evidence we have reviewed illustrates these continue well past the first decade of life. Rather than being strictly reductive, neural constructivism points to the interaction between cognitive and neural processes in development, suggesting that cognitive and neural levels of description will need to be integrated into a single explanatory framework to explain this prolonged growth. Neural constructivism thus provides a meeting ground for cognitive scientists and neuroscientists. Although we are only beginning to understand how the world and brain interact to build the mind, the story that is unfolding already makes nativist theories appear implausible. What lies ahead promises to be an exciting---and far richer---account in which the mind emerges from a prolonged interaction with a structured world.

Acknowledgments

We would like to thank Ed Callaway, Read Montague, Michael Stryker, and Stevan Harnad for their helpful comments .

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.
- Bates, E.A. & Elman, J.L.(1993) Connectionism and the study of change. In: *Brain development and cognition: A reader*, ed. M.H. Johnson. Blackwell.
- Bates, E.A. & MacWhinney, B. (1987). Competition, variation, and language learning. In: *Mechanisms of language aquisition*, ed. Brian MacWhinney. Lawrence Erlbaum Associates.
- 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.
- Bennett, M.R. & Pettigrew, A.G. (1974) The formation of synapses in striated muscle during development. *Journal of Physiology (London)* 241:515-545.
- 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. (1988) 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 University 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. (1990) Emergence and refinement of clustered horizontal connections in cat striate cortex. *Journal of Neuroscience* 10:1134-1153.
- 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*, eds. W.M. Cowan, Z.W. Hall, & E.R. Kandel. Annual Reviews.
- 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.
- Changeux, J.P. & Dehaene, S. (1989) Neuronal models of cognitive functions. *Cognition* 33:63-109.
- 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. (1988) Perspectives on Cognitive Neuroscience. *Science* 242:741-745.
- 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.
- Coggeshall, R.E., & Lekan, H.A. (1996) Methods for determining numbers of cells and synapses: A case for more uniform standards of reviews. *Journal of Comparative Neurology* 364:6-15.
- 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.
- Dailey, M.E., & Smith, S.J. (1996) The dynamics of dendritic structure in developing hippocampal slices. *Journal of Neuroscience* 16:2983-2994.
- 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.
- Easter, S.S., Jr., Purves, D., Rakic, P. & Spitzer, N.C. (1985) The changing view of neural specificity. *Science* 230:507-11.
- Edelman, G. (1987) *Neural Darwinism: The Theory of neuronal group selection*. Basic Books.
- Elman, J.L., Bates, E.A., Johnson, M.H., Karmiloff-Smith, A., Parisi, D. & Plunkett, K. (in press, 1996) *Rethinking innateness: A connectionist perspective on development*. MIT Press.
- 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.
- Finlay, B.L., & Darlington, R.B. (1995) Linked regularities in the development and evolution of mammalian brains. *Science* 268:1578-84.

- Fischer, K.W. (1980) A theory of cognitive development: The control and construction of hierarchies of skills. *Psychological Review* 87:477- 531.
- Flood, D.G. & Coleman, P.D. (1979) Demonstration of orientation columns with [14C] 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.C. (1989) Development of Y axon innervation of cortical area 18 in the cat. *Journal of Physiology (London)* 416:183-213.
- 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 (London)* 359:293-313.
- Friedlander, M.J., Martin, K.A.C. & Wassenhove-McCarthy, D. (1991) Effects of monocular visual deprivation on geniculocortical innervation 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*, eds. 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., 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.
- Globus, A., & Scheibel, A.B. (1967) The effect of visual deprivation on cortical neurons: a golgi study. *Experimental Neurology* 19:331-245.
- 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.
- Goodhill, G.J. (1992) Correlations, Competition and Optimality: Modelling the Development of Topography and Ocular Dominance. Cognitive Science Research Paper 226, University of Sussex.
- Gottlieb, G. (1976) Conceptions of prenatal development: behavioral embryology. *Psychological Review* 83: 215-34.
- Gottlieb, G. (1991) Experiential canalization of behavioral development: Theory. *Developmental Psychology* 27:4-13.
- Greenough, W.T., Black, J.E. & Wallace, C.S. (1987) Experience and brain development. *Child Development* 58:539-559.
- 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.
- Herrmann, K. & Shatz, C.J. (1995) Blockade of action potential activity alters initial arborization of thalamic axons within cortical layer 4. *Proceedings of the National Academy of Sciences* 92:11244-8.

