

# The neural basis of cognitive development: A constructivist manifesto

**Steven R. Quartz**

*Computational Neurobiology Laboratory and The Sloan Center for Theoretical Neurobiology, The Salk Institute for Biological Studies, La Jolla, CA 92037*

**Electronic mail:** [steve@salk.edu](mailto:steve@salk.edu) [www.cnl.salk.edu/cnl/](http://www.cnl.salk.edu/cnl/)

**Terrence J. Sejnowski**

*Howard Hughes Medical Institute, The Salk Institute for Biological Studies, and Department of Biology, University of California, San Diego, La Jolla, CA 92037*

**Electronic mail:** [terry@salk.edu](mailto:terry@salk.edu) [www.cnl.salk.edu/CNL/](http://www.cnl.salk.edu/CNL/)

**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.

**Keywords:** cognitive development; constructivism; evolution; learnability; mathematical learning theory; neural development; selectionism

## 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.

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 (Fig. 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 remov-

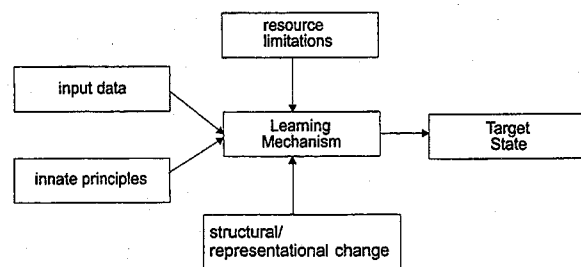


Figure 1. The logic of the problem of development. At its most abstract, the problem is defined as one of characterizing the mapping from some set of input data into the target state (the adult competence; see Wexler & Culicover, 1980, for such a treatment). This problem is complicated by two elements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

ing 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., Gottlieb 1976; Greenough et al. 1987; Hebb 1949; Purves 1988; 1994). 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., Bates & Elman 1993; Elman et al. 1996; Johnson 1990; Karmiloff-Smith 1992; Neville 1991; Plunkett & Sinha 1992; [see also multiple book review of Karmiloff-Smith's *Beyond Modularity*. *BBS* (17) 1994]). 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 (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.

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 (Fig. 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

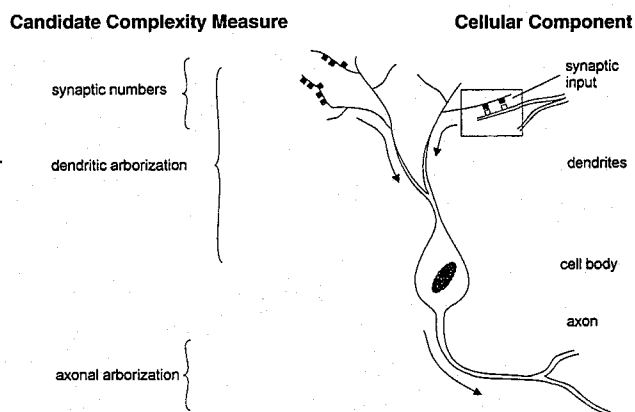


Figure 2. Candidate measures of structural/representational complexity. There are three possible (nonexclusive) measures: synaptic numbers, dendritic arborization, and axonal arborization. The figure also summarizes the basics of neural information processing elements: synaptic input travels along the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

view that there is an active interaction between the developing system and the environment in which it is embedded. Like Piaget's theory, ours also 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 preexisting, 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 grip 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 Churchland & Sejnowski 1992; Pylyshyn 1984). 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 (Bates & Elman 1993; Karmiloff-Smith 1992). This latter concern brings us back to nonstationarity. Although 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 (1883) 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., Changeux & Danchin 1976; Edelman 1987; Jerne 1967; Rakic et al. 1986).

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 (Fig. 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 (Hubel & Wiesel 1962; 1963; 1965; LeVay et al. 1980; Shtatz & Stryker 1978). To avoid a proliferation of terminology, we will refer to any view positing an initial exuberant growth followed by elim-

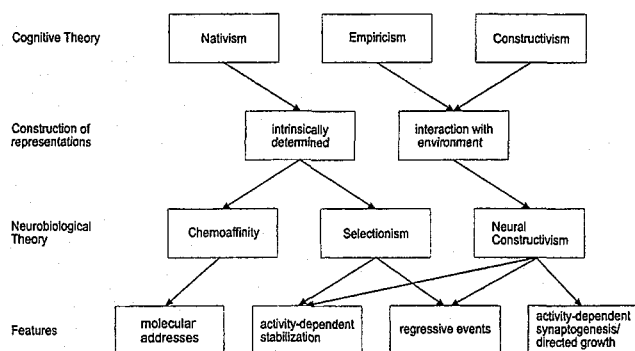


Figure 3. Summary of the relations among major cognitive and neural theories. Instructivism is often identified with Hebb's work, since in his view the environment guides or instructs brain growth. We refer to such theories as constructivist since the latter is a more widely recognized class of theories.

