BEHAVIORAL AND BRAIN SCIENCES (1997) **20,** 537–596
Printed in the United States of America

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 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 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-

© 1997 Cambridge University Press 0140-525X/XX $9.00+.10

https://doi.org/10.1017/S0140525X97301587 Published online by Cambridge University Press

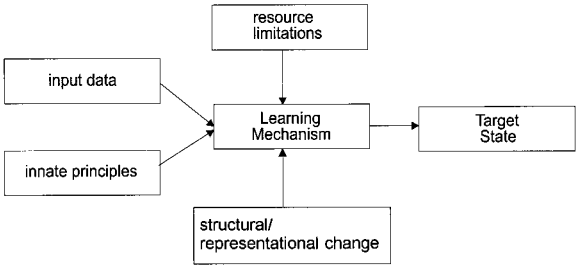


Figure 1. The logic of the problem of development. At its most abstract, the problem is defined as one of characterizing the mapping from some set of input data into the target state (the adult competence; see Wexler & Culicover, 1980, for such a treatment). This problem is complicated by two elements 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

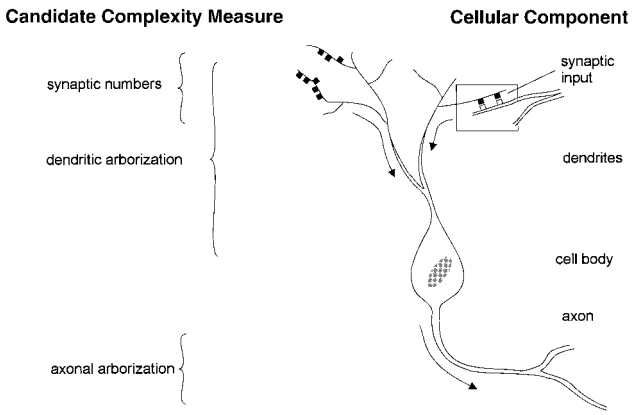


Figure 2. Candidate measures of structural/representational complexity. There are three possible (nonexclusive) measures: synaptic numbers, dendritic arborization, and axonal arborization. The figure also summarizes the basics of neural information processing elements: synaptic input travels along the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

https://doi.org/10.1017/S0140525X97301587 Published online by Cambridge University Press

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; Shatz & Stryker 1978). To avoid a proliferation of terminology, we will refer to any view positing an initial exuberant growth followed by elim-

https://doi.org/10.1017/S0140525X97301587 Published online by Cambridge University Press

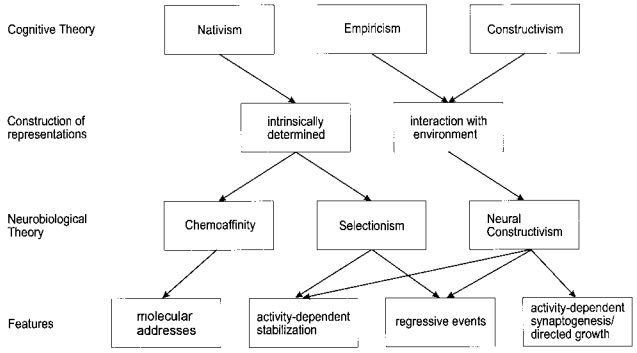


Figure 3. Summary of the relations among major cognitive and neural theories. Instructivism is often identified with Hebb's work, since in his view the environment guides or instructs brain growth. We refer to such theories as constructivist since the latter is a more widely recognized class of theories.

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 Lamarkian 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

https://doi.org/10.1017/S0140525X97301587 Published online by Cambridge University Press

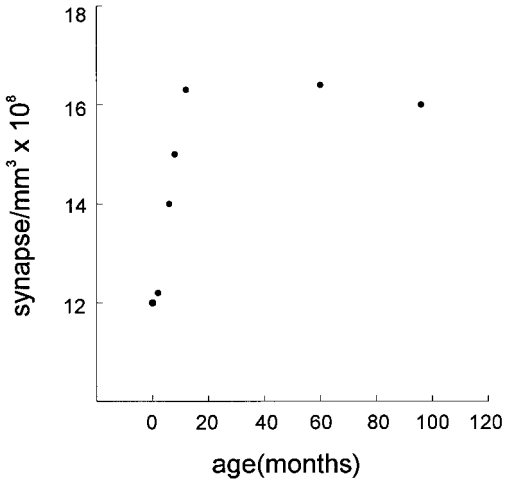


Figure 4. Synaptic density in layer III, human middle frontal gyrus from birth to 8 years. Data from Huttenlocher (1979). Note the absence of data points between 1 and 5 years of age.

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-

https://doi.org/10.1017/S0140525X97301587 Published online by Cambridge University Press

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	mouse visual cortex, stellate cells	decrease in spine density, numbers in dark rearing
Globus & Scheibel 1967	rabbit visual cortex	visual deprivation resulted in deformed spine morphology
Valverde 1971	mouse visual cortex, layer V pyramidal cells	mice raised in darkness for 20 days attain normal spine values when returned to normal environment
Globus et al. 1973	rat visual cortex	increase in spine density, numbers in complex environments
Cragg 1975	cat visual cortex	bilateral suture or deafferentation, 30% reduction in the number of synapses per neuron
Lee et al. 1980	CA1, hippocampus	increased synapse numbers with long-term potentiation
Winfield 1981	cat visual cortex	bilateral eye suture slows rate of development and decreases synapses per neuron by 32%
Chang & 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;

https://doi.org/10.1017/S0140525X97301587 Published online by Cambridge University Press

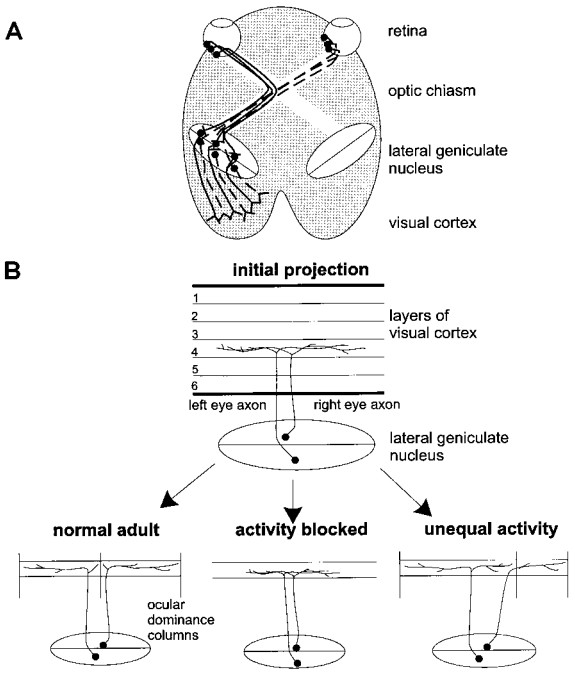


Figure 5. The generation of specificity according to the standard model. A, schematic representation of the connections between retinal ganglion cells and the lateral geniculate nucleus of the thalamus and connections to primary visual cortex. B, the two-phased process of development and perturbations 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 retinogeniculate pathway (see Table 2). More recently, Antonini and Stryker (1993) confirm this more constructivist interpretation, and place even more emphasis on directed growth than the earlier studies (see p. 3572 for discussion). Using a technique known as bulk-filling and anterograde transport, they found some retraction of early axonal structure, but they also found substantial progressive increases in axonal measures throughout development. During the first week of segregation between left and right eye input, axonal branch point numbers increase strikingly from a median value of 34.5 to one of 96, while total arbor length increased from 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:

https://doi.org/10.1017/S0140525X97301587 Published online by Cambridge University Press

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

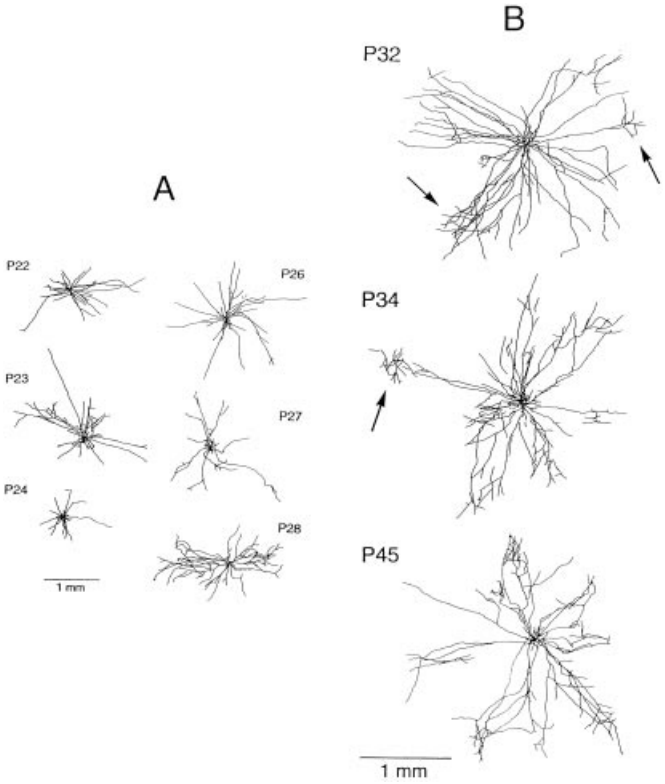


Figure 6. The development of horizontal connections in the ferret. A, at postnatal day 22 (P22) little clustering can be seen. As development proceeds, however, axon cluster formation becomes evident, particularly by P28. B, subsequent development, seen here under higher magnification, reveals the refinement 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.[1]

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

2.3. Dendritic development

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

https://doi.org/10.1017/S0140525X97301587 Published online by Cambridge University Press

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 nonlinear ways, their processing properties make them central to how information is processed by neural systems. It is essential, then, to assess the developmental processes that shape dendritic form and the role of these processes in determining the representational properties of neural circuits.[2]

2.3.2. Dendritic arborization. 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 synapses (Segev et al., 1995, contains an excellent collection of classic and contemporary views on dendritic function).

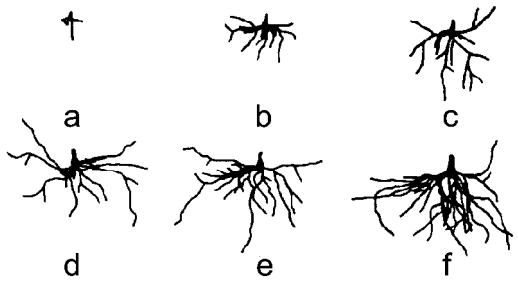


Figure 7. Camera lucida drawings of basal dendrites of layer V human pyramidal cells; *a*, newborn; *b*, 3 months; *c*, 6 months; *d*, 15 months; *e*, 24 months; *f*, adult (from Schade & van Groenigen 1961).

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.[3] In the following sections, we accordingly examine the mode of dendritic development in some detail – the extent of progressive processes at the level of dendritic structure and their malleability by changes in activity. From this, we go on to formulate some features of dendritic development, considering their cellular basis, and relating these to the learning and representational properties of cortex.

2.3.3. Environmental conditions and dendritic development. 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

https://doi.org/10.1017/S0140525X97301587 Published online by Cambridge University Press

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 & Fifkova 1979	rat visual cortex	progressive dendritic expansion of pyramidal cells, layers II-III, V
Simonds & Scheibel 1989	Broca's area	progressive dendritic expansion into the sixth year
Parnavelas & Uylings 1980	rat visual cortex, layer IV stellate cells	progressive dendritic development lasting to postnatal day 20, followed by distal expansion to postnatal day 90
Buell & 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 μm 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 μm 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-

https://doi.org/10.1017/S0140525X97301587 Published online by Cambridge University Press

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 acoutic 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.[4]

An interpretation of these results is that dendritic segments function as detectors of correlated activity and grow preferentially in such regions. Support comes from Tieman and Hirsch's (1982) finding that exposure to lines of only one orientation during development has specific effects on dendritic development. The dendritic field orientations of cells from cats raised with exposure to lines of a single orientation were significantly elongated in conformity with this shift in the visual environment.

An insight from this study is that a dendritic tree samples its input space actively in response to the environmental structure. A similar result has been obtained for layer IV stellate cells by Coleman et al. (1981), who suggest (p. 19): "[I]f an alteration of the spatio-temporal pattern of the afferent activity is sufficient to lead to dendritic alterations

https://doi.org/10.1017/S0140525X97301587 Published online by Cambridge University Press

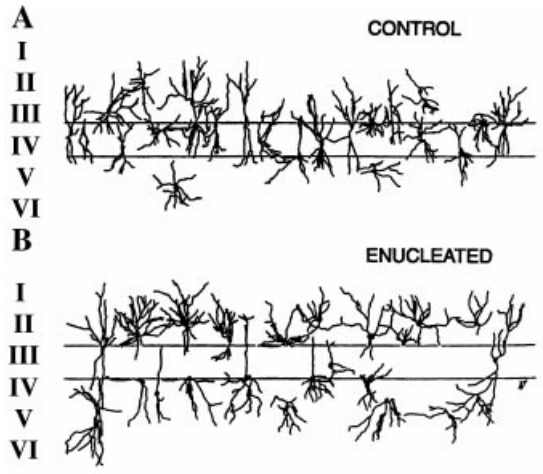


Figure 8. Dendritic organization in visual cortex of normal mice (A) and enucleated mice (B). The degenerative afferent termination is evident in B, where layer IV is sparsely covered with dendrites, whereas adjacent layers are more heavily covered, suggesting that these dendrites have reorganized according 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).[5]

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

3. Directed dendritic development and representational change

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

https://doi.org/10.1017/S0140525X97301587 Published online by Cambridge University Press

"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 Ca^{+} 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: Electrotonically close synapses encode common features of the input space and effectively fire the cell (Mel 1992a; 1992b; 1994). The presence of cluster-encoded features significantly alters both the representational properties of cortex and its computational power.

3.1. Developmental mechanisms underlying clustering

For clustering to serve an interesting information processing role, there must be an appropriate developmental scheme underlying its formation. For this, temporal patterns of activity must somehow be translated into appropriate spatial patterns of termination. As Mel states (1992b, p. 43):

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

A number of Hebbian schemes have been proposed to 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.[6] This sort of

https://doi.org/10.1017/S0140525X97301587 Published online by Cambridge University Press

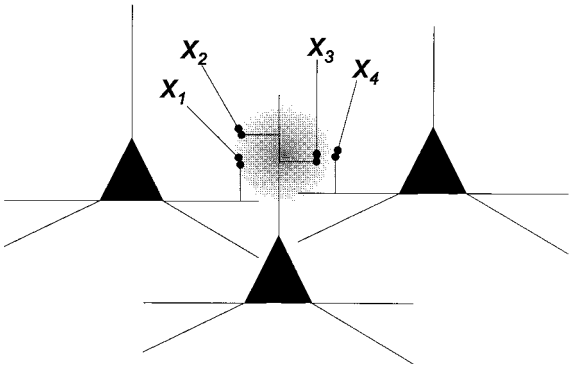


Figure 9. A diffusible substance allows synapses in a local volume of tissue to communicate whether or not they share a connection. Using such a signal, it is possible for synapse X_1 and X_4 to modify their weights according to an associative learning rule. (From Montague & Sejnowski 1994.) One such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

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

where $\Delta w\,(t)$ is the change in the "weight" or synaptic efficacy of a connection, η is a constant controlling rate of change of synaptic efficacy, $x(t)$ is a measure of presynaptic activity, and θpre is a threshold that determines whether a terminal is active at time t. *T*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

https://doi.org/10.1017/S0140525X97301587 Published online by Cambridge University Press

intrinsically generated structures.[7] Rather, selection is only one kind of process in a *dynamic interaction* between environmentally derived activity and the neural growth mechanisms that activity regulates. This changes the emphasis from synaptic elimination to synaptic turnover. New techniques that allow the same structures to be examined successively over time (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

https://doi.org/10.1017/S0140525X97301587 Published online by Cambridge University Press

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

https://doi.org/10.1017/S0140525X97301587 Published online by Cambridge University Press

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 loading 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

https://doi.org/10.1017/S0140525X97301587 Published online by Cambridge University Press

learners are impressive (e.g., Azimi-Sadjadi et al. 1993; Fahlman & Lebiere 1990; Frean 1990; Hirose et al. 1991; Kadirkamanathan & Niranjan 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

https://doi.org/10.1017/S0140525X97301587 Published online by Cambridge University Press

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 instantiability").

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.

https://doi.org/10.1017/S0140525X97301587 Published online by Cambridge University Press

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).

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 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 constuctivism 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."

https://doi.org/10.1017/S0140525X97301587 Published online by Cambridge University Press

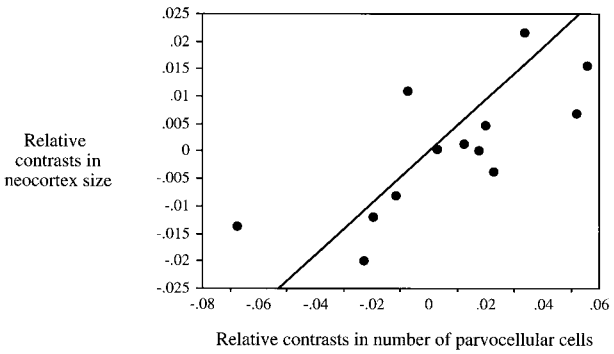


Figure 1 (Barton). Co-evolution of parvocellular laminae of the lateral geniculate nucleus and neocortex in primates. The graph plots evolutionarily independent contrasts in neocortex size and in the number of parvocellular neurons, corrected for overall brain size. Data on neocortical size and number of parvocellular neurons were extracted from Stephan et al. (1981) and Shulz (1967) respectively. Independent contrasts were computed using the C.A.I.C. computer programme (Purvis & Rambaut 1995); these contrasts are the standardised differences between taxa at each node on a bifurcating phylogeny, to which end the composite phylogeny of primates published by Purvis (1995) was used. Contrasts so produced were corrected for gross brain size by regressing them on contrasts in the size of the rest of the brain; the points on the graph are based on the residuals from this regression: $r^2 = 0.53$, $p = 0.0008$. It is important to note that no such correlation was found between neocortical size and magnocellular laminae ($p = 0.23$), indicating that neocortical adaptive radiation in primates is specifically linked to parvocellular specialisation.

In rejecting extreme nativist propositions invoking "predispositions somehow embedded *a priori* in the recipient cortex" (sect. 1.1), and "pre-wired modules" (sect. 3.2), Q&S describe a view that I doubt any evolutionary psychologist would actually espouse. Perhaps the concept of specialised modules has been oversimplified by some, and more sophisticated approaches (e.g., Karmiloff-Smith 1992) are preferable. [See also *BBS* multiple book review of Karmiloff-Smith's *Beyond Modularity*, *BBS* 17(4) 1994] A strength of the modularity concept is that it allows one to evade the trap that the target article falls right into; seeing brain evolution as a progressive process, with species ranked on a linear *scala naturae* of global processing capacity. The fact that cortical structure and function vary between species, with radically different emphasis on different sensory modalities and specialised forms of stimulus representation, means that there is more to cortical evolution than simply "a progression toward more flexible representational structures."

Constructivism, nativism, and explanatory adequacy

Derek Bickerton

Department of Linguistics, University of Hawaii, Honolulu, HI 96822.
derek@hawaii.edu

Abstract: Constructivism is the most recent in a long line of failed attempts to discredit nativism. It seeks support from true (but irrelevant) facts, wastes its energy on straw men, and jumps logical gaps; but its greatest weakness lies in its failure to match nativism's explanation of a wide range of disparate phenomena, particularly in language acquisition.

Cognitive linguistics, connectionism, constructivism: C-men continue to assault the post-Skinnerian nativist consensus. But Quartz & Sejnowski (Q&S), like their predecessors, prefer long-distance bombardment to the bloody hand-to-hand fighting it would take to storm Fort Innateness. Consequently, most of their shots sail overhead or burst harmlessly on empty trenches.

For the versions of nativism and selectionism Q&S target are blatant straw men. No serious selectionist claims neural pruning as the sole process in brain development, just as no serious nativist denies that a vast amount of learning must occur if hard-wired mechanisms are to function effectively in the real world. Barkow et al. (1992) should surely have squelched the idea that the less the brain knows to begin with, the more it can learn – on the contrary, it's only hard-wired, task-specific functions that allow any creature to know anything at all about the world.

Worse still, Q&S force their straw men into a shotgun marriage. While it's true that early versions of a language acquisition device (Chomsky 1965) could be viewed as selectionist, the two positions are wholly dissociable and clearly dissociated in updated models. A post-Borer (1984) model assumes invariant principles that do not require selective pruning or any other learning process, plus a grammatical morphology whose forms and properties have to be plain old-fashioned learned, even given the aid of some (still unspecified) constraints on the concept "possible grammatical morpheme." Such models neither assume nor need selection.

Well over a third of the target article is devoted to three measures of representational complexity. Q&S point out that selectionism does poorly on the first (synaptic numbers), but they fail to note that if, as they claim, these remain constant over the developmental period, constructivism fares equally badly. With regard to axonal arborization (sect. 2.1.1), the best they can claim is "more balance between both selective elimination and selective growth," a statement with which few will be tempted to quarrel. Only dendritic arborization seems to increase significantly over time.