- 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., Uhlich, 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.
- Johnson, M.H. (1990) Cortical maturation and the development of visual attention in early infancy. *Journal of Cognitive Neuroscience* 2:81-95.
- Jones, O.T., Kunze, D.L. & Angelides, K.J. (1989) Localization and mobility of omega-conotoxin-sensitive Ca²⁺ 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.
- Karmiloff-Smith, A. (1992) *Beyond modularity: A developmental perspective on cognitive science*. MIT Press.
- Kasper, E.M., Larkman, A.U., Lubke, J., & Blakemore, C. (1994) Pyramidal neurons in layer 5 of the rat visual cortex. II. Development of electrophysiological properties. *Journal of Comparative Neurology* 339:475-94.
- 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*, eds. 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.
- Lenneberg, E.H. (1967) *Biological Foundations of Language*. Wiley.
- LeVay, S. & Stryker, M.P. (1979). The development of ocular dominance columns in the cat. In: *Society for Neuroscience Symposium: Aspects of Developmental Neurobiology*, ed. J.A. Ferrendelli. Society for Neuroscience.
- 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.
- Lichtman, J.W. (1977) The reorganization of synaptic connexions in the rat submandibular ganglion during post-natal development. *Journal of Physiology (London)* 320:121-130.
- 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*. 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.
- Macnamara, J. (1982) *Names for things: a study of child language*. MIT Press.
- Maffei, L. & Galli-Resta, L. (1990) Correlation in the discharges of neighboring rat retinal ganglion cells during prenatal life. *Proceedings of the National Academy of Sciences* 87:2861-2864.
- Mainen, Z.F., Joerges, J., Huguenard, J.R. & Sejnowski, T.J. (1995) A model of spike initiation in neocortical pyramidal neurons. *Neuron* 15:1427-1439.
- 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. L. Erlbaum Associates.
- 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. (1996) The resource consumption principle-attention and memory in volumes of neural tissue. *Proceedings of the National Academy of Sciences* 93:3619-3623.
- 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, M.J. (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 order in synaptic learning mechanisms. *Learning and Memory* 1:1-33.
- Mooney, R.D., Nikoietseas, 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.
- Neville, H. (1991) Neurobiology of cognitive and language processing: effects of early experience. In: *Brain maturation and cognitive development*, eds. K.R. Gibson & A.C. Peterson. Aladine de Gruyter Press.
- Noback, C.R., & Purpura, D.P. (1961) Postnatal ontogenesis of neurons in cat neocortex. *Journal of Comparative Neurology* 171:291-308.
- 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 macaque 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 macaque 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. (1992) Development of connectional diversity and specificity in the mammalian brain by the pruning of collateral projections. *Current Opinion in Neurobiology* 2:70-7.
- O'Leary, D.D.M., Schlaggar, B.L. & Stanfield, B.B. (1992) The specification of sensory cortex: lessons from cortical transplantation. *Experimental Neurology* 115:121-126.
- 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 pathway of ferrets. . novel inputs to primary auditory 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.
- 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. (1979) Formal models of language learning. *Cognition* 1:217-283.
- Pinker, S. (1994) *The language instinct*. W. Morrow and Co.
- 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.
- Plunkett, K. & Sinha, C. (1992) Connectionism and developmental theory. *British Journal of Developmental Psychology* 10:209-254.
- 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 imaging 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. & Lichtman, J.W. (1985). *Principles of Neural Development*. Sinauer Associates.
- 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.
- Pylyshyn, Z. (1984) *Computation and cognition: Toward a foundation for cognitive science*. Bradford Books.
- 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., Hahn, 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.
- Rumelhart, D., McClelland, J. & the PDP research group. (1986). *Parallel distributed processing: Explorations in the microstructure of cognition*. Cambridge, MA: Bradford Books.
- 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 neurocognition: Speech and face processing in the first year of life*, ed. 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 unique to somatosensory cortex. *Science* 252:1556-1560
- Schuman, E.M. & Madison, D.V. (1991) A requirement for the intercellular messenger nitric oxide in long-term potentiation. *Science* 254:1503-6.

- Segev, I., Rinzel, J. & Shepherd, G.M. (1995). *The theoretical foundations of dendritic function: Selected papers by Wilfrid Rall with commentaries*. MIT Press.
- Shatz, C.J. (1990) Impulse activity and the patterning of connections during CNS development. *Neuron* 5:745-756.
- Shatz, C.J. (1992) How are specific connections formed between thalamus and cortex. *Current Opinion in Neurobiology* 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.
- Shultz, T.R., Mareschal, D. & Schmidt, W.C. (1994) Modeling cognitive development on balance scale phenomena. *Machine Learning* 16:57-86.
- 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.J. (1943) Effect of 180° rotation of the retinal fields on visuomotor coordination. *Journal of Experimental Zoology* 92:263-279.
- 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) Activity-dependent reorganization of afferents in the developing mammalian visual system. In: *Development of the visual system*, eds. 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) Experimentally 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.
- Swindale, N.V. (1980) A model for the formation of ocular dominance stripes. *Proceedings of the Royal Society of London, B* 208:243-264.
- 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*, eds. 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. & Culicover, 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.

Williams, C.V., Davenport, R.W., Dou, P., & Kater, S.B. (1995) Developmental regulation of plasticity along neurite shafts. *Journal of Neurobiology* 27:127-40.

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.

References

- 1.** Friedlander et al. (1991) also found a number of cellular differences between the two groups of arbors suggesting that the observed shift in autoradiographic studies might be exaggerated. Depending on the pattern of incorporation of a radiolabelled tracer, it is hence 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.** We should note that although we are emphasizing dendritic development, aspects of axonal development also satisfy these conditions. As it is from the interaction between dendrites and axons that the structure of the mature system emerges, this interaction must ultimately be characterized.
- 3.** Koester & O'Leary (1992) report a significant retraction of layer V apical dendrites, but Kasper et al. (1994) report that these apical dendrites continue to grow and that the apparent retraction is due to the expansion of cortex.
- 4.** Differences in the degree of rostral-caudal dendritic bias between normal and stripe-induced cells support 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."
- 5.** 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 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), suggests that the elaboration of axon terminals at this fine level may proceed in this way.
- 6.** Specificity is maintained by requiring that the presynaptic terminal should be coincidentally active (See Montague & Sejnowski (1994) for discussion).
- 7.** A widely used metaphor to describe this process is that of error-correction. It should not be assumed, however, that the exuberant connections are strictly in error, since they may serve a useful purpose in instances in which a changes in connectivity is required (as in the case of blindness).