ination 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 (Fig. 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 (Bourgeois et al. 1989; Huttenlocher & de Courten 1987; Huttenlocher et al. 1982; O'Kusky & Collonnier 1982a; 1982b; Rakic et al. 1986; Zecevic et al. 1989).

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 2 and 4 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 and Holbach (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 parvocellular, 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? Reinterpretation has been forced by a recent study by Rakic and colleagues (Bourgeois et al. 1994) on synaptic development in the prefrontal cortex of macaque monkeys. With a larger sample, they found that synaptic density reached a peak around 2 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.

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

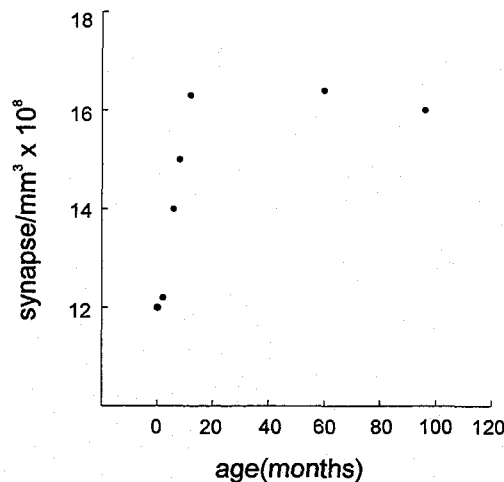


Figure 4. Synaptic density in layer III, human middle frontal gyrus from birth to 8 years. Data from Huttenlocher (1979). Note the absence of data points between 1 and 5 years of age.

development, however, are scarce. In fact, the only published quantitative study of synaptic measures in developing human nonvisual cortex is Huttenlocher's (1979) widely cited one (Fig. 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 to 5 and 8 to 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 1 year and 8 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 by Lund et al. (1990) 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).

Early work examined the effects of differential rearing conditions on synaptic numbers and density (reviewed in Boothe et al. 1979). 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 (Mainen et al. 1995; Stuart & Sak-

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

Study	System	Remarks
Valverde 1967; 1968 Globus & Scheibel 1967	mouse visual cortex, stellate cells rabbit visual cortex	decrease in spine density, numbers in dark rearing 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 & Greenough 1984	CA1, hippocampus	increased synapse numbers with long-term potentiation, but not non-LTP inducing stimulation
Turner & Greenough 1985	rat visual cortex	used electron microscopy to quantify synapse/neuron; highest numbers in complex environments, lowest in isolated environment
Bailey & Chen 1988a; 1988b	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

mann 1994), 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.** 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 (Hubel & Wiesel 1962; 1963; 1965; LeVay et al. 1980; Shatz & Stryker 1978). Ocular dominance columns are alternating, anatomically defined regions of input from primarily one eye (OD development is summarized in Fig. 5). As noted by Antonini and Stryker (1993, p. 3549), 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.

Although 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 fill intracellularly. 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., Linsker 1986; Miller et al. 1989;



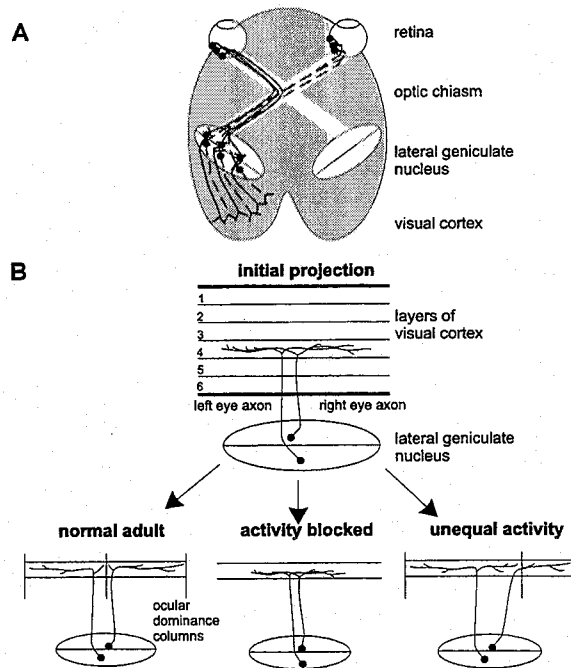


Figure 5. The generation of specificity according to the standard model. A, schematic representation of the connections between retinal ganglion cells and the lateral geniculate nucleus of the thalamus and connections to primary visual cortex. B, the two-phased process of development and perturbations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

Swindale 1980). 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 postsynaptic activity results in a strengthening of that synapse. Further conditions, supposing that anticorrelations 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 (reviewed in Stryker 1991), 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 retinogenicu-

late 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 7,538 to 12,848 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."