But how relevant is this? A logical gap yawns between the analysis of cortical development and the claim that nativists overstate the case for prespecification. For that matter, there is nowhere any clear explanation of how arboreal density and cognitive growth are related. Suppose one were to claim that denser arborization, like enlarged breasts or pubic hair, merely expresses a genetically determined developmental program, and that enhanced appreciation of environmental structure is effect rather than cause? (I'm not claiming this – just curious as to how Q&S would answer.)

Mentioning "environmental structure" opens another can of worms. Q&S's repeated use of this term (or derivatives thereof) invokes, without actually materializing, ghosts of Gibsonian "affordances": just how structured do they think the environment is, and how uniformly is it structured? In some areas, much of the structure is imposed by the observing organism, color being the obvious example. Other areas (such as syntactic input) have, if anything, too much structure and yield a plethora of plausible analyses that defy induction. At last count, there were at least seventeen different theories of syntax, and if the experts can't agree, what's the probability that naive infants will universally hit on the right answer?

Indeed, language acquisition is where Q&S really come to grief. They seem to think that the demolition of Gold's (1967) learning-theoretic argument will, like Joshua's trumpet call, bring down the stoutest outpost of innateness and perhaps the whole edifice along with it. Nothing could be further from the truth. Few linguists cite, and still fewer understand, Gold's paper. In order to effect a breach here, Q&S would have to explain at least the following:

a. How creole languages are acquired.
b. How the first human language was acquired.
c. Why Williams syndrome subjects have perfect syntax.
d. Why apes can't acquire syntax.
e. Why all languages, however superficially diverse, obey the same syntactic principles.
f. How children acquire the referential properties of unexpressed morphemes (e.g., the subject and object, marked by *e*, in a sentence such as *John wants someone e to love e*).

https://doi.org/10.1017/S0140525X97301587 Published online by Cambridge University Press

g. Why input-deprived children (Curtiss 1989) can acquire phonology and morphology, but not syntax.

By "explain" I don't mean "explain away." Scholars as astute and as familiar with the cognitive literature as Q&S should find no difficulty in chipping at each issue individually, casting doubt on the validity of one, raising apparent counter-examples to another, attributing a third to general principles ("syntactic principles are the same everywhere because the environment is the same everywhere"), and so on.

My point is precisely this: nativism can easily handle (a) through (g) (and many more phenomena) in terms of a single coherent theory. I doubt whether constructivism could give a unified account of even three or four of the phenomena listed. In the physical sciences, one key criterion for theories is the ability to coherently yoke together a wide range of apparently disparate facts. If the bottom line is (to quote Chomsky 1965 once more) "explanatory adequacy," then nativism, or at least the linguistic variety thereof, stands clearly superior to constructivism.

The nearest that Q&S come to denting nativism lies in their generalization from a syntactic module to a whole slew of innately specified modules. Indeed, outside of language and the various sensory and motor systems, evidence for modularity is much weaker. Could the authors have argued that, among mental faculties, only language constitutes an innate module, and that this module, interacting with sensory and motor systems as well as with the extrasomatic environment, yields a general learning device unique to our species that would account for our vast cognitive superiority? That would, for me at least, have made a more original, more stimulating, but very different article. For this one, the verdict must be, "No cigar, not even close."

How to build a brain: Multiple memory systems have evolved and only some of them are constructivist

James E. Black[a] and William T. Greenough[b]

[a]*Department of Psychiatry,* [a,b]*The Beckman Institute, and the Neuroscience Program;* [b]*Departments of Psychology, Psychiatry, and Cell & Structural Biology, University of Illinois at Urbana-Champaign, Urbana, IL 61801.*

greenou@s.psych.uiuc.edu

[1]soma.npa.uiuc.edu/labs/greenough/home.html

Abstract: Much of our work with enriched experience and training in animals supports the Quartz & Sejnowski (Q&S) thesis that environmental information can interact with pre-existing neural structures to produce new synapses and neural structure. However, substantial data as well as an evolutionary perspective indicate that multiple information-capture systems exist: some are constructivist, some are selectionist, and some may be tightly constrained.

In arguing against selectionist models, Q&S have overshot the mark, oversimplifying the complex processes of brain development and evolution. We have suggested that multiple systems store information in the mammalian brain (Black & Greenough 1986). Q&S minimize the messiness of evolution (sect. 4.4); brains are not designed *de novo* for elegance or efficiency. Mammalian brain structure is *enormously* complicated, and its cobbled-together qualities reflect an evolutionary history that adapted whatever brain components already existed. Given mammals' long evolutionary history, it is likely that multiple forms of neural plasticity (or modified versions of them) have been retained. (We ignore here the technical misunderstandings in the "Manifesto," including a failure to discriminate between synaptic density and synapse number.) [See also Ebbesson: "Evolution and Ontogeny of Neural Circuits" *BBS* 7(3) 1984.]

Experience-expectant processes. We have argued that the major mammalian development and learning processes group into two major types, experience-expectant and experience-dependent, with different underlying neural processes corresponding roughly to selectionism and constructivism (Black & Greenough 1986). Experience-expectant information storage is designed to capture information usually available to all members of a species (such as patterned light) in brain structure. These processes correspond to the selectionist models described by Q&S, with a developmentally timed overproduction of synapses followed by a selective pruning back of connections. These selectionist processes have been observed in many mammalian species and in many brain regions, and almost certainly occur in many other cases in which they are masked. For example, Greenough and Chang (1988), cited in the target article as support for "directed dendritic development" (see below), reported *simultaneous* loss and growth of somatosensory cortical dendrites, clearly a process involving "selectionism." Rather than being unimportant, selectionist processes seem to set the stage upon which later development builds. Experience-expectant information storage typically involves a brief time span (roughly corresponding to a "sensitive period") when the organism is maximally ready (i.e., developmentally "expects" or is sensitive) to receive appropriate information, so that experience-related neural activity can select functionally appropriate subsets of synapses. Early visual experience is an example of "expected" experience, and many of the experiments described in Q&S's article deprive animals of visual experience that is normally available for all species members. Experiments such as Flood et al.'s (1963) demonstrate the importance of developmental scaffolding that sequentially enables the organism to extract reliable information from its environment.

Experience-dependent processes. Experience-dependent information storage optimizes an individual's adaptation to specific and sometimes unique features of the environment, corresponding closely to constructivist models. Experience-dependent information storage does not involve sensitive periods, although there may be sequential dependencies. Mammals are particularly suited to and adapted for this form of information storage. We would argue, in rough agreement with Q&S, that humans, in particular, have a cerebral cortex (with associated regions) that has evolved to form a massive "memory organ," one capable of incorporating in its synaptic connections the complexity of language, culture, and enormous amounts of specialized knowledge. We also agree with Q&S that much of this information storage occurs through constructivist mechanisms after selectionist mechanisms have shaped and scaffolded the cortex.

An experimental platform for the study of constructivist processes is the complex or "enriched" environment (EC) rearing paradigm. Literature on complex environments is cited in Tables 1 and 4 of the target article, but it is mixed in with studies of dark rearing, aging, and long-term-potentiation, confusing the distinction we make here. The modification of experience in differential rearing has a character that is largely not "expected" in development, that is, individual experience in the EC environment cannot be assumed to be available (or adaptive) to all species members. In fact, the complex environment resembles to some extent the "natural" environment rats evolved in, which is full of unexpected, idiosyncratic things to learn, for example, the location of warrens, food, or predators. The specific experiences of animals raised in EC differ from those of cage-reared animals primarily in the complexity of experience available; and self-initiation of experience (e.g., exploratory activity or play) is a key determinant of the availability and quality of experience. The metabolically most efficient way to generate synapses within the system would be for patterned activity, along with some neuromodulatory signal, to trigger local synaptogenesis, a proposal consistent with constructivist models. The dendritic/synaptic alterations seen after EC experience are also seen after behavioral training (Tables 1 and 4), suggesting common experience-dependent mechanisms.

A fundamental question is whether this synaptogenesis (or dendritic growth) is "directed" as Q&S argue. No data of which we are aware rule out the possibility that the net synapse addition seen in EC or trained rats arises through local overproduction and

https://doi.org/10.1017/S0140525X97301587 Published online by Cambridge University Press

subsequent selection mechanisms. The tissue culture models cited argue precisely for this – supernumerary neuritic processes are stabilized by appropriate targets. It seems quite possible that overall growth of dendrites and synaptic numbers in mature brains involves the generation of multiple dendrites and synapses followed by activity-dependent selection of a subset, an overlapping sequence of synaptic blooming and pruning. At the cellular level, this would imply that "constructivist" processes utilize both construction and selection.

Implications for the constructivist manifesto. Q&S argue that "the brain is not a hodgepodge of specialized circuits, each chosen by evolutionary pressures" (sect. 4.4), but we believe the evidence weighs against them. It seems clear that multiple mechanisms have evolved to store memory, probably through co-opting mechanisms previously serving other purposes. The existence of various (perhaps independently evolved) mechanisms of information storage across divergent species (e.g., honeybees, sea slugs, and humans), in many different neural structures, influenced by a variety of modulatory factors, strongly suggests that evolution has established multiple and different mechanisms for neural plasticity. We believe that good evidence exists for selectionist processes to play essential roles in development, as well as many other types of systems.

More important, we believe that the case for directed neural growth underlying constructivist growth has not been made. Ultimately, the *pattern* of connections is most important, and mathematical models need not distinguish between directed neural growth and localized overproduction and selection of synapses if the overall network is growing. However, for those of us interested in the neurobiology of memory, this distinction is crucial and still unresolved.

NOTE

1. Please send correspondence to W. T. Greenough, Beckman Institute, 405 N. Mathews, Urbana, IL 61801.

Dynamical learning algorithms for neural networks and neural constructivism

Enrico Blanzieri

Centro di Scienza Cognitiva, Universita' di Torino, 10123 Turin, Italy.
blanzier@di.unito.it

Abstract: The present commentary addresses the Quartz & Sejnowski (Q&S) target article from the point of view of the dynamical learning algorithm for neural networks. These techniques implicitly adopt Q&S's neural constructivist paradigm. Their approach hence receives support from the biological and psychological evidence. Limitations of constructive learning for neural networks are discussed with an emphasis on grammar learning.

The arguments that Q&S's target article presents in favour of the *neural constructivist* paradigm and against a strong version of selectionism and nativism are mainly of two types. First, Q&S present a collection of biological and psychological evidence of the growing of the neural system during the developmental phases. Second, according to a computationally based argument, a constructivist learner should support a wider definition of learnability. In particular, the argument tries to reverse the nativism-supporting considerations about the unlearnability of grammars.

This commentary addresses both the arguments and the main claim of the target article from the viewpoint of dynamical learning algorithms for neural networks. These kinds of algorithms are constructive exactly in the sense presented in the target article: new units, typically the hidden ones, are added to the network, changing its representational power and the space in which the nondynamical learning algorithm searches for the optimal or suboptimal parameters. Constructive algorithms have been presented for standard popular architecture such as the multilayer perceptron as well as for more exotic neural networks, but always with the implicit hypothesis of the existence of a pool of unused units that can be "inserted" into the network when needed.

Alas, the "pool" hypothesis only worked mathematically because of the absence of any biological support for it. The biological and psychological evidence that Q&S present in the target article support this kind of approach. The authors move the level of the construction from the units to the connections (synapses, axons, and dendrites): there is no longer any need for a "unit pool," the structural changes are elsewhere. In particular, the considerations about the dendritic growth and the speculations about their nonlinear computational capabilities suggest that even nonstandard networks (i.e., those not based on the multilayer perceptron) can be biologically plausible. For example, radial basis function networks (RBFNs) fit Q&S's locality and stability conditions (sect. 2.3) as well as the clustering property. A constructive algorithm for RBFN has been proposed (Blanzieri & Katenkamp 1996; Fritske 1994) and shown to be effective.

Q&S cite constructive neural networks (sect. 4.2) and describe their asymptotic approximation properties. Q&S correctly emphasise that the universal approximation properties of different kinds of neural networks are based on the possibility of changing the structure of the network by adding hidden units. This fact supports constructive learning theory, but it is worth noting that the results are limited to feedforward networks and hence to the approximation of arbitrary mappings between multidimensional domains. The learning of grammars is a different and harder problem even for constructivist neural networks. The solution requires recurrent networks and the expressive power that is necessary for effectively encoding grammars is not known. For an example of the relation between finite state grammars and recurrent nonconstructivist radial basis function networks RBFN see Frasconi et al. (1996). Hence the argument about the learnability of the grammars by a constructivist learner seems to be controversial and not decisive against the nativism.

In conclusion, the target article provides biological and cognitive support for already existing neural network research. It fails to show the inconsistency of nativism, but it presents a good working hypothesis that does not require it. Further research is needed to design constructive neural networks that learn grammars.

Learning, development, and synaptic plasticity: The avian connection

Johan J. Bolhuis

Sub-Department of Animal Behaviour, University of Cambridge, Madingley, Cambridge CB3 8AA, England. **jjb19@cus.cam.ac.uk**

Abstract: Quartz & Sejnowski's target article concentrates on the development of a number of neural parameters, especially neuronal processes, in the mammalian brain. Data on learning-related changes in spines and synapses in the developing avian brain are consistent with a constructivist interpretation. The issue of an integration of selectionist and constructivist views is discussed.

Quartz & Sejnowski's (Q&S's) thoughtful target article provides us with a timely evaluation of the empirical facts about the neural basis of cognitive development. Q&S's emphasis is on experimental work in mammals. However, considerable relevant research on the development of brain and behaviour in birds largely supports a constructivist interpretation. This research has been concerned mainly with early learning in the domestic chick and song learning in passerine birds. The advantage of early learning tasks such as imprinting is that the animals can be reared in visual isolation until they are exposed to the training stimulus for the first time; structural changes can then be related directly to the learning experience (Horn 1985).

Filial imprinting in the chick led to a significant increase in the mean size of the postsynaptic density (PSD, a thickening of the

https://doi.org/10.1017/S0140525X97301587 Published online by Cambridge University Press

area of the postsynaptic membrane where there is a high density of receptors) in the intermediate and medial hyperstriatum ventrale (IMHV), a region that is crucially involved in learning during imprinting (Horn 1985). There were no significant learning-related changes in the number of presynaptic boutons or dendritic spines. The increase in PSD was limited to spine synapses; there was no such increase in synapses directly on the shafts of dendrites. Furthermore, the increase only occurred in synapses in the left IMHV, not in the right, which is consistent with other evidence that assigns a long term storage function to the left IMHV (Horn 1985). Similarly, McCabe and Horn (1988) reported a significant positive correlation between the number of NMDA receptors and the strength of imprinting in the left IMHV only.

Using the Golgi method of staining, some authors have reported changes in the number of dendritic spines in some brain areas after early learning. Patel et al. (1988) reported, *inter alia*, a small, nonsignificant increase in the number of dendritic spines per unit length of dendrite in the left IMHV of chicks that had learned a passive avoidance task, compared to chicks that had not. In contrast, Wallhäusser and Scheich (1987) reported a *reduction* in the number of dendritic spines on a particular class of neurones, after an auditory imprinting session in guinea fowl chicks. The changes were found in a forebrain region (MNH) that may have some overlap with the anterior part of the IMHV. Interestingly, a recent combined Golgi and electron-microscopic analysis suggests that although after auditory imprinting there is a reduction in the number of spines in the MNH, the size of the PSDs of the remaining synapses increases (Scheich 1996).

Unlike Q&S (sect. 3), Nottebohm (1991) and co-workers have suggested that the entire neuron is the brain's basic computational unit, when they reported neurogenesis in the brain of canaries and zebrafinches (which occurs for some time after hatching and into adulthood). It is not clear whether these new neurons, which become functionally integrated into existing circuits, are necessary for the central representation of song. For instance, there is no relationship between neurogenesis and seasonality or whether or not the birds modify their songs in adulthood. Also, neurogenesis occurs across the forebrain and not just in song-related regions.

Taken together, the avian data do not support a role for the neuron as a computational unit. However, even if we assume that "clustering" and "volume learning" (sect. 3) are important features in the developing brain, the synapse is still the ultimate unit of neural plasticity and the avian data to which I have referred suggest that changes in synaptic strength are crucially involved in neural representation of information. Although changes in the number of synaptic spines may play a role, it would seem that increases but *probably* also *decreases* in PSDs and concurrent receptor numbers are the changes that matter. Thus, changes in the strength of existing connections are important, in line with theoretical considerations that have been formally incorporated in a model of early learning (Bateson & Horn 1994). As Q&S state, for learning to occur "there must be some built-in assumptions" (sect. 4.3.1). The architecture of the Bateson and Horn model also allows for the influence of predispositions, in line with the behavioural and neurobiological data on the development of filial prefences in the chick (Bolhuis 1994; Horn 1985).

The target article acknowledges the existence of regressive events during learning and development, and their theoretical importance (e.g., sects. 2.1.1; 2.2.1; 2.3.2). How can constructive and regressive findings be reconciled in an "interactionist" theory? Greenough et al. (1987) advanced such a theory in associating synapse selection and synapse formation, respectively, with "experience-expectant" and "experience-dependent" information storage. This distinction is a functional one, based on the type of information stored, information which is ubiquitous and identical for all members of the species in the former, and unique to the individual in the latter. I have argued (Bolhuis 1994) that using these terms at two different levels of explanation limits their usefulness; morever, the evidence from the avian work does not support the distinction. In reality, both "experience-dependent" and "experience-expectant" information may be stored during learning and development. It seems impossible to tease apart the mechanisms involved in these two putative processes, let alone relate them to specific neural changes.

In defense of learning by selection: Neurobiological and behavioral evidence revisited

G. Dehaene-Lambertz[a] and Stanislas Dehaene[b]

[a]*Laboratoire de Sciences Cognitives et Psycholinguistique, 75270 Paris cedex 06, France.* **ghis@lscp.ehess.fr;** [b]*INSERM U334, CHFJ CEA, 91401 Orsay, France.* **stan@lscp.ehess.fr**

Abstract: Quartz & Sejnowski's (Q&S's) constructivist manifesto promotes a return to an extreme form of empiricism. In defense of learning by selection, we argue that at the neurobiological level all the data presented by Q&S in support of their constructive model are in fact compatible with a model comprising multiple overlapping stages of synaptic overproduction and selection. We briefly review developmental studies at the behavioral level in humans providing evidence in favor of a selectionist view of development.

Our understanding of cognitive functions such as language and calculation is likely to make considerable progress once we begin to uncover, not merely the algorithms that underlie their acquisition in the child's mind, but also their biological basis in the child's brain (Changeux & Dehaene 1989; Dehaene & Changeux 1993; Diamond 1988). What, however, is the appropriate theoretical framework in which developmental cognitive neuroscience research should be framed? Quartz & Sejnowski's "constructivist manifesto," although purportedly rejecting the classic dichotomy between empiricism and nativism, actually promotes a return to an extreme form of empiricism, in which a "largely equipotential cortex" (sect. 4.3.1) becomes specialized through "directed growth" (sect. 2.2.1) under the specification of environmental inputs. We strongly disagree with this point of view and would like to present here a brief defense of the selectionist framework of cognitive and neural development (Changeux & Dehaene 1989; Edelman 1987).

Multiple waves of synaptic overproduction and selection. Most of the arguments that Q&S present against learning by selection arise from a misconstrual of selectionist theories of cognitive and neural development. In section 2, Q&S describe a simplistic version of selectionism that "divides development into two discrete stages": first, the intrinsic production of "prerepresentations," and second, their selective elimination or stabilization. Q&S then reach the far-fetched conclusion that "for selectionism, development marks a reduction in representational complexity" (a definition that seems to better capture the manifestations of aging!). This oversimplified characterization is then easily dismissed as being incompatible with numerous data that underscore the progressive enrichment in synaptic numbers, axonal arborizations, and dendritic trees that accompany cognitive development.

Selectionist theories of learning, (e.g., Changeux 1985; Changeux & Dehaene 1989; and Edelman 1987) are considerably more subtle than suggested by Q&S's summary. Most important, they allow for multiple, overlapping waves of synaptic development followed by selection, thus providing a basis for understanding the increase in nervous system complexity during development. Here, for instance, is how Changeux (1985, pp. 248–49) describes this process:

> The 10,000 or so synapses per cortical neuron are not established immediately. On the contrary, they proliferate in successive waves from birth to puberty in man. With each wave, there is transient redundancy and selective stabilization. This causes a series of critical periods when activity exercises a regulatory effect . . . One has the impression that the system becomes more and more ordered as it receives «instructions»

https://doi.org/10.1017/S0140525X97301587 Published online by Cambridge University Press

from the environment. If the theory proposed here is correct, spontaneous or evoked activity is effective only if neurons and their connections already exist before interaction with the outside world takes place. Epigenetic selection acts on preformed synaptic substrates. To learn is to stabilize preestablished synaptic combinations, and to eliminate the surplus.

Once it is recognized that synaptic overproduction and selection can occur simultaneously and in multiple waves, rather than in a discrete two-step process, all the neurobiological phenomena that Q&S claim refute selectionism and support constructivism actually become explainable within a selectionist framework (see Fig. 1). In this framework, the progressive increase in axonal complexity, for instance, is viewed as resulting from a blind process of overproduction of axonal branches, while their increasing targeting to specific layers is taken to reflect the elimination of misplaced or nonfunctional branches.

In addition to explaining the major phenomena of neural growth, selectionism also releases axons and dendrites from a considerable computational load. According to Q&S's constructivism, the nervous system is subject to "directed growth" and dendrites act "as though they are looking for afferent inputs" (sect. 2.3.3). They even "redirect their growth to find active afferents" (sect. 2.3.3) and they "actively seek out incoming activity" (sect. 2.3.4). We think that the use of this mentalistic vocabulary, which treats single dendrites as cognitive agents, makes for a misleading and ill-specified analogy, which in selectionist theories is replaced by a simple, explicit, and empirically testable mechanism for stabilization (see Fig. 1). Like the nervous system, the immune system also seems to react "intelligently" to external inputs – but the mechanism underlying this seemingly directed response is selectionist in nature. Whether neural development will also be accounted for by selectionist mechanisms remains an open issue at present, but our point is that selectionism remains a viable theory in the face of current neurobiological data (for a recent in depth discussion, see Katz & Shatz 1996).

Cognitive developmental evidence for learning by selection. At the cognitive level, the Piagetian notion that mental representations are constructed through a progressive instruction by environmental inputs is now obsolete. Ample evidence indicates that human infants exhibit, very early on, a degree of competence in specific domains of cognitive functioning that seems hard to explain by "constructive learning." One such domain is number

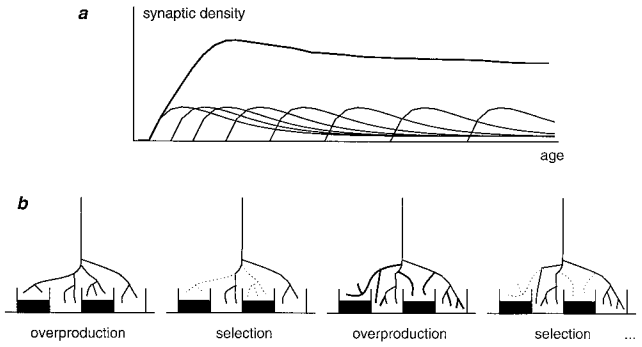


Figure 1 (Dehaene-Lambertz & Dehaene). Contrary to Quartz & Sejnowski's suggestions, learning by selection, when proceeding in multiple stages, is not incompatible with neurodevelopmental evidence. (*a*) Multiple waves of synaptic overproduction followed by selective elimination (bottom curves) add up to a curve for total synaptic density that remains stable over a long period of time, as shown in Fig. 4 of Q&S's article (Huttenlocher 1979). (*b*) Multiple, possibly overlapping phases of blind axonal sprouting followed by selection result in an increasingly refined targetting of connections, thus giving the impression of directed growth.

processing. Although Piaget viewed numerical development as a slow, protracted process, we now know that size- and location-invariant discrimination of number is within the grasp of 2-month-old infants (Starkey & Cooper 1980; Treiber & Wilcox 1984) and even of newborns (Antell & Keating 1983; Bijeljac-Babic et al. 1991). Four-and-one-half-month-olds exhibit elementary set addition and substraction abilities (Koechlin et al., submitted; Wynn 1992), and 6-month-olds show evidence of cross-modal matching of auditory and visual numerosity (Starkey et al. 1983; 1990; see also Moore et al. 1987). Such data seem difficult to explain by constructive learning, for it is hard to see what kind of environmental input available to young infants could teach them object addition or cross-modal numerical correspondence (two objects do not necessarily make two sounds, nor do two sounds necessarily come from two objects). Much of early numerical development, however, can be accounted for by a model in which an innately specified array of numerosity detectors serves as the basis for subsequent learning by selection (Dehaene & Changeux 1993).

Language acquisition during the first year of life also provides a rich set of data that militates against constructive learning and supports a selectionist view of development, most likely with multiple critical periods. From birth, human infants discriminate phonemic contrasts (Bertoncini et al. 1987; Eimas et al. 1971), including those that are not used in their maternal language (Trehub 1976). This capacity has been shown to depend on an early left-hemispheric lateralization for linguistic stimuli (Bertoncini et al. 1989; Dehaene-Lambertz & Dehaene 1994), thus contradicting Q&S's notion of a "largely equipotential cortex." Most important, the first year of life shows a decrease in phoneme discrimination abilities. At 8 to 10 months of age, infants still discriminate non-native phonemic contrasts, but by 10 to 12 months they lose this ability and maintain only native phonemic categories (Werker & Lalonde 1988). This provides direct evidence for a selective, environment-dependent loss of initial abilities, as predicted by selectionist theory. A similar phenomenon occurs for vowels at an even earlier age (six months, Kuhl et al. 1992). It is now suggested that discrimination of language prosody follows a similar developmental pattern (Mehler & Christophe 1995). Newborns discriminate sentences drawn from two different languages (Mehler et al. 1988), including languages that they have never heard (Nazzi et al., in press), whereas two-month-olds are only able to discriminate their maternal language from another language, but not two foreign languages from one another (Mehler et al. 1988), again suggesting a selective loss of speech perception abilities in the course of acquiring a specific language. To be sure, learning plays a major role in language development during the first year of life – but it is a form of learning that capitalizes on an initial repertoire of dedicated processes that are already functional right from birth.

Neurotrophic factors, neuronal selectionism, and neuronal proliferation

T. Elliott[1] and N. R. Shadbolt

Department of Psychology, University of Nottingham, Nottingham, NG7 2RD, United Kingdom. **te@proteus.psyc.nott.ac.uk; nrs@psyc.nott.ac.uk**

Abstract: Quartz & Sejnowski (Q&S) disregard evidence that suggests that their view of dendrites is inadequate and they ignore recent results concerning the role of neurotrophic factors in synaptic remodelling. They misrepresent neuronal selectionism and thus erect a straw-man argument. Finally, the results discussed in section 4.2 require neuronal proliferation, but this does not occur during the period of neuronal development of relevance here.

Selectionist accounts are, for the most part, inadequate theories of neuronal development: much evidence suggests that constructive events are at least as important as regressive events (e.g., Purves 1994). For this reason, we are developing models of the formation

https://doi.org/10.1017/S0140525X97301587 Published online by Cambridge University Press

of ocular dominance columns based on the sprouting and retraction of axonal processes (Elliott et al. 1996; 1997). We therefore welcome any paper that argues for a nonselectionist approach to neuronal development. But we have a number of difficulties with the present paper.

Q&S argue that the computational complexity and slow growth of dendrites compared to axons make them good candidates for a measure of representational complexity. We agree that dendritic plasticity must be taken into account in theories of neuronal development (Elliott et al. 1997), but we do not agree that the case for a dendritic measure is as well established as Q&S seem to believe. First, they specifically overlook presynaptic plasticity. The way that a presynaptic terminal responds to depolarisation is complicated and can be highly nonlinear. For example, paired–pulse facilitation, post-tetanic potentiation, presynaptic facilitation, presynaptic depression, and presynaptic inhibition all represent ways that a presynaptic terminal's response can be modified. Second, ocular dominance plasticity is possible even when cortical cells are pharmacologically inhibited (Hata & Stryker 1994). There is admittedly a paradoxical shift towards the deprived eye, but this could be due to interference with inhibitory circuits. This result suggests that the full computational complexity of dendrites is not necessary for at least certain aspects of neuronal plasticity. Third, the known mechanisms of development, plasticity, and competition at the vertebrate neuromuscular junction are similar to those in other systems (e.g., Grinnell 1995). Since muscle cells do not possess dendrites, the role of dendrites in plasticity in general is unclear. Fourth, while Q&S discuss the activity-dependent plasticity of dendrites, they do not mention that the evidence they cite is consistent with a model of dendritic plasticity in which changes in dendritic morphology occur in response to changes in axonal morphology (Elliott et al. 1997). This model accounts for why dendrites remodel on a longer time-scale than axons, and it is a more economical account than Q&S's model of dendrites as detectors of correlated activity. Fifth, Q&S suggest that the N-methyl-D-aspartate receptor might be involved in regulating dendritic growth. Recent evidence, however, suggests that retrograde neurotrophic factors (NTFs), particularly the neurotrophins, might play a role in dendritic growth (McAllister et al. 1995). This brings us to a more general point.

Of late there has been much excitement about the role of NTFs in neuronal development after the phase of target innervation and neuronal death, since NTFs might mediate competition between afferents during synaptic remodelling (e.g., Thoenen 1995). Given that a characteristic feature of NTFs is their ability to promote neurite outgrowth (e.g., Cohen-Cory & Fraser 1995), it seems odd that Q&S do not even mention them: NTFs are a very intriguing class of molecules for anyone interested in neuronal growth, in particular, in the construction of neuronal circuits. A possible model is that NTFs are released from target cells and taken up by afferents in an activity-dependent manner. It is assumed that elevated levels of NTFs promote afferent sprouting, while depressed levels promote afferent retraction. Furthermore, if the target cells' own dendrites are influenced by the NTFs released from them, then the dendrites will be long and ramified in regions of high innervation density and short and relatively unramified in regions of low innervation density (Elliott et al. 1997).

Another problem with Q&S's target article is that they seriously misrepresent neuronal selectionism. It is not true that selectionist theories of neuronal development assume the existence of prerepresentations: they do assume an initial pattern of connectivity, but it is misleading to call this a prerepresentation since all theories, whether selectionist or nonselectionist, which address neuronal development after target innervation must assume such initial connectivity. Selectionist models take such connectivity and chisel out a final pattern by selectively weakening some connections and strengthening others. This process is analogous to the way that a sculptor takes a block of marble and sculpts, that is, constructs, a statue. Yet Q&S do not accept that neuronal selectionism is constructivist in the ordinary sense of the word just used; we do not have the space to pursue an analysis of Q&S's many and distinct uses of the word "construct" in their target article.

Finally, there is a very noticeable logical gap between the end of section 3 and the beginning of section 4. Indeed, the results discussed in section 4 seem to bear little relation to anything preceding it. Section 4.2 discusses recent results in neural network theory relating to adding units to networks. Yet adding a unit is equivalent to adding a new neuron, not elaborating or growing a dendrite or sprouting a new axonal process or terminal branch. In fact, for the phase of neuronal development of relevance to this discussion, neuronal proliferation does not occur. Q&S must therefore demonstrate that axonal and dendritic plasticity are as powerful, mathematically speaking, as adding new neurons. Thus, the critical step from neuronal constructivism to cognitive constructivism, which is supposed to be the basis of their paper, is in fact absent.

In conclusion, while we are broadly sympathetic towards any paper that advocates taking neuronal growth seriously, we believe that, like most manifestos, the present constructivist manifesto promises far more than it can deliver.

ACKNOWLEDGMENT

TE thanks the Royal Society for the support of a University Research Fellowship.

NOTE

1. Address correspondence to TE at **te@proteus.psyc.nott.ac.uk**

Constraining the brain: The role of developmental psychology in developmental cognitive neuroscience

David Estes and Karen Bartsch

Department of Psychology, University of Wyoming, Laramie, WY 82071-3415. **estes@uwyo.edu**

Abstract: Developmental psychology should play an essential constraining role in developmental cognitive neuroscience. Theories of neural development must account explicitly for the early emergence of knowledge and abilities in infants and young children documented in developmental research. Especially in need of explanation at the neural level is the early emergence of meta-representation.

Recent work in cognitive neuroscience suggests an increasing appreciation for the value of a developmental perspective. There also seems to be increasing recognition that progress in cognitive neuroscience requires multiple levels of description because empirical findings at one level are needed to constrain theorizing at other levels. Just as new findings about the brain should guide and constrain cognitive theories, so should research at the cognitive level guide and constrain theories at the neural level. These are promising trends, but what do they mean in practice? What can a developmental perspective contribute to cognitive neuroscience? More specifically, what is the role of developmental psychology in developmental cognitive neuroscience?

We believe that in the endeavor to understand how mind emerges from brain, developmental psychology has a certain priority. The description of children's developing knowledge and abilities provided by basic research in developmental psychology should not just guide theories of neural development; it should tightly constrain them. If the goal is to understand the neural basis of cognitive development, we have to start with cognitive development.

Failure to adhere to this principle weakens theories of neural development. Quartz & Sejnowski (Q&S), for example, although enthusiastic advocates for a developmental cognitive neuroscience, offer a theory of neural development that would be stronger had they dealt more explicitly with some of the behavioral data from recent cognitive development research. In their effort to

https://doi.org/10.1017/S0140525X97301587 Published online by Cambridge University Press

avoid nativism, Q&S fail to address adequately one of the major factors responsible for the rise of nativistic and strongly domain-specific theories of cognitive development. A primary motivation for such theories is the abundant evidence, amassed by developmental psychologists over the past two decades, that many kinds of knowledge and abilities emerge earlier, and sometimes much earlier, than previously recognized (Hirschfeld & Gelman 1994).

Innatism is often justified by citing the impressive capacities of infants. But just as impressive and as needful of explanation, at both the neural and cognitive levels, are abilities first manifested in early childhood rather than in infancy, especially those that can be characterized as "metarepresentational." To illustrate the kinds of early abilities that theories of neural development ignore at their peril, we point to some recent findings in two areas of research we happen to find close at hand – preschoolers' understanding of imagination and their talk about the mind.

By about 3 years of age, or about as soon as children can actively participate in an experiment, they can demonstrate quite sophisticated conceptual understanding of mental images (Estes et al. 1989). They understand the subjective, immaterial nature of mental images and sharply distinguish between this form of mental representation and physical representations such as photographs. They also recognize that mental images do not necessarily correspond to reality, frequently invoking this principle to explain their responses (e.g., "It's not really real. I was just imagining it.")

Very young children also begin talking about mental states spontaneously in their natural discourse (Bartsch & Wellman 1995). References to thinking and knowing emerge in the talk of 3-year-olds. More remarkably, 3-year-olds spontaneously and explicitly contrast the content of their mental states with external reality, and by 4 years of age they make statements such as the following: "Now she knows that I know. She used to think that I don't know when I really did" (Wellman & Bartsch 1995, p. 338).

Experimental and naturalistic evidence thus demonstrates that preschool children understand that thoughts do not always represent the world veridically, that one's mental world and the physical world are independent in this sense. This has important implications for theories of neural development. The pace of cortical development may indeed be slower than previously recognized, as Q&S claim, but an awful lot must be going on inside the skull to support the rapid emergence of the representational and metarepresentational competence manifested in the behavior of very young children. Brain mechanisms must obviously underlie these early competencies, and, because they involve actively constructing and consciously evaluating counterfactual scenarios, prefrontal cortex is strongly implicated (Knight & Grabowecky 1995).

It may be tempting for developmentalists to embrace neural constructivism uncritically because it endows humans with the capacity for flexible adaptation, in developmental rather than evolutionary time, to changes in the natural and cultural environment. Development is central, in contrast to extreme forms of nativism. But regardless of the inherent appeal of constructivism, the extent to which cognitive development is guided by innate predispositions remains an open question, and any viable theory of neural development must come to terms with the exceedingly rapid pace of cognitive development in infants and young children. The ultimate explanation of cognitive development will most likely involve an integration of aspects of nativism and constructivism. Karmiloff-Smith (1992) has already shown us what such a theory might look like. [See *BBS* multiple book review of Karmiloff-Smith's *Beyond Modularity*, *BBS* 17(4) 1994.]

We share the excitement over the prospect of a vibrant and multi-faceted developmental cognitive neuroscience. Maybe it can move us beyond the current post-Piagetian period of scandalous diversity in developmental theorizing. In this decade of the brain the status of neuroscience is not in question. But in this heady atmosphere, when new technologies are finally opening up the black box and neuroscientists seek to understand not just adult cognition but the neural basis of cognitive development, it is easy to lose sight of the fact that behavior still has priority. Description must precede explanation, as Carey (1990) has pointed out, and providing an accurate behavioral description of cognitive development is a daunting task in itself. Solid progress has been made in this endeavor, however, and old-fashioned research in developmental psychology using only behavioral measures to document the early emergence of specific kinds of knowledge and abilities still presents a kind of imperative to neuroscientists: Here is what children can do. Now tell us how the brain does it.

Radical empiricism is not constructive

Jerome A. Feldman

EECS, University of California, Berkeley and International Computer Science Institute, Berkeley, CA 94704. **jfeldman@icsi.berkeley.edu**

Abstract: The radical empiricist theory of the Quartz & Sejnowski target article would result in a brain that could not act. The attempt to bolster this position with computational arguments is misleading and often just wrong. Fortunately, other efforts are making progress in linking neural and cognitive development.

Quartz & Sejnowski (Q&S) have the ambitious goal of providing a neural explanation of cognitive development. The goal is excellent; cognitive science needs a foundational neural theory of cognition paralleling the molecular theory of matter, the germ theory of disease, and so on. More specifically, the target article promises to exploit recent advances in developmental neurobiology and computational learning theory to eliminate from cognitive science the last vestiges of functionalism (the brain is irrelevant) and nativism (everything important is innate). What it delivers, however, is an extreme empiricism (nothing interesting is genetic) that is equally old, simplistic, and hopeless. This commentary will focus on computational issues and then conclude with some pointers to research that substantively addresses the neural basis for cognitive development.

The task of the target article could be made easier by dropping the arbitrary separation of development from learning. For decades, cognitive development has been viewed as a life-long process. Extending this in the obvious way to the neural level would eliminate any *a priori* distinction between development and learning and seems a necessary step in any effort to unify the fields. It would also be helpful to eliminate the teleological fallacy, the notion that everything in the brain is an optimal solution to some current evolutionary problem.

Q&S's proposal is as radically empiricist as possible, maybe more so. The block diagram in their Figure 1 has a target state, but no initial state – just "innate principles." The manifesto is made explicit in section 3.2. "Cortical areas farther away from the sensory periphery wait in anticipation of increasingly complex patterns of activity resulting from the development in lower areas" – apparently with no initial state and no influence from any functional requirements. A Q&S brain could develop ever more elaborate visual descriptions of a scene but is given no way to make any inferences or actions with the information. A coherent version of the equipotential cortex story could go like this: early development yields a complete, but primitive, brain – possibly all subcortical. Cortical development occurs as follows: each primitive functional module recruits nearby cortex to expand its functionality and the recruited cortical regions learn to communicate with one another by directed growth. There are compelling reasons not to accept this story (Johnson 1996), but at least it provides a basis for computational modeling that is not inherently inadequate.

About a quarter of Q&S's article is about computational learning theory. The argument is that theoretical results show the superiority of data-driven development over selectionist methods in abstract neural networks. There is a lot of excellent work on learning systems that add structure incrementally but it does not support Q&S's contentions. They talk of "impressive" results for constructive methods, but essentially all of the thousands of

https://doi.org/10.1017/S0140525X97301587 Published online by Cambridge University Press

applications of neural networks are exclusively selectionist. Moreover, much of the technical discussion is misleading and some of it is badly mistaken. No two people can be fully informed over the full range covered by the target article, but one could hope that the authors or editors would seek advice.

The paradigm problem addressed is the learning of formal grammars. By ignoring the issue of negative evidence, the target article misses the whole point of Gold's early work and the resulting subdiscipline (Carrasco & Oncina 94). The learning problem that Q&S actually state is easily solved by selecting a grammar that treats any sentence as acceptable. Since children receive little explicit information about ungrammatical examples, the central question is how they reject such over-general grammars. There are non-nativist proposals for this (à la Occam) but the target article does not give us either the question or the answer. It is true that no device with bounded resources can learn all the grammars in a standard class (such as context free), because each class includes languages with arbitrarily large grammars. However, all of the cited constructive learning results apply only to functions with inputs of fixed size and not to grammars, which must accept sentences of any length. The stronger results (e.g., Barron 1994) also require continuous functions, which grammars decidedly are not – a small change in input (say a spelling error) can make a large difference in grammaticality. The constructive approaches that have made progress on grammar learning (cf. Carrasco & Oncina 1994) are not mentioned. Although biological relevance is invoked as a key argument, none of the algorithms cited makes any pretense of biological plausibility.

There is a more basic problem with the computational section. By identifying language acquisition with the learning of formal grammars (which are by definition content-free), Q&S preclude any discussion of conceptual content. In fact, despite the title of the piece, it says literally nothing about human conceptual systems and how they might be neurally represented and learned. Fortunately, there is a good deal of progress on these important issues. One core problem for any neural theory of cognition concerns how mature brains learn to associate new words with images (or other experiences), which they clearly do not do by directed growth. There is a literature on recruitment learning (e.g., Diederich 1990) that addresses this and other core problems of neural cognitive development.