Similarly, in describing the development of intrinsic horizontal connections in area 17 (see Fig. 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.

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:

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

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 & Martin 1989	cat Y-pathway, area 18	progressive expansion of terminal arbors
Callaway & 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 & 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 & Stryker 1993	cat, X and Y-pathways, area 17	expansion of arbor area and branch numbers

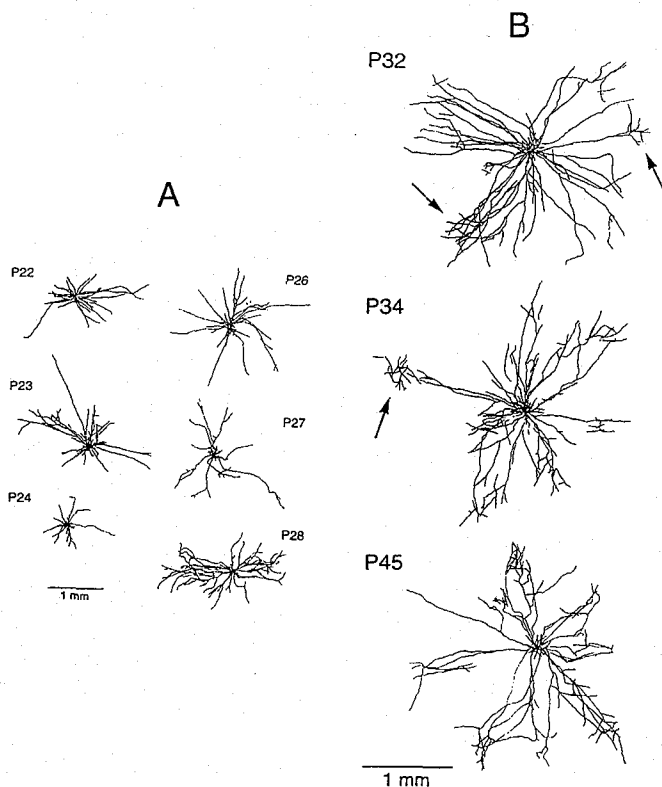


Figure 6. The development of horizontal connections in the ferret. A, at postnatal day 22 (P22) little clustering can be seen. As development proceeds, however, axon cluster formation becomes evident, particularly by P28. B, subsequent development, seen here under higher magnification, reveals the refinement 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 with permission from Durack & Katz (1996).

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 nondeprived eye. This expansion is caused by relatively more correlated activity from the nondeprived eye, and so its arbors compete favorably for territory that is initially shared by overlapping arbors. The deprived eye columns, in contrast, shrink (see Fig. 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.<sup>1</sup>

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, although 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:

(a) *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;

(b) *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 M. J. Katz et al. 1984; Uylings et al. 1990). 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 non-linear 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.<sup>2</sup>

**2.3.2. Dendritic arborization.** 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 nineteenth 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

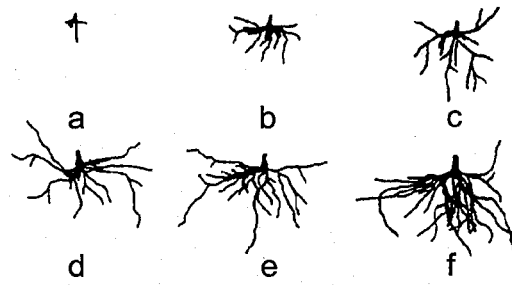


Figure 7. Camera lucida drawings of basal dendrites of layer V human pyramidal cells; a, newborn; b, 3 months; c, 6 months; d, 15 months; e, 24 months; f, adult (from Schade & van Groenigen 1961).

synapses (Segev et al., 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 6 months of age, this cell's dendrites expand over 10 times to a total length of over 2000 microns, or 2 mm. By 2 years of age, this cell's dendrites add another 1000 microns. Perhaps most startling, even by 2 years of age these dendrites have not yet reached their halfway point, but will more than double between 2 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 30 times their length at birth and undergo the majority of this growth after 2 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.

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.<sup>3</sup> 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.

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 (Caviness & Rakic 1978; Mariani et al. 1977; Pinto-Lord & Caviness 1979; Rakic & Sidman 1973).

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

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

Study	System	Remarks
Mathers 1979	rabbit visual cortex, layer V pyramidal cells	postnatal expansion of dendritic arbor and length
Juraska & Fífkova 1979	rat visual cortex	progressive dendritic expansion of pyramidal cells, layers II-III, V
Simonds & Scheibel 1989 Parnavelas & Uylings 1980	Broca's area rat visual cortex, layer IV stellate cells	progressive dendritic expansion into the sixth year progressive dendritic development lasting to postnatal day 20, followed by distal expansion to postnatal day 90
Buell & 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, layers III and V	layer III did not reach mature values until 2 years of age, followed by a nonsignificant 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 $\mu\text{m}$ total length to 1600
Lubke & 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 $\mu\text{m}$ total dendritic length

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, extended 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."

Whereas studies such as Valverde's illustrate the dependence of dendritic form on afferent pathways, the study by Mooney et al. (1992) illustrates the striking malleability of developing dendrites. Mooney et al. 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 led 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 sect. 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 L. C. 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 L. C. 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 prespecifying it. Similar themes of dendritic de-

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

Study	System	Remarks
Valverde 1968	mouse visual cortex, stellate cells	enucleation resulted in dendrites redirected toward remaining afferents
Ruiz-Marcos & 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 & Greenough 1972	occipital cortex	enriched environmental rearing resulted in increased higher order branches
Greenough & Volkmar 1973	occipital cortex	follow up of Volkmar & Greenough (1972); found most increase in basal dendrites of pyramidal cells
Borges & Berry 1976; 1978	rat visual cortex, layer IV stellate cell	dark rearing reoriented dendrites toward residual input
Uylings 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 & Woolsey 1981	mouse somatosensory cortex	vibrissal removal results in reduced representation in corresponding barrel cortex with increase in spared vibrissae
Conlee & 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 4 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

velopment 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 (L. C. Katz & 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.<sup>4</sup>

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

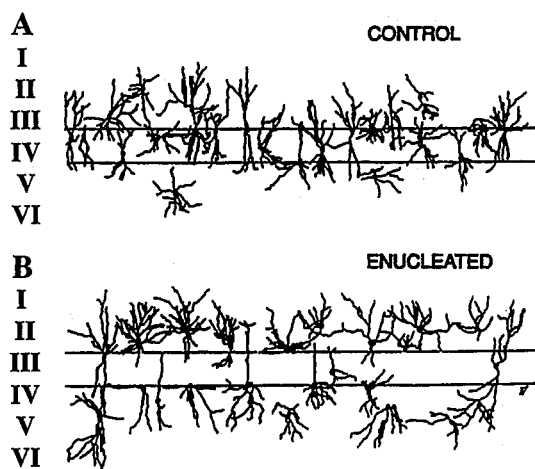


Figure 8. Dendritic organization in visual cortex of normal mice (A) and enucleated mice (B). The degenerative afferent termination is evident in B, where layer IV is sparsely covered with dendrites, whereas adjacent layers are more heavily covered, suggesting that these dendrites have reorganized according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

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 nondeprived 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) concluded, 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 coactive synapses has come to be a central developmental theme in both experimental work (reviewed in Cline 1991; Shatz 1990) 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 Cline 1991; Shatz 1990) 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 3 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, L. C. Katz and Constantine-Paton (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 (1989) 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., Balazs et al. 1989; Bulloch & Hauser 1990; Pearce et al. 1987).<sup>5</sup>

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" (Durbin & Mitchison 1990; reviewed in Churchland & Sejnowski 1992; Mead 1989; Mitchison & Durbin 1986). 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 subregions that function as autonomous processors. Local dendritic segments might be the brain's basic computational units (see also Jaslove 1992; Koch et al. 1982; 1983; Mel 1992a; 1992b; 1994; Segev et al. 1995; Shepherd & Brayton 1987). 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  $\text{Ca}^{2+}$  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 (Bernardo et al. 1982; Wong et al. 1979). 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 (Durbin & Rumelhart 1989; Feldman & Ballard 1982; Koch & Poggio 1992; Mel & Koch 1990). 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 connect this 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: Electronically 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 subserve 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 Fig. 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 (Manzoni et al. 1992; Montague et al. 1994; O'Dell et al. 1991).