Johnson (1996) presents an overview of human neural development that is sympathetic to empiricist claims but is much more balanced and careful than the target article. The developmental cognitive neuroscience cited in the introduction is outgrowing its *tabula rasa* phase and moving to a more plausible interactionist stance (cf. Elman et al. 1996). There is also a great deal of work that takes seriously how the brain might represent and learn concepts (cf. Shastri & Ajanagadde 1993 and the continuing commentary). For connectionist modeling of early cognitive development and language learning, Regier (1996) is a good place to start.

Despite its many flaws, the target article does have redeeming social value. Connectionist computational models can and should play a significant role in theories of neural development and Q&S's paper can be read as a call for more work in this area. Q&S are right that current approaches will not suffice. But serious research will need to take explicitly into account the initial and final state of each developmental episode as well as the accompanying experience. It is certainly easier to work with equipotential (random or uniform) networks and many people will do so. This is fine as abstract theory, but as a manifesto for linking neural to cognitive development, it is definitely not constructive.

So many problems, so little time: Evolution and the dendrite

Barbara L. Finlay

Department of Psychology, Cornell University, Ithaca, NY 14853.
blf2@cornell.edu

Abstract: The multiple levels of analysis that Quartz & Sejnowski (Q&S) bring to bear on the phenomenon of activity-driven dendritic growth show the tight linkage of explanations from the cellular to the cognitive level. To show how multiple control regimes can intersect at the same site, I further elaborate an example of a developmental problem solved at the axodendritic connection: that of population matching.

Developmental biology, as invented by evolution, acts as if determined to confound standard hypothesis-testing methods. When asked whether Mechanism A or B is used to implement a particular developmental task (for tasks ranging from the production of trophic factors, to the organization of topographic maps, to establishment of the circuitry for syntax in language), the answer increasingly appears to be A and B (and C, and D, etc.). Conversely, if a phenomenon asks for explanation, such as developmental cell death, or in Q&S's case, activity-driven dendritic growth, it is becoming commonplace that what appears to be a single mechanism subserves a number of different functions and exists at a number of organizational levels.

Things become more difficult still when natural selection, having little interest in the form of intellectual bookkeeping we call "levels of analysis," not only multiplies mechanisms but mixes their levels. The formal separability of levels of analysis has been much discussed (Fodor 1992). In the particular case where an organizing force such as evolution acts across "levels," philosophic separation of levels has produced an *apparent* intellectual clarity that has in fact hopelessly muddled understanding of the actual state of affairs. Sejnowski and colleagues are to be highly congratulated for this article, as well as others, which show another of the many ways perception, cognition, development, and physical mechanism overlay: Gold's theorem and rules for dendritic growth do belong in the same article! In this commentary, I would like to point out another multiple-level, multiple-mechanism conjunction in the case of developing dendrites.

In addition to the unusual challenge of representing the informational structure of the outside world, neurons are posed with some physical challenges in early development. They must send axons to vicinities where they have some chance of encountering their proper targets, determine whether that has occurred, and roughly calibrate the numbers of neurons in connecting populations, which depends upon a reciprocal supply of trophic factors as one of several mechanisms. Q&S have discussed the problem of pathway selection and reduction of developmental "errors" in their target article; I would like to discuss briefly the problem of population matching, as it shows some differences in the behavior of axons and dendrites from those highlighted by Q&S. There are several games in progress on the dendritic playing field.

Convergence, the fan-in of presynaptic cells to postsynaptic cells, varies widely in the nervous system. In some cases, for example, the ratio of projecting thalamic axons to isocortical neurons, convergence seems virtually unconstrained and a potential source of plasticity (Finlay & Darlington 1995); in others, as in regions of the retina (Rowe 1991) and in some aspects of cerebellar circuity (Williams & Herrup 1988), convergence seems highly constrained and of immediate functional consequence. The developmental regulation of convergence ratios from retina to tectum in developing mammals (Xiong et al. 1994; reviewed in Xiong & Finlay 1996) and during regeneration (Hayes & Meyer 1988) shows that the action is in the axon. The visual receptive field size of a normal tectal neuron does not show wide variability in the adult, and this receptive field size is defended against a wide range of experimentally induced variations in the total number of retinal neurons compared to tectal neurons (Pallas & Finlay 1989). During development, this problem must be solved in the perplex-

https://doi.org/10.1017/S0140525X97301587 Published online by Cambridge University Press

ing context of early overabundance of neurons. In this case, the numbers of both retinal and tectal neurons are several times in excess of their mature numbers. What matches what? Essentially, tectal neurons proceed to their adult number of neurons through apoptosis and to their adult state of synaptic density uninfluenced by excessive numbers of axons in the input retinal population. Retinal neurons buffer their own rate of survival through their axons: fan-out is sacrificed to fan-in, so that axonal arbors are much reduced in size (up to tenfold) when convergence ratios are high. The number of contacts from any one retinal neuron to a tectal neuron is about the same, but many fewer tectal neurons are contacted by each retinal axon. Conserved target dendritic volume sets the upper and lower limit on retinal neuron number, with plasticity within this range permitted by axonal variability. It is of interest in light of Q&S's article that a functional reset of the size of a system, as might happen in the case of neural structures influenced by gonadal steroids (Sengelaub 1989) or by unusual learning regimes (Turner & Greenough 1985), can and does happen through dendritic structure. More time could be spent on the specific empirical question about the active role of the axon, somewhat neglected in Q&S's discussion, but that is a separate issue.

Overall, Q&S's multilevel style of analysis is to be much applauded, and the example set out here adds an additional regulatory problem for the axo-dendritic interface. At least three problems involve the axo-dendritic interface in early development: target selection; population matching, and the activity-dependent dendritic structuring that was the main focus of Q&S. Each problem uses the information of the correlated firing of the input array, but each under a different control regime. We might construe some problems as more biological in level (e.g., acquiring adequate trophic support), others as more "cognitive" (e.g., representing the correlational structure of the input), but, in all cases, it's the same dendrite solving the problem.

Irresistible environment meets immovable neurons

Jeffrey Foss

Department of Philosophy, University of Victoria, Victoria, BC V8Y 3P4 Canada. **june19@uvvm.uvic.ca**

Abstract: Quartz & Sejnowski's (Q&S's) main accomplishment is the presentation of increasing complexity in the developing brain. Although this cuts a colorful swath through current theories of learning, it leaves the central question untouched: *How* does the environment direct neural structure? In answer, Q&S offer us only Hebb's half-century-old suggestion once again.

Excellent manifesto! Selectionism, long supported by the developing brain's decreasing number of neurons, is chastened by data showing considerable growth during learning of number of synapses (30–50%), axon branching (300%), axon length (70%), dendritic length (35–3000%, see sect. 2.3.1), and consequently a hefty growth (?%) in neural *complexity* (about which more below). The age-old dispute between hardwired innatism and *tabula rasa* empiricism is answered by a plausible plague on both houses: since neural complexity increases during learning, any native hardwiring is later wired over, and sensory experience can hardly write upon a continually metamorphosing tablet – or, if it does, this is not empiricism, but constructivism. Abstract learning theory is forced to face data; Chomsky's legacy takes another hit; and the ever-ready theory of evolution is adapted to Quartz & Sejnowski's support. Constructivism rules.

But, kudos done, we owe the new proto-paradigm some critical analysis. The axial idea at the motionless hub of the spinning constructivist wheel is this: the increasing complexity of the brain is in fact growing representational space structured by the environment. This idea has the important virtue of truth – or verisimilitude – but, just as important, it commits the sin of merely restating the very problem which constructivism is supposed to solve. Surely people do learn from their environment, and surely on any scientific, physicalistic account, learning (except for ephemera) is implemented by structural changes in the brain. Thus everybody, whether selectionist, nativist, empiricist, learning theorist, or whatever, can agree with Q&S that "learning guides brain development in very specific ways" (sect. 4). Very specific ways indeed, just as the ways of all actual things are very specific – but Q&S offer no actual, specific case of *the way* brain development is guided by the environment; no cases, that is, of learning.

Q&S's main accomplishment is persuasive presentation of evidence of the increasing complexity of the learning brain. However, this merely cuts a swath through competing accounts, thereby clearing ground for their own account, which, as portended by this manifesto, would need to show three things: (1) that the additional complexity of the brain is representational, (2) that it is directed by sensory input, and (3) how sensory input does such a thing.

Point 1 is merely asserted and assumed: "the physical structure of a neural area corresponds to a representational space," so "[b]uilding neural circuits with directed growth thereby builds the brain's representational properties" (sect. 3). Well, okay; though we might note that even for manifestly representational entities, such as DNA, increasing complexity might just yield junk, not more representational power. But charity and good sense dictate that we not nip the constructivist rose in the bud. After all, connectionism is plainly running just under Q&S's constructivist surface, and this provides the basis for the assumption in point 1: since representation is distributed through synaptic weights on the neural net, increasing the size and connectedness of the net generally increases its representational power. The crudity of this generalization is refined in section 4.2, where Q&S cite studies which argue that "the effect of *allowing* [my emphasis] a system to add structure – to build representations – as it learns" is that it "can learn essentially any arbitrary mapping." This in itself is a wonderful result (though not Q&S's). However, allowing a system to build representations is just too easy. The question is whether growing neural structures *are* representations – and if so, *what* is building them, and *how*.

Point 2, as noted above, is a broadly shared article of faith. In the happy spirit that faith should seek support in reason, Q&S muster substantial evidence that increasing neural complexity requires sensory input (sect. 2.3.4) – but that it is directed by that input is not so clear. The growing complexity of the embryo requires many things such as nutrients, a sheltering egg or womb, and so forth, but these things do not direct this complexity. We can see that it is directed by DNA and associated nuclear processes just to the extent that we can see how this is done. Likewise until we can begin to see how sensory input directs neural complexification, we cannot see that it does.

So everything rides on point 3, the question of how the learning environment shapes neural development. Q&S answer by references to Hebbian learning (sect. 2.3.5), bolstered by suggestions about a "diffusible substance" (sect. 3.1.1) which might add to its neural realizability. What is conspicuous in its absence is any discussion of convincing victories for Hebbian learning, though a few interesting works are listed. It is fair to suggest, however, that Hebbian successes have been modest. The best examples of learning in connectionist nets depend upon distinctly non-Hebbian, god-like algorithms, which not only know how close a given output is to the target output and which synaptic adjustments to make, but also have the power to make them. This model of learning, however, is grossly unrealistic neurally.

Hebbian learning, by contrast, is neurally realistic because it is local: processes in the locale of the desired, but as yet unformed, synapse can coax it into existence. Hebb's idea, conceived deep in the stimulus-response heyday of behaviorism, is that when the activity of one neuron repeatedly "takes part in firing" a second, the first should synapse onto the second (Hebb 1949, p. 62). Thus, as in Pavlov's dogs, neural assemblies registering ringing bells

https://doi.org/10.1017/S0140525X97301587 Published online by Cambridge University Press

should synapse onto those for salivation, perhaps through a series of adjacent intermediaries should they be too widely separated. Unfortunately, what Hebbian learning gains in locality and realizability, it loses in directedness and speed. When Sylvie's teacher says, "No, no, six times eight is forty-eight, not forty-nine," Sylvie may in the twinkling of an eye shed her error and learn the right answer for good. Neither she, nor her teacher, nor anyone else, has any idea which synaptic connections to weaken or which to strengthen – so how do the neurons know? Neither Hebbian nor neo-quasi-Hebbian learning have much to tell us in answer to this question – and it is not manifest what else constructivism offers.

Apparently the environment is irresistible: learning, hence neural structure, is directed by sensory input. But, apparently, neural patterns are immovable: the environment has no mechanism for pushing neurons into line. The point of a theory of learning is to go beyond these appearances to show us *how* the environment directs neural structure.

Neural models of development and learning

Stephen Grossberg

Department of Cognitive and Neural Systems, Boston University, Boston, MA 02215: **steve@cns.bu.edu; cns-web@bu.edu/profiles/grossberg.html**

Abstract: I agree with Quartz & Sejnowski's (Q&S's) points, which are familiar to many scientists. A number of models with the sought-after properties, however, are overlooked, while models without them are highlighted. I will review nonstationary learning, links between development and learning, locality, stability, learning throughout life, hypothesis testing that models the learner's problem domain, and active dendritic processes.

I introduced adaptive resonance theory, or ART, in 1975–1976. It has since been developed by many scientists. Sejnowski helped to popularize backpropagation and cites Rumelhart et al. (1986) for "placing learning and the creation of internal representations once again in the fore of cognitive science." Backpropagation, however, is a slow off-line, learning algorithm which uses non-local operations that become unstable in nonstationary environments, and cannot function without a teacher. Developing cognitive systems learn on-line, quickly and locally; they are stable in nonstationary environments, and function without teachers. Q&S nonetheless cite backpropagation and related models to illustrate developmental constructivism, but not ART-like models that have developmental/cognitive properties (Carpenter & Grossberg 1991; 1993).

ART models gain these properties by solving the stability-plasticity dilemma of how to learn quickly and stably without suffering "catastrophic interference." Learning top-down expectations helps to achieve this balance and makes sense of why top-down corticocortical and corticothalamic pathways exist. Self-regulating memory search, or hypothesis testing, creates representations that match nonstationary environmental trends. Selectionists like Edelman (1987) posit huge numbers of "pre-representations" and select good ones via competition. Their theories select local minima (or non-optimal representations) because they omit top-down attentional focusing and hypothesis testing. ART realizes these processes using interacting attentional and orienting subsystems whose properties interpret temporal/prefrontal and hippocampal data, respectively (Grossberg & Merrill 1996). Many studies illustrate how synchronizing resonances between bottom-up and top-down signals drive self-stabilizing learning while stimulating cognitive and neural data; for example, Grossberg (1987; 1988; 1995).

ART describes cellular activation and learning laws as well as circuits to control them contextually. Its laws for adult learning are formally similar to developmental laws, the former seamlessly continuing the latter, within a "universal developmental code." Two early articles illustrate this developmental constructivism. Grossberg (1978a) described a neural framework for developmental and cognitive processing that stimulated much subsequent work. This framework emerged from studies during 1972–1978 by Malsburg, Willshaw, and myself that introduced modelling of activity-dependent map development. My discovery that development could become unstable led to ART models of map self-stabilization. Grossberg (1978b) suggested that the laws of neural synaptic development and map learning also control cell growth, movement, and adhesion in both neural and non-neural morphogenetic processes. It was proposed that associative (Hebbian and anti-Hebbian) laws modeling synaptic learning direct the formation of new cell outgrowths and connections, one of Q&S's key points.

Q&S mention the type of Hebbian and anti-Hebbian learning that works well in sensory and cognitive processes as part of match-based learning architectures such as ART. Associative mismatch-based learning occurs in spatial and motor processes (Grossberg 1997). ART-style learning stabilizes sensory and cognitive representations. Mismatch-learning lets spatial and motor representations learn new parameters continuously to control changing bodily parts. Grossberg and Kuperstein (1986) modelled how a cascade of such processes, ranging from superior colliculus (SC) to parietal and prefrontal cortex, develops progressively through life to control visually reactive, multimodal, attentive, and planned eye movements. Grossberg et al. (1996) described how such adaptive circuits simulate SC burst and build up cell profiles in various experimental paradigms.

Selective synaptic learning on dendritic spines can use active dendritic spikes to dissociate read-out of presynaptic signals from read-in of their postsynaptic consequences (Grossberg 1975; Grossberg & Schmajuk 1987). Stability considerations suggested that competitive decisions reorganize the total presynaptic read-out across a cell network before surviving signals use postsynaptic dendritic spikes to read-in new synaptic values. Clustered associations can thereby form, as Q&S note.

Models of developmental learning in a nonstationary world suggest why the brain uses nonlinear feedback processes. These processes require diligence to understand, even though their heuristics are simple. Articles like this may make studying them more fashionable. The neural modelling field has gone through several such manifestos and fashion cycles since 1980. They have typically sold old designs to new markets and renamed the label. I hope this won't happen again here. [See also Grossberg: "Neural Models of Reaching" *BBS* 20(2) 1997.]

Processing limitations can help neural growth build hierarchical representations

Gary Haith

Department of Psychology, Stanford University, Stanford, CA 94305. **haith@psych.stanford.edu; white.stanford.edu/~haith**

Abstract: Processing limitations can be an advantage for a learner (Elman 1993; Newport 1990; Turkewitz 1982). They can filter input to the learner so that the relations to be learned increase in complexity only gradually. The time-course of filtered input can complement the growing neural representations discussed by Quartz & Sejnowski.

Quartz & Sejnowski (Q&S) have presented an impressive and powerful framework for understanding development. They aptly demonstrate that developmental neuroscience can and should inform research in cognitive development. In suggesting how mechanisms of neural development potentiate hierarchical representations, the authors outline a clear alternative to some strong nativist assumptions. On this point the target article meshes with some like-minded theoretical and connectionist work in the psychology literature.

The work of Elman (1991; 1993), Newport (1990; Goldowsky & Newport 1993) and Turkewitz & Kenny (1982) addresses the role

https://doi.org/10.1017/S0140525X97301587 Published online by Cambridge University Press

of information processing limitations that change over the course of learning. Contrary to expectation, they suggest that limitations can actually work to the learner's advantage, essentially by helping the learner to decompose complex tasks into simpler ones. In their view, processing limitations can act as filters that "select" input appropriate to the learner's level of expertise. Moreover, if a domain is hierarchical, the learner can extract useful components of the more complex complete task from this selected input. Both Elman (1993) and Goldowsky and Newport (1993) have demonstrated that neural network models can be made more effective in artificial language tasks by *limiting* their capacities in realistic ways.

An illustration might help clarify the basis of this surprising finding. A prelinguistic child with adult memory and attention capacities (a stationary learner) would be faced with long strings of phonemes that could be parsed and related to each other and any potential referents in a staggering number of ways. This profusion of possible relations is discussed by Q&S as an intractably large hypothesis space (sect. 4.2), that remains large due to the "poverty of the stimulus" (sect. 4.2.1, para. 3). In contrast, an actual child with limited memory and attention can only consider the relations between those phonemes that fall within its limited span of memory and attention. These limitations focus the child on relations at the phrase level, a more approachable problem, with a much smaller hypothesis space. In addition, understanding the structure of simple phrases helps the learner understand more complex sentences.

This work provides a psychological mechanism by which hierarchical representations might be built. Processing limitations initially select for the child the most basic, tractable components of the task domain. Once these components are mastered, the child can apply them to understanding the increasingly large portions of the domain to which it is exposed as the limitations abate. Because of this process, the child might be expected to build complex representations using simpler ones that were mastered earlier. This prediction adds a psychological dimension to the neural details provided by Q&S. It suggests that the growing neural structures described by Q&S might first represent simple components and then progressively combine and interlink these building blocks to build more complex representations.

It is important to note that the processing limitations are manifestations of neural constraints just as representational limitations are. Thus, neural development establishes the timecourse of the structure in the input that reaches the brain and the structure that can be represented by the brain. In this way, the representational capacity of the brain could be tightly coupled to the complexity of the input that impinges upon it, with both increasing gradually over neural development. Q&S highlight the "dynamic interaction between a structured environment and the neural mechanisms that are responsive to that structure" (sect. 4.3, last para.). The addition of gradually abating processing limitations to their framework allows the neural mechanisms not only to respond to the structure in the environment but actually to shape and amplify the environmental structure to which it will react. Though the flexibility of filtering is limited, it probably contributes to the development of complex and powerful neural representations.

Quartz & Sejnowski's target article is part of an exciting trend in studying development embracing the complexity of a dynamically changing system that shapes and is shaped by its environment. It seems that the time is ripe for neuroscience and psychology to enrich each other's understanding of development.

Evolution might select constructivism

James Hurford, Sam Joseph, Simon Kirby, and Alastair Reid

Department of Linguistics, University of Edinburgh, Edinburgh EH8 9LL, United Kingdom. **jim@ling.ed.ac.uk** **www.ling.ed.ac.uk/~jim**

Abstract: There is evidence for increase, followed by decline, in synaptic numbers during development. Dendrites do not function in isolation. A constructive neuronal process may underpin a selectionist cognitive process. The environment shapes both ontogeny and phylogeny. Phylogenetic natural selection and neural selection are compatible. Natural selection can yield both constructivist and selectionist solution to adaptuive problems.

Ontogeny. In order to promote their hypothesis of neural constructivism and demote the process of neural selectionism Quartz & Sejnowski (Q&S) make several unjustified attacks on the role of the synapse in development. To begin with, they point out that the studies of pruning in human prefrontal cortex (sect. 2.1.1 para. 5) are inadequate and do not have data covering ages 1–5 and 8–15. It is unfair of Q&S, however, to demand human data in this instance when almost all of their other references to neuroanatomical studies rely on nonhuman data. A more important point, which they avoid making, is that studies (including Rakic et al. 1994, whom they cite) of other brain regions show a definite increase, and then decline in synaptic numbers, with different regions having different time courses. Since cognitive processes are not restricted to the prefrontal cortex and develop at different rates, the aforementioned study contradicts the claim that synapse elimination does not underlie cognitive development.