Such a spatial signal has a number of attractive properties from a developmental and computational perspective (Montague et al. 1991; Montague 1996) 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.<sup>6</sup> This sort of

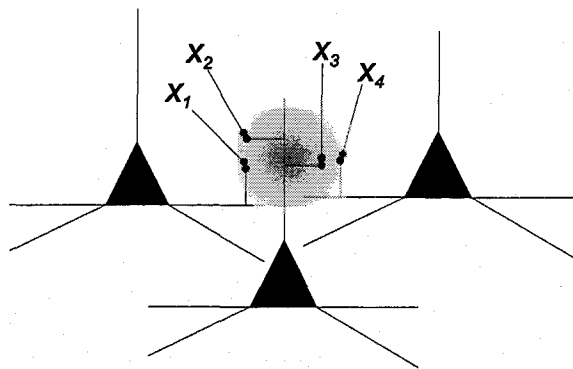


Figure 9. A diffusible substance allows synapses in a local volume of tissue to communicate whether or not they share a connection. Using such a signal, it is possible for synapse  $X_1$  and  $X_4$  to modify their weights according to an associative learning rule. (From Montague & Sejnowski 1994.) One such rule is:

$$\Delta w(t) = \eta [x(t) - \theta_{pre}] [\mu(r; t) - T_{pre}],$$

where  $\Delta w(t)$  is the change in the "weight" or synaptic efficacy of a connection,  $\eta$  is a constant controlling rate of change of synaptic efficacy,  $x(t)$  is a measure of presynaptic activity, and  $\theta_{pre}$  is a threshold that determines whether a terminal is active at time  $t$ .  $T_{pre}$  is a threshold, dependent on the activity of the presynaptic terminal, which determines the direction of synaptic change. The postsynaptic factor of typical Hebbian rules has been replaced by a term for substance concentration,  $\mu(r; t)$  at time  $t$  located at position  $r$ .

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.

**3.1.1. Simulating dendritic growth.** 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 4 months compared to 18 months for layer III cells.

Much of nonvisual 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 2 years of age (Schade & van Groenigen 1961). In addition, the extent of their postnatal development is dramatic – they grow to over 30 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 to 8 years. Why, then, is human nonvisual 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 prewired modules, patterned activity builds up increasingly complex circuits, with areas staging their development. Cortical areas farther 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 6 months of age; but those in language areas, farther 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 (Sperry 1963). 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, is 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.<sup>7</sup> 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 (Dailey & Smith 1996; O'Rourke et al. 1994; O'Rourke & Fraser 1986; 1990; Purves & Hadley 1985; Purves et al. 1986; 1987) 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., Barkow et al. 1992; Fodor 1983; Gazzaniga 1992; Hirschfeld & Gelman 1994). Although Piaget's legacy remains undeniable in developmental psychology (e.g., Bates & MacWhinney 1987; Fischer 1980; 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 is constrained by the learning-theoretic pressures confronting any learner (Gold 1967; Osherson et al. 1986; Pinker 1979; Wexler & Culicover 1980). 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 worst-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  $10^{100}$  possible grammars even in an extremely simplified case – a computation that could never actually be performed. Learners are so slow because of the general strategy they adopt. 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 the consequences are ambiguous because of some major limitations. Even ignoring such dubious assumptions regarding the psychology of learning, there are two internal limitations: the concern of Gold's model merely for convergence in the limit and its 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 and which, accordingly, became 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?

**4.1.1. Adding a neurobiological constraint.** 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 (Schlaggar & O'Leary 1991; reviewed in O'Leary et al. 1992; Stanfield & O'Leary 1985), cross modal rewiring (Frost 1982; Pallas et al. 1990; reviewed in Shatz 1992; Sur et al. 1990; Roe et al. 1990, 1992; Sur et al. 1988), 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. Although 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 section 4.1, 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 postnatal 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 predefined 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; Girosi & Poggio 1990; Hornik et al. 1989). 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 that 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 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., Azimi-Sadjadi et al. 1993; Fahlman & Lebiere 1990; Frean 1990; Hirose et al. 1991; Kadirkamanathan & Niranjana 1993; Platt 1991; Shin & Ghosh 1995; Shultz et al. 1994; Wynne-Jones 1993).

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, tailoring 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 decoding problems. 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 the 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 suggest 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.