Another problem arises with Q&S's argument that it is the dendritic arbor which provides the necessary measure of representational complexity. Dendrites do not function in isolation; the representational power of a neuron must be a composite of dendrite and synapse. Thus factors affecting synapse formation and location are just as important as those affecting dendritic growth. Following from this is a more general point that the addition of structure does not necessarily change the one-to-one connectivity that may have arisen from pruning. For example, in the neuromuscular junction, axons show an increase in branching and synapse number but this certainly does not represent an increase in representational complexity (J. Lichtman, personal communication). A further illustration is seen in the cerebellum where a *single* climbing fibre connects with a *single* Purkinje cell despite the complicated arborisations of both (Purves & Lichtman 1985). Our main point here is that increase in structural complexity (e.g., dendritic arborisation) does not necessarily give rise to an increase in representational complexity.

Finally there is the issue of levels of processing. Q&S wish to discuss the ways in which "cognitive and neural processes interact during development." It is reasonable to say that the acquisition of cognitive processes and their performance are inseparable components of cognitive development and that "learning guides brain development." This does not mean that learning and brain development proceed by the same process. It is perfectly feasible that a constructive process at the neuronal level (e.g., extension of a dendritic arbor) may underpin a selectionist process at the cognitive level (e.g., hypothesis elimination; Levine 1966). Thus, although neural constructivism and constructive learning are both valid concepts, neither one entails the other. The interaction between neural and cognitive processes in development is unlikely to be so simple as to be described by a single concept, be it constructivism or selectionism.

Phylogeny. The target article is negative about the prospects of understanding the human brain through understanding its evolution. Q&S's only positive statement is that there is a direction to evolution, that the "hallmark of cortical evolution is . . . an increasing representational flexibility." Q&S do not actually provide any evidence for this view; they simply argue that the only alternative is

https://doi.org/10.1017/S0140525X97301587 Published online by Cambridge University Press

to see the brain as a "hodgepodge of specialized circuits," which is not in accord with the neurobiological evidence.

Although Q&S's view places prime importance on the structure of the environment in ontogeny, it demotes its importance in phylogeny. We believe this is an impoverished position. Instead, evolutionary psychologists (and evolutionary linguists) should be encouraged to take up Q&S's timely challenge to be responsive to neurobiological constraints, and explore the way in which the structure of the environment could influence the evolution of constructivist learning itself. Q&S do suggest that there are still some initial biases "that are consistent with a largely equipotential cortex," but it is crucial to relate these biases with the environment in which humans evolved.

It is likely that the constructivist approach to learning in particular problem domains is a particular evolutionary response to those problems. We need to understand which features of the environment and the problems the environment poses lead to the evolution of a constructivist strategy. For example, the "long period of dendritic development in Broca's area" is an evolutionary response to the pressure to learn language. As Q&S themselves point out, this constructivist response is bought at the cost of "increased vulnerability during a protracted developmental period."

This leads to a view of domain specific innateness in which development plays a central role. What is coded for genetically is a maturational program which leads to certain areas of the developing brain being particularly sensitive (in a constructivist sense) to input. We can now talk about a new kind of neural "specialisation": a specialisation or dovetailing of developmental strategies to the environment in which organisms evolve. Extended cortical dendritic growth will dovetail with problem domains that require flexible representation construction. This will not be true for all problem domains, in which case it should be unsurprising that this extended development is not uniform over all cortical areas. Kirby and Hurford (1997) show that this type of evolutionary adaptation of developmental programs can be modelled computationally. For language learning they predict an early phase of "constructivist" development leading to a critical period for acquisition.

Topics for evolutionary research in the light of Q&S's paper could include an attempt to understand the relationship between constructivist learning and the Baldwin Effect (Baldwin 1896), in which flexible learning should be selected for when the environment is unpredictable but should be selected against when it is not, *whenever increased plasticity has a cost*. The effect predicts that learning can guide the evolution of innate solutions (see, e.g., Hinton & Nowlan 1987) by increasing the phenotypic plasticity of organisms and broadening the search space of a particular genome. Without further understanding, however, of the parameters along which constructivist development can vary (Q&S list: "generic initial cortical circuitry, conduction velocities, subcortical organization, learning rates, and hierarchical development"), we cannot model exactly how learning and evolution will interact.

Finally, Q&S state that "[neural] selectionism . . . is strictly incompatible with the claim that evolutionary pressures have picked out specialized circuits." Although those advocating neural selectionism may emphasise its epigenetic component, it seems clear that it is not incompatible with natural selection. Moreover, natural selection could pick out the initial set of neural circuits (or even the mechanism for their generation) over phylogenetic time, which would then be subject to selectionist processes in the ontogenesis of the individual.

In summary, we are left with a view that the environment can impinge on the cortex on both the ontogenetic timescale (through several possible developmental processes at both a neural and cognitive level) and a phylogenetic timescale (through the selection of developmental processes which are appropriate to the environment). The important contribution of the target article is in putting the structure of the environment and the dynamics of development back on centre stage.

ACKNOWLEDGMENT

The authors would like to acknowledge the members of the Adaptive Systems Workshop at the University of Edinburgh, whose discussion of the target article clarified many of the issues raised. Corresponding author: JH.

"Differentiationism" can reconcile selectionism and constructivism

G. M. Innocenti

Institut de Biologie Cellulaire et de Morphologie, 1005 Lausanne, Switzerland. **giorgio.innocenti@ibcm.unil.ch**

Abstract: Increased complexity of representations in development probably results from the differentiation of distributed neural circuits. Axonal differentiation plays a crucial role in this process. Axonal differentiation appears to be achieved in stages, each involving combinations of constructive and regressive events controlled by cell intrinsic and cell extrinsic information.

Quartz & Sejnowski (Q&S) emphasize the notion that, beyond the empire of the genes, development and learning interact with each other. This is a lovely idea, developed with "mucho gusto" and coming at the right time. But that is not all. They address other fundamental issues in brain sciences, including the relation between neural structure and representations (sect. 2), the nature of the processes underlying the development of neural structure and the role of intrinsic versus extrinsic information in these processes (sect. 3). In this theory-laden journey through the neurosciences, they lost me in a few places. I shall speak to some of them.

Some measure of neural structure, possibly dendritic complexity, ought to correlate with representational complexity. Why dendrites? And which dendrites? I was totally lost here. Most of the arguments advanced to support the role of dendrites could be extended to axons, when the complexity of their terminal arbors is considered, along with their possible computational properties (Innocenti 1995). Furthermore, there may be few, if any, crucial differences in the way dendritic and axonal arbors develop. Both differentiate through (often) exuberant growth and regression, show directed growth, respond to environmental conditions and can undergo local changes over extended life spans, including adulthood. Neither of them seems to call the shots in development; rather, they adapt to each other. Indeed, representations might not relate to any neuronal compartment in isolation, but rather to distributed neural circuits where axons, dendrites, and synapses combine interactively with each other, and with structural parameters, including the number of neurons and biochemical determinants of synaptic transmission. If increased complexity of representations means more numerous representations, each involving more elements, then it might be related to the number of *different* neural circuits available, to the number of *different* neurons belonging to each of these circuits, and perhaps to some measure of their connectiveness. As Edelman (1987) stressed the notion of "different" is important; I will come back to it. For now, there can be no doubt that:

Regressive events do play a crucial role in the development of neural circuits. On this point, the target article is biased in its presentation and quotation of the literature. More complete accounts of regressive events in the formation of neural circuits can be found (e.g., Clarke 1994; Innocenti 1991; O'Leary 1992). In short, development involves massive reduction in the number of neurons, dendrites, dendritic spines, axonal branches, and synapses and, in addition, elimination of receptors, of neurotransmitters, and so on. There are some methodological difficulties here. One has to do with the quantification of neural structures. As rightly noticed by Q&S, the assessment of synaptic densities in a changing volume does not provide results unequivocally interpretable in terms of connectivity, and sometimes not even in terms of total numbers. In general, regression of synapses (or of other

https://doi.org/10.1017/S0140525X97301587 Published online by Cambridge University Press

neuronal structures) is difficult to rule out since it could take place without decrease (or even with an increase) in their total number. Q&S call these situations "turn over," but this term should be reserved for steady-state conditions in which elimination and regrowth leave the functional structure of the circuit unchanged. This is not what happens in development.

Does any of the above imply that we should embrace a "strict selectionistic position"? Not to me! For I cannot believe that the developing brain has all the circuits of the adult brain plus many others. Nor have I ever met an infant who was endowed with all the representations of an adult, plus more. (If I become an infant again, I'll check.) In neural development, as in immunology and in evolutionary biology, selection is only the initial step, which would have minimal consequences were it not followed by the amplification of the phenotypes surviving the selection. Selection and construction are, therefore, strictly intertwined, even in areas to which selection applies unchallenged.

But do regressive events in development imply selection? I suspect so, in most cases. Indeed the criteria according to which the selection operates are sometimes known and the selection can be modulated. Scientists compete for grants and their neurons for food. Changes in the availability of trophic factors alter the number of neurons (axons or synapses) surviving at a given stage of development. A different type of regression is that of neurons and neural circuits which, rather like toys, are eliminated at some point, after having taken part in the juvenile behavior. The latter examples are not restricted to the metamorphic development of insects (Iriki et al. 1988). However, we do not know how widespread they are in the developing mammalian brain.

Exuberant development and selection are probably a way of coping with the lack of the detailed information required for certain developmental choices, as has been proposed several times (Changeux & Danchin 1976; Innocenti 1991, among others). It should nevertheless be stressed that the selection in development differs from Darwinian selection in some fundamental ways. Darwinian selection operates on differentiated phenotypes, mutated by genetic variation. It decreases phenotypic variation. Developmental selection operates on differentiating phenotypes, at various stages of the differentiation process. At each stage it increases phenotypic variation.

Our experience is with visual callosal axons, probably the only cortical axon so far studied quantitatively from the phases of elongation to that of synaptic elimination and into adulthood (reviewed in Innocenti & Tettoni 1997). These axons have well-defined adult phenotypes, well suited to implementing several computational operations. The phenotypes are acquired in stages. The stages (elongation, subcortical branching, ingrowth and branching in the gray matter, synaptogenesis, synaptic elimination) are probably the expression of cell autonomous programs but they need to be triggered by external signals. At each stage, exuberant growth occurs, followed by regression, both probably controlled by external signals. The overall process is one of cellular "conditional differentiation," leading to the differentiation of neural circuits.

Personally, I would support a differentiationist rather than a constructivist manifesto. I mean that what is important in development is that neurons and the circuits to which they belong become different from one another. Differentiation uses intrinsic and extrinsic information in imaginative combinations of construction and regression. This principle has the advantage of applying to all cellular and system development, not only to the nervous system, and the disadvantage of not being particularly new, although it has been somewhat neglected in developmental neurobiology.

Constraints on the construction of cognition

Mark H. Johnson,[a] Liz Bates,[b] Jeff Elman,[b], Annette Karmiloff-Smith,[a] and Kim Plunkett[c]

[a]*MRC Cognitive Development Unit, London WC1H OBT, England.* **mark@cdu.ucl.ac.uk www.cdu.ucl.ac.uk/mark.home.htm;** [b]*Cognitive Science, University of California, San Diego, La Jolla, CA 92093.* [c]*Department of Experimental Psychology, University of Oxford, Oxford OX1 3UD, England.*

Abstract: We add to the constructivist approach of Quartz & Sejnowski (Q&S) by outlining a specific classification of sources of constraint on the emergence of representations from Elman et al. (1996). We suggest that it is important to consider behavioral constructivism in addition to neural constructivism.

In our view Q&S provide an excellent manifesto for a constructivist approach to neurocognitive development. The approach they outline has much in common with that advanced in our recent book (Elman et al. 1996). Like Q&S, we believe in the importance of constraining accounts of representational change in cognitive development with evidence from developmental neurobiology and neuropsychology, and through considerations from computational (connectionist) modeling. We also agree with Q&S that recent evidence from developmental neurobiology makes nativist or "evolutionary psychological" views of cognitive development distinctly implausible. Although there is substantive overlap in our perspectives, in this commentary we focus on the unique aspects of each of the approaches with the view that, when taken together, they provide a more complete constructivist account of neurocognitive development.

Q&S put forward a specific account of directed dendritic growth, which they suggest is the primary mechanism responsible for the emergence and transformation of representations in cerebral cortex during ontogeny. This mechanism, when taken together with selective loss, provides a powerful basis for representational plasticity, and we believe it significantly extends the plausibility of the constructivist approach. There is a source of evidence discussed in our book, however, that poses an apparent challenge to constructivist views based simply on powerful learning mechanisms: namely, the relative consistency of the outcome of cortical specialisation. In other words, most (but not all) of the time adults tend to have (1) similar types of representations that emerge in (2) approximately similar regions of cortex. In our book we address this apparent paradox by identifying sources of constraint on the emergence of representations in cortex. Without acknowledging and identifying these sources of constraint, neural constructivism is in danger of becoming a sophisticated form of empiricism. Indeed, Q&S acknowledge this:

> 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. (sect. 4.3.1, para. 2)

This is what we attempt in Elman et al. (1996). We also consider that understanding these sources of constraint is essential; one objective of our book is to classify them as shown in Table 1 (Table 1.3 in Elman et al. 1996). By comparison with artificial neural networks, we classified the intrinsic sources of constraint on the emergence of representation in real neural tissue. Like Q&S, we dismiss the possibility of prespecified (innate) representations in cerebral cortex (though not necessarily in subcortical structures) instantiated as detailed microcircuitry including synaptic and dendritic connectivity. We then identified three levels of Architectural constraints referring to the properties of units/neurons, their overall connectivity within a region, and the connectivity between regions. Our review of the evidence suggests that these aspects of structure are not malleable in response to input in the same way

https://doi.org/10.1017/S0140525X97301587 Published online by Cambridge University Press

Table 1 (Johnson et al.). *Sources of prespecification*

Source of constraint		Examples in brains	Examples in networks
Representations		synapses; specific microcircuitry	weights on connections
	unit	cytoarchitecture (neuron types); firing thresholds; transmitter types; heterosynaptic depression; learning rules (e.g., LTP)	activation function; learning algorithm; temperature; momentum; learning rate
Architectures	local	number of layers; packing density; recurrence; basic (recurring) cortical circuitry	network type (e.g., recurrent, feedforward); number of layers; number of units in layers
	global	connections between brain regions; location of sensory and motor afferents/efferents	expert networks; separate input/output channels
Timing		number of cell divisions during neurogenesis; spatio-temporal waves of synaptic growth and pruning/decay; temporal development of sensory systems	incremental presentation of data; cell division in growing networks; intrinsic changes resulting from node saturation; adaptive learning rates

From Elman et al. 1996.

that more detailed aspects of cortical circuitry are, and that they constrain considerably the types of representations that eventually emerge during development. This, then, is one limitation on plasticity. Another source of constraint, also briefly alluded to by Q&S, is timing. In our book we give a number of examples from real and artificial neural networks in which the relative timing of events critically influences the solutions arrived at. Likewise, Q&S stress need for prolonged gradual exposure and regulated rates of growth of cortical (dendritic) networks.

In addition to these intrinsic constraints on the formation of representations, we suggest that initial biases in the behavior of the infant select particular aspects of the environment for input, thus shaping subsequent brain development. For example, the tendency of newborns to orient to faces biases their input from the environment to that class of stimulus (Johnson & Morton 1991). Thus, we argue for behavioral constructivism in addition to neural constructivism. Of course, these biases in behavior have a neural basis which, we suggest, involves mainly subcortical structures. Specific biases in subcortical structures may provide an additional constraint on the representations that subsequently emerge in cerebral cortex. In short, Q&S's neural constructivism needs to be matched with a mechanism for behavioral constructivism in the infant.

While we are encouraged that our book and Q&S's target article offer a vision of a future constructivist approach to neurocognitive development, there is clearly a long and challenging path ahead. Specifically, there is a need for computational models which incorporate both dynamic and static aspects of brain structure during development, and which speak to the issue of the formation and transformation of hierarchical representations during postnatal cortical development. Only when we have such models is there the possibility of accounting for interesting cognitive changes.

Cortical development: A progressive and selective mesh, with or without constructivism

Henry Kennedy and Colette Dehay

INSERM U371 Cerveau et Vision, 69675 Bron, France.
kennedy@lyon151.inserm.fr

Abstract: A credible account of the neurobiology underlying cognitive development cannot afford to ignore the recently demonstrated innate regionalisation of the neocortex as well as the ontogeny of corticocortical connections. Although progressive processes are implicated in both phenomena, only for the latter does the timing of development permit control by external events and this is most likely to occur at later stages in the fine tuning of cortical microcircuitry.

Relating mind and brain is laudable. Relating the emergence of mind and brain is positively heroic. Quartz & Sejnowski's (Q&S's) constructivist manifesto attempts just that and takes "environment" as the common denominator. Minds provide representations (of the environment) and environmental factors shape gene expression, which in turn controls development. Hence the philosophers dream: what is to be learnt instructs the means of learning. The notion that environmental factors provide an impetus for progressive as well as selective processes makes for a welcome conceptual change from the excessive reference over the past few years to so called "regressive phenomena" (cell death, axon and synapse elimination). However, a major confusion in the manifesto is made concerning the the environment of minds (which is the object of learning and what we refer to as external events) and the environment of cells (which in a multicellular organism is largely specified by the genome and referred to as the "internal environment"). The constructivist manifesto requires that external events control growth mechanisms. However, much of cortical development occurs in the absence of evoked activity signaling external events, and, although spontaneous activity is frequently reported to play a potentially important role (e.g., McAllister et al. 1996; reviewed in Goodman & Shatz 1993), it is part of the internal environment as defined above.

The development of thalamocortical connectivity shows a progressive acquisition of structure rather than the massive pruning required by selectionist theories; given what we know about the development of ocular dominance columns, Q&S are right to mention it. However, why have they chosen to ignore the literature on the development of connectivity linking cortical areas? First, such connections underlie much of the processing which is *inherently* cortical (synapses driven by the thalamus represent less than 1.6% of total synapses in the cortex; Peters & Payne 1993). Second, unlike in thalamocortical connectivity, where the consensus overwhelmingly favours a limited role of pruning, it has been repeatedly claimed that the adult pattern of corticocortical connections develops via the elimination of widespread connections.

It is surprising that so few studies have examined the development of the functionally defined pathways linking cortical areas. In kittens, several studies did show late elimination of connections between visual cortical areas (e.g., Price & Blakemore 1985). However, progressive phenomena have also been reported. As

https://doi.org/10.1017/S0140525X97301587 Published online by Cambridge University Press

Q&S point out, a cardinal feature of representation resides in the topography of connections. A recent quantitative study of the topography of extrinsic and intrinsic cortical connections in cats showed that the divergence values of the bulk of connections in developing pathways were within the adult range, suggesting that this feature of connectivity develops largely independently of selective processes (Kennedy et al. 1994). These studies in the kitten fail to distinguish feedforward and feedback connections. This difficulty can be overcome in the exquisitely compartimentalised extrastriate visual cortex of the primate. Using this model we examined the development of functional feedforward pathways and showed that directed growth plays a major role; the adult pattern of connectivity is accordingly acquired early in development with little or no elimination of inappropriate target axons (Barone et al. 1996).

So far, the development of association feedforward pathways could well be in accord with the constructivist manifesto, although it needs to be stated that even the small amount of developmental pruning observed could reflect important selective processes. However, feedback connections do not show the early specification of the feedforward pathways. Here selection leads to a massive reorganization of the laminar distribution of corticocortical neurons participating in feedback projections (Barone et al. 1995), suggesting that the relative role of progressive and selective mechanisms differs according to the functional role of the cortical pathway (Singer 1995). These findings suggest that understanding the functional role of feedback and feedforward connections gives a better understanding of where one could expect constructivist mechanisms to operate.

The manifesto gives short shrift to the early specification of cortical areas and declares: "the cortex . . . is largely equipotential at early stages" (sect. 4.1.1, para. 2). This is curious because it suggests a selective stance: if the cortex were equipotential, it would have a uniform connectivity so that selective axon loss must be proposed for the emergence of specific regional connections. In any case, the hypothesis of the equipotential cortex is no longer viable in view of recent results with molecular markers showing early regionalization prior to thalamic innervation (Arimatsu et al. 1992; Barbe & Levitt 1991; Cohen-Tannoudji et al. 1994) and axonal tracers showing adult-like distributions of immature corticospinal projections (Meissirel et al. 1993; Oudega et al. 1994). Finally, the concept of an equipotential immature cortex has been definitively refuted by the quantitative analysis of grafting experiments (Ebrahimi-Gaillard & Roger 1996).

The difficult task which remains is to conceptualise the afferent specification by the internal environment (Killackey 1990; O'Leary 1989) with early cortical regionalization (Rakic 1982). Attempts to do this have largely referred to prenatal enucleation experiments in the monkey (Kennedy & Dehay 1993a; Rakic 1988). These experiments show that the dimensions of striate cortex are determined by the number of thalamic afferents and that in their absence, cortex which was destined to acquire striate features instead takes on the phenotypic features of extrastriate cortex, which, however, still continues to display a number of unusual features, possibly of striate origin (Dehay et al. 1996a; 1996b; Rakic et al. 1991). Clearly, the early regionalisation of cortex places important constraints on the amount of instruction that can be derived either from external events or the internal environment.

In conclusion, there are many progressive phenomena in cortical development, particularly in feedforward connections. However, for this to correspond to a constructivist model, external events need to influence the underlying growth processes. The paradigm for such a phenomenon is curiously suggested by numerous reports of cortical plasticity in the adult (reviewed in Kaas 1995), largely unmentioned by the manifesto. That field of research reminds us that the cortex is a universal learning machine and raises renewed interest in the continuity hypothesis (Pinker 1984). It links adult and developmental plasticity (Cramer & Sur 1995) and suggests a neotenous phenomenon as a central feature of cortical function. More specifically, the manifesto needs to pay particular attention to the timing of cortical development and to recognise the improbability of external events influencing *early* development (Kennedy & Dehay 1993b). Because much of cortical pathway formation occurs prior to activity evoked by the external world reaching the cortex, it seems likely that constructivist control concerns the late fine tuning of the feedback cortical conections required for cortical function.

From neural constructivism to children's cognitive development: Bridging the gap

Denis Mareschal[a] and Thomas R. Shultz[b]

[a]*Department of Psychology, Washington Singer Laboratories, Exeter University, Exeter, EX4 4QG, United Kingdom.* **d.mareschal@exeter.ac.uk www.ex.ac.uk/psychology/staff/dmaresch.htm;** [b]*Department of Psychology, McGill University, Montreal, Quebec, H3A 1B1 Canada.* **shultz@psych.mcgill.ca; www.psych.mcgill.ca/labs/lnsc/html/lab-home.html**

Abstract: Missing from Quartz & Sejnowski's (Q&S's) unique and valuable effort to relate cognitive development to neural constructivism is an examination of the global emergent properties of adding new neural circuits. Such emergent properties can be studied with computational models. Modeling with generative connectionist networks shows that synaptogenic mechanisms can account for progressive increases in children's representational power.

Quartz & Sejnowski describe an exciting new approach to relating brain development and cognitive development. Although constructivist accounts of cognitive development have a long history (e.g., Case 1985; Fischer 1980; Piaget 1970), there have been few attempts to ground the principles of constructivist development in a neural or biological substrate. The few attempts have either failed to relate development to specific brain processes or have failed to show how neural changes increase representational power. For example, Piaget (1980) describes cognitive development in terms of biologically plausible competitive mechanisms but fails to specify how these mechanisms are implemented in the brain. In contrast, others – such as Case (1992), who attributes much of cognitive development to greater mylenation in the frontal lobes, or Fischer (1994), who argues that patterns of cyclical brain activity correspond to stage like transitions in brain reorganization – fail to specify how these neural processes increase representational power. Q&S are to be commended for tackling both sides of the issue.

However, an important element is missing from Q&S's argument if they wish to support the hypothesis that the neural mechanisms they describe actually underlie the development of behaviors observed in children. They suggest a means by which constructivist development can occur at the neural level but it is possible that other, more global processes dominate children's cognitive development. Indeed, the global interactions that emerge from a complex system may overshadow the effects of constructivist development at the level of individual neurons. To substantiate the claim that a constructivist neural mechanism empowers a system to learn complex relationships in the same way children do, it is necessary to explore whether the developmental profiles of a generative neural network match the developmental profiles observed in real children when presented with identical learning tasks.

One way to test this is through computer modeling. Computer models of child development provide a way to investigate the mechanisms that may underlie observed behaviors (Halford & Simon 1995). A mechanism that does not produce the correct developmental profile can be ruled out as a candidate. Generative connectionist networks (networks that grow their own architectures as part of learning) are being explored as models of constructivist cognitive development (Mareschal & Shultz 1996). Models

https://doi.org/10.1017/S0140525X97301587 Published online by Cambridge University Press

of children's performance on the balance-scale task, the seriation task, the integration of time, distance, and velocity information, a pronoun reference acquisition task, and a causal judgment task concerning the potency and resistance of an effect have all been developed using the cascade-correlation learning algorithm (see Shultz et al. 1995 for a review). In many cases, the generative connectionist models capture the developmental profiles of children better than do static feed-forward connectionist networks, which often miss important stages (Buckingham & Shultz 1996; Mareschal & Shultz 1996; Shultz et al. 1994).

Generative connectionist algorithms like cascade-correlation allow for both quantitative changes (through connection weight adjustments) and qualitative restructuring (through recruitment of hidden units). The former permits knowledge acquisition within an existing representational framework, and the latter an increase in representational power. The question of whether development is driven by quantitative or qualitative change is central to cognitive development (Keil 1990). The cascade-correlation models suggest that both types of mechanisms are necessary to account for development over a variety of domains, although particular developments may be attributed to one or the other. In addition to conventional quantitative adjustments of synaptic strength, the neural mechanisms reviewed by Q&S allow for qualitative restructuring, thus putting this fundamental psychological distinction on a firm neurological basis.

The cascade-correlation algorithm forces a network to develop initially through a phase of limited representational power and then through successive phases of increasing representational power. That is, the system must develop limited representations and then build on them when added power is developed. The simulations suggest that this is a necessary feature for capturing realistic cognitive development in many domains.

The renewed research interest in the periodic need to increase the representational power of a system (as illustrated in both the target article and the modeling work) suggests that research in cognitive development should move away from the search for ever more precocious proto-representations in children and return to the study of what children can and cannot assimilate at particular stages of development. Constraints on early representations are provided by what the child cannot represent as well as by what the child can represent. Finally, these constraints should be yoked to neurophysiological constraints like those described by Quartz & Sejnowski.

More mathematics: Bodily-kinaesthetic intelligence

Gin McCollum

R. S. Dow Neurological Sciences Institute, 1120 NW 20th Avenue, Portland, OR 97209. **mccollum@ohsu.edu; www.ohsu.edu/s■m-neuroscience/sys/ mccollum.html**

Abstract: Although the idea that cognitive structure changes as we learn is welcome, a variety of mathematical structures are needed to model the neural and cognitive processes involved. A specific example of bodily-kinaesthetic intelligence is given, building on a formalism given elsewhere. As the structure of cognition changes, previous learning can become tacit, adding to the complexity of cognition and its modeling.

Quartz & Sejnowski's (Q&S's) philosophy that we build on our learning is liberating, as is the less strongly stated point that our behavior is not genetically determined (sect. 4.4) and that minds are not computers (sect. 1). A further liberation is needed from narrow and fashionable views of representation and from dependence on artificial neural network modeling. "Selectionism brings increased tractability to modeling efforts" (sect. 2.2.1) only in a narrow range of modeling types.

It is not clear that activity in the cerebral cortex is "representing" anything; it is more accurate to speak of neural ensembles as "transforming" activity. The question is then to characterize the transformation mathematically. That is, a better characterization is needed of the mathematical structure(s) of the cerebral cortex itself. While it is possible that "representational complexity" depends monotonically on synaptic numbers, axonal arborization, or dendritic development, there is no necessary relationship. One can use more words to say less. The complexity of function depends instead on the complexity of patterns among the synapses, axons, and dendrites and on their particular relationship to the neural and cognitive environment.

It follows from Q&S's views that learning is not only meeting some task criterion but also elaborating a set of behavioral (or imaging) strategies. Learning is creation, not limitation. In order to make sense of the role of the cerebral cortex in learning, it is necessary to formalize the interactions of cortical ensembles with other parts of the nervous system and with the environment. An example of a unified mathematical characterization of a simple combined neural-behavioral system is given by Roberts and McCollum (1996).

The cortex's environment in the nervous system is enough to vary the type of mathematics needed. In our studies of movement-related parts of the nervous system, different mathematical systems fit different neural ensembles (Holly & McCollum 1997; McCollum 1992; 1993). Combined mathematical characterizations of cortical and extra-cortical systems will therefore differ, depending on the parts of the nervous system involved.

A good area to model is bodily-kinaesthetic intelligence, to use Gardner's (1983) expression. The modeling of movement is grounded in physics, with its agreed-upon formalism. Brenière et al. (1989), conclude that children in the early stages of walking integrate physical principles of movement sequentially. That conclusion fits well with our own study of the physics of learning to walk (McCollum et al. 1995).

One can imagine extending that study to include the cortical role in movement. For example, the twist style of early walking involves a movement like that of a drawing compass, in which early efforts at walking alternate from foot to foot after each twist. The pattern of reafference encountered in early practice gives rise to a hypothesis (sect. 4.2) that there is a relationship between aspects of sensory reafference and progression toward some goal. For instance, it would be natural to use the vestibular system to gauge the extent of twist. If we assume that the role of the cerebral cortex is to grasp connections and envision possibilities, it would make sense that the cortex would oversee (1) the integration of the connection between reafference and twist (after the fact) and (2) the calibration of twists to reach predetermined goals.

This is an example of second-order learning (I wouldn't call it nonstationary): of finding a goal during learning or laying down the road in walking (Varela et al. 1991). The various elements are available for modeling and many are even available for observation. In constructing this specific example, I am extending Q&S's point that the cognitive structure changes with learning. I would like to emphasize another aspect of real-world learning: the surprises that come in trying something out, for example, walking. Because of such surprises, second-order learning may not be a refinement of an original goal but may take entirely new turns.

Second-order learning also involves burying earlier conclusions: previously learnt skills and procedures become tacit and unavailable for modification. This is a way to understand some of the difficulties faced by recovering stroke patients. A patient who once *learned* to guide footfall by the vestibular system may no longer be able to walk successfully following that procedure (McCollum 1994).

Elaborating and specifying second-order learning will require patience for complexity. One must not get stuck in one's metaphors and methods.

https://doi.org/10.1017/S0140525X97301587 Published online by Cambridge University Press

Neural construction: Two and a half cheers for Quartz & Sejnowski!

Dale Purves

Department of Neurobiology, Duke University Medical Center, Durham, NC 27710. **purves@neuro.duke.edu www.neuro.duke.edu/faculty/purves/purves.htm**

Abstract: A wealth of evidence supports the idea that the circuitry of the developing nervous system is gradually elaborated under the instructive influence of neural activity. Some care must be taken, however, not to handicap the acceptance of this useful perspective with too onerous a theoretical superstructure.

Having struggled – with only modest success – to promote the idea that constructionism is a well supported and eminently sensible framework for understanding neural development (Purves 1988a; 1994; Purves et al. 1996), I am obviously sympathetic to the theses of Quartz & Sejnowski's (Q&S's) "manifesto." It has always been surprising to me that these ideas have not been more widely embraced; indeed, one of my initial fears was that they seemed too obvious to warrant much discussion. But clearly the notion that the nervous system is gradually built in an activity-dependent manner in order to instantiate the lessons of experience is not everyone's cup of tea. In any event, Q&S present an excellent review of the evidence for this. Moreover, by grounding much of it in computational neuroscience, they will no doubt encourage interest in new quarters. There are, nonetheless, a few things to worry about in what Q&S say.

First, the sword of computational neuroscience cuts both ways. Whereas those more familiar with software than synapses will prick up their ears, there is some danger of lapsing into the sort of obscurantism that has bedeviled theories of selectionism (see, for example, Barlow 1988; Crick 1989; Purves 1988b). Although Q&S generally avoid the incomprehensibility that makes the literature on selectionism heavy going, they sometimes struggle to balance a natural affinity for the argot of AI with the knowledge that these ideas are pretty straightforward and don't really need to be gussied up. The desire to examine how "nonstationarity in development can be characterized from a neurocomputational perspective" in terms of "environmentally derived neural activity carrying domain-specific information" (sect. 1.1) is altogether worthy, but one shouldn't have to work so hard to understand the meaning of such statements. The basic idea is that once experience is able to guide the neural connections that generate our awareness of complex stimuli, some pretty neat things are possible.

A second quibble, which is to some degree self-serving, is that Q&S fail to discuss evidence that makes their case stronger (Riddle et al. 1992; 1993; Zheng & Purves 1995). These studies – which examine the effects of normal and augmented levels of neural activity on brain growth in the somatic sensory system – show that particularly active brain regions grow more during development than less active areas. Perhaps Q&S are suspicious of observations on populations of neurons that depend on the macroscopic effects of differential activity, but that do not allude to the more fashionable (but poorly understood) molecular mechanisms that link activity and growth. In my view, such studies not only add teeth to the argument by demonstrating the remarkable effects of differential activity, but they suggest how cortical construction could eventually be studied noninvasively in humans (see also Zheng et al. 1991).

A third problem concerns the assertion that cortical evolution reflects "increasing representational flexibility" rather than "a progressive increase in specialized structures." (sect. 4.4). This part of the narrative is the only one, it seems to me, in which the sense of the ideas (as opposed to the style in which they are expressed) may be off the mark. The notion that representational flexibility and brain specialization are antagonistic evolutionary goals is ill considered. Clearly, complex brains pull off both of these stunts, to the considerable advantage of the brain's possessor (witness the primate visual system; Allman 1985). The teleology of neural development, particularly its evolutionary purposes, is a slippery slope, and venturing onto it here seems gratuitous.

Perhaps the most important implication of the constructionist take on brain development is that experience can instruct the structure and function of the developing brain (in contrast to the view that neural activity is primarily permissive, which, as Q&S point out, has been implicit in mainstream visual physiology for the last 30 years). The most impressive figures among many who have enjoined this debate were archrivals Ewald Hering and Herman von Helmholtz, who in the nineteenth century battled over this general issue for decades (Helmholtz 1924; Turner 1995). Hering, an unabashed "nativist," argued that the intrinsic rules of brain organization supply the visual system with all the analytic ability it needs from the outset. Helmholtz took the "empiricist" stance that many visual perceptions rely on judgments made possible by learning about actual objects and their visual consequences. Despite these minor reservations, Quartz & Sejnowski have done a fine job pulling together the burgeoning evidence for the reality and purposes of activity-driven neural construction, evidence which is rapidly vindicating Helmholtz.

Is the learning paradox resolved?

M. E. J. Raijmakers

Department of Psychology, University of Amsterdam, 1018 WB Amsterdam, The Netherlands. **op_aijmakers@macmail.psy.uva.nl**

Abstract: We argue that on logical grounds the constructivist algorithms mentioned by Quartz & Sejnowski (Q&S) do not resolve the learning paradox. In contrast, a neural network might acquire a more powerful structure by means of phase transitions. The latter kind of developmental mechanism can be in agreement with the constructivist manifesto.

One of the classical arguments against constructivist approaches to development is the logical impossibility of acquiring more powerful structures (Fodor 1980; Q&S refer to this as the "learnability problem"). Has Q&S's constructivist manifesto formally solved this fundamental problem? Or, more concretely, do the constructivist algorithms listed in section 4.2 implement a mechanism for acquiring more powerful structures?

One of the key points of the learning paradox is that the hypothesis space of a system cannot be increased by learning, but only by adding conceptual resources to the language or the algorithm of the system (e.g., by means of maturation). More generally, a system can only become more powerful if the algorithm itself changes qualitatively. As Quartz (1993) shows convincingly, the power of a system can be formalized in terms of the PAC model of learning. A qualitative, as opposed to quantitative, change of an algorithm indicates that: (1) A discrete event is taking place, (2) improvements in power can be predicted from the changes in the algorithm, and (3) no arbitrarily small change of the algorithm can have the same effect in a nontrivial way.

According to Q&S, the power of a biological neural network can be improved by (among other things) extending the dendritic arborization of individual neurons. The constructive algorithms that they mention are neural network models that extend their architecture by recruiting new nodes. Is this a qualitative or a quantitative change of the algorithm of the network? The addition of a node is a discrete event that specifically increases the power, and there is no arbitrarily small change that might cause an equivalent increase. Moreover, it increases the dimensionality of the state space of the network. Hence, these constructivist algorithms implement a qualitative change in the algorithm; in this respect, they do not resolve the learning paradox.

There may be a way, however, to increase the hypothesis space of a system by changing the algorithm only quantitatively. The general idea is that a nonlinear dynamic system can acquire more powerful structures by self-organization, that is, spontaneous organization from a lower to a higher level of complexity by means

https://doi.org/10.1017/S0140525X97301587 Published online by Cambridge University Press

of a bifurcation or phase transition. A phase transition is a discontinuous change in the equilibrium behavior of the system under quantitative variation of a control parameter. The underlying change is quantitative for the following reasons. First, the value of a control parameter can be changed continuously. Second, this change is not specific in the sense that it does not in itself show when a phase transition will occur. Third, the change of the value of the control parameter might be arbitrarily small. Crutchfield (1994) shows a concrete example of an increase in the power (i.e., computational complexity) of a system, the logistic map, by means of a phase transition. In neural network models, phase transitions are common. However, rare are the neural networks that show qualitatively different functional behavior in two dynamic regimes that are separated by a phase transition. Only a few neural networks obey the above description (cf. Pollack 1991). One example is Exact ART (Raijmakers et al., 1996a), which is related to Q&S's constructivist approach. By varying the range of inhibitory connections (not as a function of learning), the activity of the representational layer shows phase transitions. Here, structural properties of the network are implemented as low-dimensional real-value parameters. It appears that, in different dynamic regimes, the learned representations vary from local to distributed. It has not yet been shown formally, however, that these different representations result in networks with different power. It is obviously far more difficult to simulate qualitative improvements with networks that learn or develop through phase transitions rather than constructive algorithms.

Development by means of phase transitions resolves crucial aspects of the learning paradox and hence might be a constructive alternative for the kinds of constructive algorithms mentioned by Q&S. As we tried to expound, there exist crucial dynamic differences between the two kinds of constructive mechanisms. For this reason, they should be considered as two different mechanisms of constructive development which might exist alongside one another. Empirical criteria can distinguish between different kinds of dynamics underlying a specific developmental process (van der Maas & Molenaar 1992). Important empirical evidence is shown for transitions in the acquisition of the liquid conservation task and the balance scale task, although no absolute demonstration exists yet (Raijmakers et al. 1996b; van der Maas & Molenaar 1996).

Constructivism: Can directed mutation improve on classical neural selection?

George N. Reeke, Jr.

Laboratory of Biological Modelling, The Rockefeller University, New York, NY 10021. **reeke@lobimo.rockefeller.edu; www.rockefeller.edu/labheads/reeke/ reeke-frame.html**

Abstract: Quartz & Sejnowski (Q&S) find flaws in standard theories of neural selection, which they propose to repair by introducing Lamarckian mechanisms for anatomical refinement that are analogous to directed mutation in evolution. The reversal of cause and effect that these mechanisms require is no more plausible in an explanation of cognition than it is in an explanation of evolution.

1. Mischaracterization of selectionist theories. Quartz & Sejnowski choose to "refer to any view positing an initial exuberant growth followed by elimination as 'selectionism'" (sect. 2). This definition seriously mischaracterizes mainstream selectionist theories, such as Edelman's (1978; 1987) theory of neuronal group selection (TNGS). First, it has never been claimed that all growth needs to be completed before selection begins. The possibility and even necessity of continuing diversification of neural systems during the lifetime of the organism has been explicitly discussed by Edelman (1987, p. 19) as well as included in computer simulations based on the TNGS (Montague et al. 1991). Second, it has never been claimed that all neural selection is *negative* selection. Positive mechanisms related to the so-called Hebb rule have been included in all computer models based on the TNGS to date (e.g., Reeke et al. 1990). Third, it has never been claimed that regression of neural structures is a necessary concomitant of negative selection. It is enough that synaptic efficacy be reduced. Mechanisms of this kind have also been included in all computer models based on the TNGS from the very earliest (Edelman & Reeke 1982). Q&S neglect the very attractive possibility that structures not selected in one environmental encounter may in this way remain available for selection in future encounters. In any event, it is no surprise to a selectionist that large scale regression of neural structures is not seen during postnatal maturation in typical mammals.

In addition, Q&S for the most part equate increases in complexity (function) with increases in size (anatomy). In fact, complexity and size are quite different concepts (Tononi et al. 1996). Sheer size does not in itself guarantee complexity if the components of the system are connected in a regular fashion. Furthermore, reorganization within a structure of fixed size can lead to increased complexity. Accordingly, the statistics on neurite growth and number of synapses presented by Q&S are interesting, but not to the point.

2. Constructivist growth is directed growth. The end results Q&S expect of constructivist growth are quite clear: "dendritic development under the influence of environmentally derived activity conforms to cognitive schemes for the construction of mental representations" (sect. 1.1). It is less clear how activity is supposed to direct neurite growth and synapse formation to construct structures embodying appropriate mental representations. The target article can be read in two very different ways: either Q&S have merely suggested a possibly more efficient variant of mechanisms already proposed and extensively studied in selectionist simulations, or they have proposed (but not described) a radically new mechanism of directed growth in which the effect of making a new connection is somehow known before the connection is formed.