**4.3.1. Reconsidering the environment's role.** 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 the 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 show, however, there appears to be far more structure latent in the environment than the poverty of the stimulus arguments supposes. In addition, we think the details of human cortical development provide 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; Tooby & Cosmides 1992) 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 evolve 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 Fig. 3). Unfortunately, neurobiologists abandoned this extreme view of neural specificity 25 years ago (see Easter et al. 1985

for a review). The recent comparative analysis of Finlay and Darlington (1995) lends 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 instantiation").

The extent and duration of large-scale brain changes during development have also been underappreciated. Whereas many researchers believe that the major events in brain development end by 2 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. Please address correspondence to Quartz.

# NOTES

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 nondeprived 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 and 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 L. C. Katz & Constantine-Paton 1988, p. 3178). The conclusion according to L. C. 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 & 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 and 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).

more flexible representational structures. Modular architectures can develop and evolve in tandem with sub-cortical specialisation. I present comparative evidence for the co-evolution of specific thalamic and cortical visual pathways.

The Quartz & Sejnowski target article contains the following chain of argument:

1. Constructivism – the creation of brain organisation through developmental interaction with the sensory environment – is a crucial and perhaps the dominant process in cortical development.

2. Hence the mind/brain cannot contain specialised innate modules.

3. Mammalian brain evolution therefore does not consist of an increase in innate, specialised circuits but is best thought of as "a progression toward more flexible representational structures" (Abstract).

Although there may be something in the first point, the second and third points do not follow from it. At the heart of the misunderstanding lies that old chestnut, the dichotomy between nature and nurture. Although Q&S state that "neural constructivism rejects this dichotomy" (sect. 1.1) and that the cortex is not a *tabula rasa*, they write as if the environment had some intrinsic structure that was impressed upon cortical development subject only to general constraints of neural development. This cannot be correct because the developing cortex's experience of the environment has to be mediated by a system that already has its own intrinsic structure: sensory input. Otherwise, a mole and a rat reared in identical conditions would have the same basic cortical structure, whereas, of course, that is totally implausible; moles could never develop the cortical structures for visual processing that rats have, if only because their eyes and subcortical systems are so different. Whereas the developing cortex may have a high degree of equipotentiality *in theory*, in practice this is never realised because, from the point of view of the developing cortex, environmental structure is inseparable from subcortical structure.

This is not just a pedantic point; it is crucial for understanding how constructive processes could build innate modules – albeit modules that learn – and how modular architectures can evolve. Q&S recognise this point when they refer to "enslavement by the periphery," but then, puzzlingly, they ignore its role in creating a modular architecture. Equally puzzling is the logic behind their emphasis on the effects of experimental manipulations of afferents on the information processing properties of specific cortical regions. Are Q&S really saying, for example, that colour vision should not be considered an innate specialisation of some species because ablating the parvocellular inputs to V4 during development would make its neurons responsive to something other than wavelength? The point is that modular architectures can develop without a precise blueprint for cortical wiring; the design is partly inherent in the structure and properties of subcortical connections. The evolutionary implication is that cortical specialisation goes hand-in-hand with subcortical specialisation. It is a common mistake to view mammalian brain evolution as being solely about the elaboration of the neocortex, with subcortical "primitive" structures remaining essentially unchanged. In fact, the two have co-evolved. For example, in a recent comparative study of primates (Barton, in preparation), I found a close link between neocortical evolution and thalamic specialisation: controlling for differences in overall brain size, evolutionary changes in neocortical size are correlated with evolutionary changes in the parvocellular, but not the magnocellular, laminae of the lateral geniculate nucleus (Fig. 1). What this shows is that the primate neocortex has co-evolved with a specific visual input system. Hence neocortical evolution involves increasing specialisation, not just an all-round progressive increase in global processing power. The brain does, after all, consist of "a hodgepodge of specialized circuits, each chosen by evolutionary pressures" (sect. 4.4), though I would prefer to use the phrase "highly interconnected set" rather than "hodgepodge."

## Open Peer Commentary

*Commentary submitted by the qualified professional readership of this journal will be considered for publication in a later issue as Continuing Commentary on this article. Integrative overviews and syntheses are especially encouraged.*

### Neural constructivism: How mammals make modules

Robert A. Barton

Department of Anthropology, University of Durham, Durham DH1 3HN, England. r.a.barton@durham.ac.uk

**Abstract:** Although the developmental arguments in the Quartz & Sejnowski (Q&S) target article may have intrinsic merit, they do not warrant the authors' conclusion that innate modular architectures are absent or minimal, and that neocortical evolution is simply a progression toward