The heart of the issue is the discussion in sections 3.1 and 3.1.1 of unpublished simulations (with P. R. Montague) in which neurite growth is said to be directed in an activity-dependent manner. Several mechanisms apparently control this growth. In one of these, the probability of terminal segment neurite branching is made to depend on conditions like those used to modulate synaptic strength in the earlier models of Montague et al. (1991). These conditions are based on suggestions by Gally et al. (1990) regarding the possible role of nitric oxide in the development of neural connectivity. This mechanism conforms to the classical selectionist picture in that connection formation is nonspecific (although modulated by local activity); functional specificity is a result of inactivating those connections whose activity, *which can occur only after the connection has been formed*, is not correlated with that of other connections to the same postsynaptic cell. While Q&S may have improved the efficiency of this mechanism by transferring local modulation from synapse formation to the earlier stage of neurite branching, it should still be clear that the integrated volume signal, which, due to diffusion limitations, cannot differ significantly at neighboring synaptic sites (whether it is carried by nitric oxide or some other agent), and cannot contain the specific information needed to determine the individual fate of each synapse in the region.

Q&S allude without details to other mechanisms which might come into play. Supposing that these mechanisms operate without the pruning that Q&S have criticized so strongly, we would then have to accept the more radical possibility that synapses could be formed with appropriate connectivity prior to the initiation of synaptic transmission through those same synapses, which is necessary in the earlier theories to guide local selection events. This possibility stands to neural selection as directed mutations stands to natural selection. While specialized mechanisms might permit the appearance of directed construction (mutation) in very specific circumstances, this type of mechanism is simply not viable as a general basis for cognitive development (evolution) because it requires that an effect occur before its cause. One is therefore

https://doi.org/10.1017/S0140525X97301587 Published online by Cambridge University Press

forced to conclude that undirected or broadly but not specifically directed growth followed by pruning is the necessary principal mechanism of neural selection. The theory presented by Q&S, once shorn of its unsupportable suggestions of directed mutation, differs from earlier theories in its emphasis on a closer temporal coupling of growth and selection processes, but it is otherwise substantially the same theory distinguished with a new name.

3. Summary. Q&S make a valuable contribution by reminding us that postnatal development leads to increased anatomical complexity of the nervous system which must be taken into account in constructing theories of learning. They have also very properly emphasized the importance of forming mappings with local coherence, which are generally lacking in connectionist networks. However, they have taken a "directed mutation" stance in regard to the mechanisms for formation of new anatomical diversity, which essentially defines away the very real problem of how the system can "discover" appropriate behaviors in previously unexperienced situations. This approach exaggerates the differences between their theory and earlier selectionist approaches.

The right way, the wrong way, and the army way: A dendritic parable

Arnold B. Scheibel

Departments of Neurobiology and Psychiatry and Biobehavioral Sciences and Brain Research Institute, U.C.L.A. Center for the Health Sciences, Los Angeles, CA 90024. **ascheibe@bri.medsch.ucla.edu**

Abstract: We suggest that neither selectionism nor constructivism alone are responsible for learning-based changes in the brain. On the basis of quantitative structural studies of human brain tissue it has been possible to find evidence of both increase and decrease in tissue mass at synaptic and dendritic levels. It would appear that both processes are involved in the course of learning-dependent changes.

The neurosciences have seen more than their share of impassioned conceptual dualities. Reticularism versus neuronism and "soup versus spark" synaptic transmission dynamics are two among many that come to mind. It is interesting to recall that neural reality was finally determined to encompass both poles of each duality. Neurons were indubitably separate entities but in the case of gap junctions, virtually continuous through the agency of connexions establishing structural continuity for ion flow. Neurons clearly communicate through the agency of neurotransmitter release but "electrical" transmission remains a reality at gap junctions. I would suggest that we may be dealing with another impassioned duality in the matter of "selectionism versus constructivism."

Quartz & Sejnowski (Q&S) have presented powerful evidence supporting environmentally engendered dendritic and axonal growth. Ideationally implicit in the work of Ramon y Cajal (1952) and further explored at the conceptual level by Hebb (1949), the conjectures became reality at the chemical level in the studies of Krech et al. (1960) and, at the structural level, in the extended series of studies by Diamond and her collaborators (1964; 1975; 1988) and in a number of other findings. Our own quantitative histological studies of human cerebral cortex argued strongly for causal links between computational complexity and structural complexity (Scheibel et al. 1990). Thus dendritic elaboration in the primary sensory cortical representational area for hand and fingers was significantly greater than that in the adjacent area for trunk representation. Furthermore, there were "suggestive associations between the complexity of dendrite systems of the hand-finger zone of the primary receptive area and the nature of the work with which the individual had been associated during his/her working life" (Scheibel et al. 1990, p. 85). Furthermore the conjoint development of language facility and waxing dendrite elaboration in Broca's area of the language-dominant hemisphere (Simmonds & Scheibel 1989) provided correlative if not causal relations between escalating cognitive demands and expanding neuropil.

Arguments can also be advanced for selectionism, however. In several series of electron microscope studies performed on rodents, measurable and significant decreases in the number of synaptic terminals in cortical axo-spino-dendritic synapses accompanied exposure to enriched environments (e.g., Mollgard et al. 1971). Individual synaptic terminals showed significant increase in the length of the postsynaptic thickening, thereby suggesting the presence of fewer, but larger and more effective synapses in environmentally enriched animals. Further analysis of these changes indicated that the effects of enriched environmental input as expressed in loss of synaptic terminals and enlargement of the remainder actually increased with age (Diamond et al. 1975). And the enriched rats were quicker maze-learners than their non-enriched mates (Diamond 1988)!

Assuming that a complex interweaving of dendritic/synaptic gain and loss are involved in the maturation-learning process, a third mechanism seems intertwined with these two, adding to the richness and subtlety of the process. Quantitative comparisons of dendritic tissue in Broca's area of left and right hemispheres revealed an unexpected result (Scheibel et al. 1985). There was no significant difference between the total dendritic length of neurons on either side. What did differ was the amount of dendritic length 'invested' in various portions of each dendritic tree. On the right, the non-language-dominant side, most of the dendrite length was involved in the first three orders of dendrite branching. On the language-dominant side, a much greater proportion of dendrite length was devoted to the outer branches (fourth, fifth, sixth order dendrite branches, etc.). Note that the inner, lower order branches developed earlier in the developmental history of the individual, while the outer branching segments developed later. Thus both *temporal patterns* of development and *position* on the dendrite tree were significant parameters in CNS growth and maturation. Note also, that successive additions to the periphery of the dendrite ensemble should (at least theoretically) not affect the more central parts of the dendrite system where synaptic patterns had presumably already been established. However, more than a tidy "add-on" effect was noted here. Our data (Simmonds & Scheibel 1989) strongly suggested that along with the pattern of use-dependent centrifugal *growth* there was also a related (and presumably use-dependent) partial *resorption* of lower order branches more centrally located within the dendrite ensemble. Simultaneous involvement of cortical dendritic tissue gain and loss during the maturation-learning process argues for the inextricable combination of constructivist and selectionist processes.

Neural constraints on cognitive modularity?

Brian J. Scholl

Rutgers Center for Cognitive Science, Rutgers University, New Brunswick, NJ 08903. **scholl@ruccs.rutgers.edu**

Abstract: Is (some) innate *cognitive* modularity consistent with a lack of innate *neural* modularity? Quartz & Sejnowski's (Q&S's) implicit negative answer to this question fuels their antinativist and antimodular cognitive conclusions. I attempt here to suggest a positive answer and to solicit discussion of this crucial issue.

Cognitive development figures prominently in the title and concluding paragraphs of this target article, but receives too little discussion in the middle. I would like to solicit a more explicit discussion of the precise relationship between the neural evidence and cognitive conclusions proffered by Q&S. The burning question: *Is (some) innate cognitive modularity consistent with a lack of innate neural modularity?* The answer, for Q&S, appears to be (a quite implicit) "no." I would like to question this assumption.

The essence of modularity is a restriction on information flow. A function is modular to the degree that it is "informationally

https://doi.org/10.1017/S0140525X97301587 Published online by Cambridge University Press

encapsulated" – such that some of the information outside the module is not accessible from within. *Cognitive* modularity then concerns the impenetrability of the function by outside cognitive influences (e.g., Pylyshyn 1985). The converse may also hold: some of the "inter-levels" within the module may not be available from the outside. These restrictions on information flow engender several other symptoms: modules are typically fast, unconscious, obligatory, computationally shallow, and possessed of characteristic breakdown patterns. Some writers have suggested that the conjunction of all these symptoms is required for modularity (e.g., Garfield 1994), but this is not an apt characterization. Fodor is sometimes ambiguous on this question, but it seems clear that "informational encapsulation is . . . the essence of . . . modularity" (Fodor 1983, p. 71). [See also *BBS* multiple book review of Fodor's *The Modularity of Mind BBS* 18(1) 1985.]

Cognitive modularity thus imposes no constraints on how these patterns of information flow are realized in the underlying biological architecture (neural or otherwise): the relevant cognitive functions could be neurally implemented in a highly distributed fashion, even though they comprise a module at the cognitive level. This is not to say that cognitive modules *cannot* be neural modules – indeed, that they are often seems to be the case, and neural localization may often be useful evidence *for* cognitive modularity. I am instead trying to emphasize our vast ignorance concerning the mechanics of the biological implementation of cognition by pointing out that cognitive modules *need not* be neurally localized. This is often recognized by cognitive researchers. Segal (1996), for example, comments in a discussion of the modularity of "theory of mind" that "it's at least *a priori* possible that distributed, global characteristics of the brain, rather than [neurally] modular ones, realise computational or intentional modules" (pp. 145–146).

Q&S fail to address such considerations. They suggest that "the view that strong, domain-specific knowledge is built into cortical structures runs into severe difficulties from developmental neurobiological evidence" (sect. 4.1.1, para. 2). This depends, I think, on an overly restricted view of how exactly this "knowledge" could be realized neurally. Q&S do not discuss precisely what they mean by modularity (though they cite Fodor 1983), but they have elsewhere clarified that from their perspective "the modularity thesis becomes a question of whether some cortical regions are prespecified for domain-specific knowledge that determines *a priori* the computation a region may perform" (Quartz & Sejnowski 1994, p. 726).

I think this is simply incorrect. The modularity thesis (in this sense) concerns whether there exists domain-specific knowledge, *period* – and imposes no constraints on how that knowledge might be implemented in the underlying neural architecture. (Actually, that's not quite right either. Cognitive modules are also often domain-specific, but this need not always be the case: domain specificity refers to the sorts of inputs a system can process, while informational encapsulation refers to the information that the system can make use of whilst processing. This misreading of modularity will not matter here, since the same considerations about implementation apply whether you're talking about encapsulation or domain-specificity.) In any case, I think that in characterizing modularity this way, Q&S greatly diminish the impact of their arguments for the many cognitive scientists who use different notions of modularity. Q&S manage to draw anti-modular cognitive conclusions from their neurobiological evidence, but only because the relevance of that neural evidence crept illicitly into their initial (and implicit) conception of cognitive modularity in the first place.

Now, as Karmiloff-Smith (1992) and others have stressed, it is conceivable that some sorts of cognitive modularity may actually be acquired, without any innate predisposition. [See also *BBS* multiple book review of Karmiloff-Smith's *Beyond Modularity BBS* 17(4) 1994.] My comments thus far have been agnostic with regard to this issue. I have been suggesting that some cognitive modularity – however acquired – may be consistent with a lack of neural modularity. The same sorts of considerations, however, may also bear on the question of innate cognitive structure more generally. Q&S marshal an impressive array of evidence that cortical development can be affected by the environment in all sorts of ways (see their Table 4). Is this consistent with a degree of innate *cognitive* structure? Q&S require a negative answer, which again is simply assumed. But the fact that particular cells' functions depend on their interaction with the environment (see their sect. 2.3.3) may have no implications at all for how these cells will end up being used by innately determined *cognitive* processes. In any case, the specific types of environmental effects matter crucially here. No nativist ever denied the importance of environmental impact – witness appeals to triggering, imprinting, parameters, critical periods, and so on. To address nativism about cognitive structure, neurobiological evidence would have to demonstrate that bona fide *learning* was involved (see Pylyshyn 1985, especially pp. 409–14).

It remains an open question whether neurobiological equipotentiality is relevant to the evaluation of theories of cognitive modularity: at this point, we simply haven't a clue how cognition is built into the underlying, implementing biological architecture. At bottom, therefore, I think many cognitive scientists and neuroscientists simply assume different answers to the burning question above, in pursuing their research. This issue is still quite controversial, even within cognitive neuroscience (cf. Sarter et al. 1996), and this may be the ideal forum in which to devote some explicit discussion to these assumptions. Quartz & Sejnowski suggest that their theory "provides a meeting ground for cognitive scientists and neuroscientists" (sect. 5, para. 2). Maybe so; I would like to suggest the issues raised here as among the first items on the meeting's agenda.

Deconstructing neural constructivism

Olaf Sporns

The Neurosciences Institute, 10640 John J. Hopkins Drive, San Diego, CA 92121. **sporns@nsi.edu** **www.nsi.edu**

Abstract: Activity-dependent processes play an active role in shaping the structure of neuronal circuitry and therefore contribute to neural and cognitive development. Neural constructivism claims to be able to account for increases in the complexity of cognitive representations in terms of directed growth of neurons. This claim is overstated, rests on biased interpretations of the evidence, and is based on serious misapprehensions of the nature of somatic variation and selection.

Much evidence points toward the importance of correlated neuronal activity in forming and maintaining neuronal architecture. Such correlations occur throughout an animal's lifetime, in parallel with developmental processes such as structural regression and growth of neural connectivity. Correlated neural activity may reflect important aspects of the statistical structure of the environment and is thought to be a key factor in neural plasticity. In their target article, Quartz & Sejnowski (Q&S) review a variety of experimental and computational results and link structural neuronal plasticity to learning theory. Their aim is a comprehensive account of cognitive development in terms of environmentally guided directed growth of neurons and connections. This so-called "neural constructivism" is counterposed to what is called the dominant or popular view: selectionism. Q&S's reasoning is flawed, however, and the synthesis they attempt to reach ultimately eludes their grasp.

Q&S's "selectionism" is a caricature at best, a straw man deliberately erected only to be knocked down. No one has ever proposed that all of brain development can be subsumed in a two-stage process, initial overproduction and subsequent elimination. Quite the contrary, according to selectionism (Edelman 1987; 1993; Sporns & Tononi 1994; Sporns 1997a; 1997b), the basic processes of selection in the nervous system overlap temporally throughout

https://doi.org/10.1017/S0140525X97301587 Published online by Cambridge University Press

the lifetime of the organism. Selectionism stresses the simultaneous and concurrent nature of (1) the ongoing generation of *variability*, (2) the *interaction* of variant neuronal elements with an environment and (3) the differential *amplification* of elements, mediated by local correlations or by diffuse ascending (value) systems. Selectionism transcends arbitrary boundaries between "developing" and "adult" neural function, sidesteps silly discussions of "nature versus nurture," is clearly consistent with progressive elaborations of neuronal circuitry, and naturally incorporates environmental influences and an organism's behavioral activity as key components in any structural and functional modification of neural architectures. Selectionism is neither "nativist" nor "empiricist," in that it does not deny the existence of some specialized circuitry as products of evolutionary change (as is incorrectly indicated in sect. 4.4), while at the same time emphasizing the need for flexibility of developmental and functional processes to accommodate unpredicted environmental change. The degree to which Q&S (and others: see Purves et al. 1996) ignore the important role played by concurrent variation and selection in the generation of complex structure, be it in evolution (a case that can hardly be denied) or in brain function, is remarkable. Not only do they deny the creative dimension of variation and selection, but they do not seem to be aware of the rich spectrum of selectional modes, which includes eliminative selection and many other mechanisms.

Central to Q&S's constructivist agenda and their attack on selectionism is the concept of "directed growth," according to which neural activity (experience) "modulates neuropil growth, rather than functioning to stabilize useful connections from an initial excess" (Purves 1994, p. 91). But is "directed growth" really borne out by the evidence? Detailed anatomical observations (e.g., Wolff et al. 1995) strongly support the continuing formation and elimination of synaptic connections, that is, the parallel occurrence of degenerative and progressive processes throughout the entire lifetime of the organism. Direct observations of growth processes involving synaptogenesis (Ziv & Smith 1996) and structural changes during axonal remodeling (Witte et al. 1996) and plasticity (Bailey & Kandel 1993) indicate that numerous seemingly stochastic microevents may form a basis for global changes in the morphology of cell processes. While some recent work on neurotrophins seems consistent with their action as an activity-dependent signal to induce sprouting (e.g., Cohen-Cory & Fraser 1995), volume effects such as these might also be expected if the balance between construction and destruction of neuronal processes is subtly altered. Nowhere does the inconsistency of neural constructivism become more apparent than in the muddled and confused discussion of "regressive events in development" (sect. 3.3). Q&S all but acknowledge the importance of variation and selection (here demoted to "stochastic sampling" and "error correction"), only to slip back into their fundamental (pre-Darwinian) error of reasoning ("If development is directed it can't be selectionist").

Neural constructivism proposes a direct link between the development of the brain's presumed representational properties and progressive elaborations of circuitry (in particular, of dendrites) by directed growth. However, the mechanism by which any but the simplest environmentally derived correlations could be translated into useful new representations remains totally unclear; in fact, such mechanisms smack of neural Lamarckism (Changeux 1997). Although it is undeniable that activity-dependent correlations can have strong effects on development and plasticity of neural firing patterns, according to neural constructivism the environment and its changing patterns of input somehow *instruct* brain connectivity to grow in ways that are meaningful and beneficial to the organism. Within the framework of modern biology, such instructive mechanisms appear contrived and unrealistic. No such problem exists, however, if neural and cognitive development is viewed from a selectionist perspective. An integral part of selectionist thinking is a set of value systems, anatomically defined as diffuse ascending systems of the brain, which provides modulatory input to widespread regions of the brain and influence neuronal activity and synaptic change. Their power as constraints on somatic selection, analogous to similar constraints on fitness in evolution, has been demonstrated in numerous computational models (Edelman et al. 1992; Reeke et al. 1990; Rucci et al. 1997; Sporns & Edelman 1993; Tononi et al. 1992), including models that allow for the gradual emergence of value-related signals as environmental conditions change (Friston et al. 1994).

Q&S draw a parallel between the perceived increase in the structural complexity of neural circuits and the complexity of cognitive representations. Detailed analyses of key structural parameters of axonal and dendritic morphologies have indeed been carried out by others (e.g., Greenough & Chang 1988), although the relationship of these studies to the issue of complexity still remains unclear. Unfortunately, Q&S, do not attempt to define or quantify complexity in relation to either anatomy or cognitive representations. This is admittedly a difficult enterprise, but one that would put neural constructivism to the test. Q&S totally sidestep the fundamental question of how anatomical patterns (and their complexity) relate to functional interactions between neurons (i.e., in terms of correlations). Studies of the complexity of functional interactions in the brain suggest that realistic anatomical patterns give rise to highly complex dynamics (Tononi et al. 1994). It is this link between structure and function and its quantitative relationship to neural complexity, as well as the relationship between environmental and intrinsic (neural) patterns of correlations (Tononi et al. 1996) that need to be explored further.

Whereas the evidence by others reviewed by Q&S is illuminating and significant for our understanding of neural and cognitive development, their synthesis, unfortunately, is not. Based on false premises and serious misapprehensions of the nature of selection, this "constructivist manifesto" does not present a single convincing argument for abandoning experimental or theoretical approaches inspired by selectionism. On the contrary, in light of the evidence, selectionism with its theoretical pillars of variability, environmental interaction, differential amplification, and value is more appealing than ever.

Learning is remembering

George Székely

Department of Anatomy, University Medical School, H-4012, Debrecen, Hungary. **szekely@chondron.anat.dote.hu**

Abstract: The strong correlation between the geometry of the dendritic tree and the specific function of motoneurons suggests that their synaptic contacts are established on a *selective* stochastic basis with the characteristic form of dendrites being the source of selection in the frog. A compromise is suggested according to which specific structures may have evolved on a *selective* stochastic basis and "constructive learning" could be the source of selection in the cortex.

The fundamental thesis of the target article is so captivating that, despite my ignorance in cognitive sciences, it has lured me into commenting on this elegant though provocative manifesto. Quartz & Sejnowski (Q&S) state that constructive learning is a dynamic interaction between environment and neural mechanisms that also change while they learn. This minimizes the need for prewired domain-specific structures in the cortex. I fully appreciate the far-reaching consequences of this suggestion, especially from the point of view of a neuroscientist who pursues alternatives to Sperry's (1963) neuronal specificity in the organization of neuronal connections. Q&S are aware of the difficulties of "pure tabula rasa learning," and hasten to add that there must be some built-in structure to start with.

The first question that arises for a neurohistologist who studies the minute organization of lower centers is: How much prewired structure is indispensable initially? In this context one cannot help

https://doi.org/10.1017/S0140525X97301587 Published online by Cambridge University Press

but think of the vague statements of parents about the special talent of their children for music, mathematics, languages, and so on, partly in consolation for the child's failure in other subjects, but mainly because there really is a special aptitude in many children in one or more specific areas. The Bach family in music and the Bolyai family in mathematics are examples of the substantial degree of innately specified skill that exists in some domains. [Cf. Howe et al. "Innate Talents: Reality or Myth?" *BBS* 21(2) 1998.] In sections 2 and 3 of the target article, data demonstrate the effect of function on the form of neurons. I would like to add a complementary discussion of how form relates to a specific function, not in the cortex, but in lower centers. I will conclude my commentary with a suggested compromise.

Working on the spinal control of limb movements in amphibia, we have found a pronounced morphological heterogeneity, associated with functional differences, in the motoneuron column (Székely & Czéh 1976). There is a special type of motoneuron that controls limb movements; a different type coordinates swimming. The morphological differences in the brainstem motor column are more pronounced and are closely associated with the specific function of the motor nucleus (Székely & Matesz 1993). There are two interesting aspects of this morphofunctional property of brainstem motor neurons that may be relevant to the themes of the target article.

The strong correlation of dendritic geometry with a particular function suggests that dendrites may contain information about how to engage with the surrounding synaptic field in a given volume. A frog motoneuron carries 4–6 $\times$ 10^4 synaptic contacts (Birinyi et al. 1992). The exact site of this vast number of synapses originating from a variety of places would be difficult to specify with address-encoding surface markers of the kind suggested by Sperry (1963). If the chemospecificity hypothesis is abandoned, one must rely on the statistical arrangement of synaptic connections. These statistics, however, cannot be completely random; they must be *selective* in the sense of Simon (1962) in one way or another. A possible source of selection can be the characteristic geometry of the dendritic tree. This hypothesis was tested in the frog, an animal which can only close and open the mouth. Motoneurons of the closing and opening muscles are dispersed in the trigeminal and facial motor nuclei that lie quite close to one another in the pons. Neurons (No 52) were labeled with a Co^{3+} complex salt which is supposed to reveal the whole of the dendritic tree. The geometry of the dendritic arbor of each neuron was quantitatively characterized with 32 morphological parameters. Multivariant statistical techniques classified the neurons according to their function (Matesz et al. 1995). Although the results alone do not demonstrate that the trigeminal and facial motoneurons are linked to a premotor structure entirely on a statistical basis, they do indicate the instrumental role of the dendritic arbor in endowing the neuron with a specific function.

The second aspect of this morpho-functional correlation of the motoneuron dendritic tree reveals a relation to evolution as well. The very pronounced and obvious changes that can be observed in the head region during phylogenesis are mirrored in the structure of the cranial nerve motor nuclei (Székely & Matesz 1993). For example, a muscular tongue appears first in mammals, and so does the hypoglossal nucleus with its uniquely characteristic form of neurons innervating intrinsic tongue muscles. The joints and muscles in the mammalian jaw undergo a profound alteration in phylogenesis, resulting in the capacity for making complex grinding movements during mastication. The type of motoneuron that controls the simple closing of the mouth entirely disappears and is replaced by a new type of "masticatory" neuron in the mammalian brain stem. These neurons innervate the masticatory muscles, which are novel acquisitions in phylogenesis (Starck 1982). It is interesting to note that the mammalian jaw-opener muscles are homologuous to the amphibian openers, and the cell-type of the amphibian "opener" neuron is found in the mammalian accessory trigeminal and facial nuclei innervating the suprahyoid muscles. The mammalian facial nucleus is also a novel acquisition and innervates the muscles of facial expression that appear first in mammals. These and many similar examples indicate that evolution discards old nonfunction elements and creates new elements, both central and peripheral, to implement new functions.

The implication of these observations for the dynamism of constructive learning is that the possibility of establishing ordered synaptic connections on a stochastic basis should not be overlooked and that selective pressure can shape the structure giving rise to new functions in evolution. During the long process of cortical evolution, similar mechanisms may have played a role in shaping its continuous differentiation. Specific structures may have evolved on a *selective* stochastic basis, and constructive learning (and other functions) could be the source of selection.

Broca's area in the left operculum may represent one such innate structure. It is not my intention to rekindle the dispute about how languages are acquired (certainly not in this area), but the well known fact that this area is irreplaceable in case of damage suggests that it must have differentiated by some innate mechanism. Even if it needs instruction to become the motor speech center, this, and only this, receives the instruction. The special talents and mental gifts revealed in early childhood may be explicable on the basis of innate structures that are especially amenable to mastering some particular art or science. Some nativist ideas could help us understand how a special gift can be passed from one generation to the next (but it remains uncertain how the innate structures originally come about).

From these considerations a compromise can probably be reached to the effect that there are many more "built-in assumptions" than one takes home from Q&S's target article: But there is the upbeat message about the structure-building capacity of "constructive learning." If one lets one's fantasies roam in Lamarckian pasture, previous interactions with the structured environment may have left their traces in the uncommitted parts of the brain and formed cores of "built-in assumptions" put to use at a later time. At risk of misusing the authority of ancient Greek philosophers, Plato's suggestion that all learning is remembering could be embodied by a brain well supplied with "built-in assumptions" and a capacity for "constructive learning."

ACKNOWLEDGMENT

This work was supported from a grant by the Hungarian National Research Found No. T 020257.

Authors' Response

Controversies and issues in developmental theories of mind: Some constructive remarks

Steven R. Quartz[a] and T. J. Sejnowski[b]

[a]*Computational Neurobiology Laboratory and The Sloan Center for Theoretical Neurobiology, The Salk Institute for Biological Studies, La Jolla, CA 92037;* [b]*Howard Hughes Medical, The Salk Institute for Biological Studies and Department of Biology, University of California at San Diego, La Jolla, CA 92037.* **steve@salk.edu** ■*and* **terry@salk.edu; www.cnl.salk.edu/CNL/**

Abstract: As the commentaries reveal, cognitive neuroscience's first steps toward a theory of development are marked by vigorous debate, ranging from basic points of definition to the fine details of mechanism. In this Response, we present the neural constructivist position on this broad spectrum of issues, from basic questions such as what sets constructivism apart from other theories (particularly selectionism) to its relation to behavioral theories and to its underlying mechanisms. We conclude that the real value of global theories at this stage of cognitive neuroscience is not just their answers but the new set of research questions they pose.

https://doi.org/10.1017/S0140525X97301587 Published online by Cambridge University Press

As recently as a decade ago, a target article like ours would have been met with nearly universal criticism. Our mixed levels of description – exploring how the cognitive and neural interact – would have been labeled a conceptual confusion by many. And to say that neural structures emerged as a result of learning would have violated many of the core assumptions of mainstream cognitive science. Hence, before we answer the commentators' interesting queries and challenges, we think that it is worth reflecting on the quite striking fact that the overwhelming majority of commentators are sympathetic to the basic aims of the target article. By our count, only five (**Bickerton; Dehaene-Lambertz & Dehaene; Feldman; Reeke; Sporns**) are categorically against the position. Of these, Sporns, Reeke, and Dehaene-Lambertz & Dehaene favor selectionism over neural constructivism. Bickerton defends nativist linguistics and Feldman, though supportive of the idea of constructive development, has technical objections.

None of the commentators defended an intrinsic maturational view of the brain. **Bickerton** asks us how we can distinguish between intrinsically regulated and environmentally regulated growth, but he is careful to point out that he is not thereby endorsing the maturational position. Similarly, none of the commentators have charged us with confusing levels of description, although **Scholl** does defend a weaker version of the independence of levels argument. That there exists an autonomous "cognitive" level of explanation irreducible to the neural is the core assumption of traditional cognitive science (Pylyshyn's [1984] "autonomy thesis").[1] The move away from the maturational view and its autonomy thesis of not so long ago is to us one of the most significant themes of the commentaries. Some, including a few of the commentators themselves, may find this surprising. But, then, we think that some researchers outside the Chomskian tradition in cognitive science underestimate how extreme its developmental claims were. Here are two representative examples:

> Linguistic information is innate in humans if and only if it is represented in the gene code (Fodor et al. 1974, p. 450).
>
> Grammar genes would be stretches of DNA that code for proteins or trigger the transcription of proteins, in certain times and places in the brain, that guide, attract, or glue neurons into networks that, in combination with the synaptic tuning that takes place during learning, are necessary to compute the solution to some grammatical problem (Pinker 1994, p. 322).

This debate between genes and environment, or intrinsic versus extrinsic maturation, at last appears to be over. In this respect, cognitive science has caught up with neurobiology and no longer concerns itself with *whether* the environment thoroughly influences brain development, but *how* it does. Indeed, the latter question was by far the most vigorously pursued issue, aligning the commentators either toward selectionism or constructivism. Before we consider where selectionism and neural constructivism differ, it is important to remark that they do agree in some basic ways. Both, for example, agree that an understanding of development is required to understand the mature state, and so both urge for a developmentally centered approach to the mind. And both attempt to account for mature neural structures by characterizing developmental mechanisms that move beyond the genetic determinism of traditional cognitive science.

This is not to say that there are no differences between selectionism and constructivism. Neural constructivism retains the information-processing and computational approach of traditional cognitive science, albeit with a very different computational paradigm. Selectionism, on the other hand, abandons even this framework, dropping the information-processing view (since it denies the world is informationally structured) and declaring that the brain is not computational.

It is not surprising, then, that in this shifting debate the sharpest criticisms came from selectionists (**Dehaene-Lambertz & Dehaene; Reeke; Sporns**). It is the sign of a young and vigorous new approach that fundamental issues of definition are being thrashed out. And it was no surprise that there was little consensus even where the positions differed or where their core assumptions lay. Since these questions – What is the real difference between these theories? and Where does the empirical support fall? – were the major points of contention, they will be our starting point.

R1. Deconstructing constructivism or reconstructing selectionism?

The selectionist commentators (**Dehaene-Lambertz & Dehaene; Reeke; Sporns**) all charged us with creating a straw man, a caricature we propped up only to knock over. The crux of their objection was that we unfairly characterized selectionist development as a two-phased process: an initial overproduction of neural structure followed by selective stabilization of a proper subset of that structure. In place of this view, they suggested that selectionism involves multiple waves of exuberant growth, each followed by selection. We have three responses: (1) throughout the selectionist literature there are recurring treatments of development as a two-phased process; (2) the bulk of supporting empirical evidence selectionists cite characterizes development as a two-phased process; (3) our objections to selectionism do not depend on this characterization, and so hold even for selectionism with multiple waves of exuberant growth/retraction. Indeed, we think that once selectionists admit progressive growth, they open the floodgates to the very "instructivist" developmental mechanisms they seek to banish from biology as a pre-Darwinian confusion. We will present these three replies in turn.

R1.1. The two phases of selectionism. Selectionism advertises itself as "a radically new view of the function of the brain" (Edelman 1987). We think much of the perception of selectionism as a radical alternative, and therefore appealing, lies in the idea of development as an initial overproduction followed by selection. Indeed, we are not the only ones to suggest that selectionists present development as a two-phased process. The idea that selectionism involves a developmentally timed overproduction of synapses followed by a selective pruning back of connections can be found in the commentaries by **Black & Greenough, Hurford et al., Innocenti,** and **Purves.** Beyond the commentaries two well known critics of selectionism, Crick (1989; p. 240) and Purves et al. (1996) also extracted this two-phased interpretation from selectionist presentations.

Is this interpretation simply a rampant misunderstanding? If it is such a radically different theory, why is there so much disagreement over what distinguishes it from others? After so many programmatic statements of selectionism, we doubt that its dominant interpretation owes itself to nothing

https://doi.org/10.1017/S0140525X97301587 Published online by Cambridge University Press

more than widespread confusion. Looking over the selectionist literature we find many instances of development being presented as a two-phased process. For example, a two-staged process is explicitly presented by Edelman (1987) in one of the major statements of the position:

> Diversification of anatomical connectivity occurs epigenetically during development, leading to the formation by selection of primary repertoires of structurally neuronal groups. . . . A second selective process occurs during postnatal behavior through epigenetic modification in the strength of synaptic connections within and between neuronal groups. . . . This selection occurs within the original ensemble of anatomically variant groups (the *primary repertoire*), and it results in the formation of a *secondary repertoire* consisting of functioning groups that are more likely to be used in future behavior. (p. 5)

According to Edelman (1987), then, genetic and epigenetic processes internal to the brain create a "primary repertoire." He then states, "[I]n order to account for map changes by selection upon a preexisting [primary] repertoire to create a secondary repertoire, it is necessary to show the prior existence of arbor extension, variability, and overlap." Cortical maps are thus created as selective mechanisms operating on a primary repertoire to sculpt a secondary repertoire, the mature map. Since the primary repertoire exists prior to the operation of selective mechanisms, the suggestion that these are two discrete processes seems not unreasonable.

It is here that **Elliott & Shadbolt** accuse us of misrepresenting selectionism. In particular, they find it misleading to call the initial connectivity "pre-representations." The term "pre-representations" is not one of our own invention; it is a standard selectionist one (e.g., Changeux & Dehaene 1989).

What about the formation of the secondary repertoire – do later waves of growth/retraction occur? Regarding later forms of plasticity, Edelman (1987, p. 138) states, "while some neuronal sprouting may occur after early development in such systems, it is likely that the major competitive forces are the result of synaptic changes." In other words, synaptic weight changes underlie mature forms of learning (Crick 1989 gleans the same conclusion). Edelman does allow the possibility of later progressive development, but the following statement downplays its importance:

> While the original degeneracy of the system is statistically likely to decrease as learning occurs, it remains statistically possible that new variations can occur within interacting neural networks and hierarchies of networks for the lifetime of the organism. (Edelman 1987, p. 19).

Is it unreasonable to infer from this that a process that "remains statistically possible" is not considered to be a major component of the theory? Other selectionists explicitly endorse a two staged theory. Piattelli-Palmarini (1989), for example, states, "the present trend is to grant a very rich innate repertoire and then look for the mechanisms of internal selection." In fairness, some of the selectionist literature makes explicit claims regarding multiple waves of growth (e.g., Changeux & Dehaene 1989, p. 83). But even there, illustrations such as their Figure 1 (p. 80), present the model as a discrete, two-phased process.

Is there any other reason why the two-phased interpretation of selectionism would gain favor? We suspect that one answer lies in its correspondence to a popular theory of cognitive development. On this point, **Dehaene-Lambertz & Dehaene** find our conclusion that development marks a reduction in representational complexity far-fetched. This is a popular developmental view, however, one in which very rich initial states are whittled down under the influence of environmental triggers. Mehler's (1985) "knowing by unlearning" is one such example. Reflecting a similar view, Piattelli-Palmarini (1989, p. 12) states, "we seem to receive 'from the hands of nature' (in Hume's expression) much more at birth than we ultimately use in the adult." We find this a surprising criticism from Dehaene-Lambertz & Dehaene since Dehaene himself remarked, "to learn is to eliminate" (Changeux & Dehaene 1989, p. 82).

Although space limits prevent us from examining the selectionist literature in more detail, we think the two-phased developmental view of selectionism is plainly there. What's more, we think the appeal of selectionism lies partly in the idea that development has two phases. It is for this reason that some linguists, such as Lightfoot (1991), cite selectionism as a biological implementation of their position. The linguistic notion that a child is born with a universal grammar that is pared down by parameter setting to a grammar for a specific language is an example of a reduction in representational complexity, one that fits with the popular view of selectionism's two-phased development.

R1.2. The evidence cited for selectionism. As we have suggested so far, the selectionist literature contains many examples of development as a two-phased process. In addition, much of the empirical evidence selectionists marshal for support comes from what we called the "standard model of development" in the target article (see Changeux & Dehaene 1989, p. 82 for numerous citations to this view). Originally based on studies of the neuromuscular junction and ocular dominance column formation, this standard model explicitly divides development into a period of exuberant growth followed by selective pruning. Within the commentaries, **Hurford et al.** argue against our position by pointing out some of examples of early exuberant growth followed by decline as evidence for selectionism, including the neuromuscular junction.

The standard model was bolstered by the influential Rakic et al. (1986) study of synaptogenesis. Even Rakic et al. (1986, p. 234) saw the affinity between their study and selectionism. In the conclusion, they state, "if experience alters synaptic number during development it does so by causing selective survival of certain synapses, not by regulating their initial formation." So too Huttenlocher's (1979) study of human neural development is widely cited to support an initial burst of neural development followed by selective elimination.

Based on the substantial review of the literature we included in the target article and what many of the commentators took to be selectionism's core, we suggest that, far from being a straw man, there are substantial grounds from within selectionists' own writings to attribute a two-phased developmental view to at least the classical statement of selectionism. The selectionist commentaries' newfound emphasis on growth is, we will suggest, an attempt by them to reconstruct selectionism and bring it more into line with new neurobiological evidence. We believe, however, that this move only further weakens their position, as we consider next.

R1.3. Multiple waves of growth/retraction hurt, not help, selectionism. Despite our observation that there are good grounds for attributing a two-phased developmental pro-

https://doi.org/10.1017/S0140525X97301587 Published online by Cambridge University Press

cess to selectionism, we do not deny that selectionism is compatible with multiple waves of growth/retraction. None of the arguments we made against selectionism, however, depend on the two-staged interpretation of selectionism, and so are unaffected by the addition of multiple waves of growth/retraction to selectionism. The reason for this, as we illustrated in Figure 3 of the target article, is that the crucial difference between selectionism and neural constructivism is not whether there is a net increase in representational complexity over development but whether such an increase is regulated by activity-dependent growth. Here, then, is what we see to be the fundamental issue: *If development is a progressive expansion of neural structure, is the process of expansion itself insensitive to environmentally derived patterns of neural activity, or does this activity play a central and specific role in building neural circuits?*

The answer from selectionism is clear enough. Again, to quote Edelman (1987, p. 19), "while the pattern of neural circuitry depends upon evolutionary, developmental, and behavioral variables, it is neither established nor rearranged instructively in response to external influences." Although Changeux and Dehaene (1989, p. 82) allow for waves of growth, in their words, "activity does not create novel connections but, rather, contributes to the elimination of pre-existing ones." The Rakic et al. (1986, p. 234) quotation above is similarly unambiguous in the role of activity.

For selectionism, allowing environmentally derived patterns of neural activity to regulate neural growth, as opposed to merely selecting pre-existing circuits, is more than a mistaken empirical claim. It is a conceptual blunder, a pre-Darwinian confusion, an instructivist fallacy. For this reason, for the selectionist, any theory positing activity-dependent outgrowth is admitting "Lamarckian" processes of directed mutation, as **Reeke** charges. These charges belong to a well worn selectionist tradition against learning. Piattelli-Palmarini (1989, p. 21) goes so far as to suggest that "'transfer of structure' and its equivalents (interiorization, assimilation, etc.) become very close to a *nomological* impossibility."

At this point, the selectionist framework seems an unnecessary burden. Along with Purves et al. (1996), we too fail to see the utility of imposing these ideas on brain development. What happens during an individual's lifetime is by definition "Lamarckian," in that some episodes specific to that individual's life history alter the brain's representations. There is nothing conceptually problematic about the proposal that the neural activity helps construct brain circuits in specific – even directed – ways.

Selectionist commentators had little to say about the learning-theoretic section of the target article (sect. 4). There we explored many well-defined neural network algorithms and their learning properties, which bear a strong relation to the more biological proposals we made in section 3. This work shows the "instructive" role of environmental structure in constructing a learner's representations, demonstrating its plausibility and providing reasons it is a stronger approach to learning in development than selectionism. Indeed, as **Grossberg** points out, selectionist models are typically poor learners, generally leading to nonoptimal solutions in practice. Selectionism cannot simply assert that their systems have these learning characteristics because they allow multiple waves of growth. This form of learning is possible only if activity itself regulates growth, something that, as we saw in the quotations above, selectionists strictly disavow.

We think there is a deeper problem for selectionism in the multiple waves of growth/retraction view. The real problem, and why it appears so problematic to point to what makes a difference between these ideas, is that selectionists appear to recognize the centrality of a progressive expansion of cortical structures but have a dearth of consistent mechanisms to account for the progressive nature of development. The problem is to identify mechanisms that add structure in such a way so as not to be "instructive" or Lamarckian. Hence, as **Reeke** remarks, "undirected or broadly but not specifically directed growth followed by pruning is the necessary principal mechanism of neural selection."

But what mechanisms are identified in **Reeke**'s rather legalistic phrase "undirected or broadly but not specifically directed growth?" Elsewhere in his commentary, **Sporns** states that selectionist mechanisms "includes eliminative selection but range over many other mechanisms." Unfortunately, "other mechanisms" is as specific as he gets. Reeke does get a bit more specific. In dipping into the selectionist obscurity **Purves** alluded to, Reeke talked of positive selection and positive mechanisms.

What are "positive" mechanisms? Edelman (1987) mentions "positive selection." This sounds very much like constructive mechanisms in selective clothing, rather the way government might describe job loss as "negative increases in the workforce." To select is to pick out from pre-existing alternatives. It is not to create novel structure. That is construction, and, when guided by the environment, it is the very thing selectionists disallow. Indeed, we were surprised to see **Reeke** suggest that the algorithms for activity-dependent dendritic growth we explored in section 3 are consistent with selectionism. The algorithms we explored are paradigmatically Hebbian – instructivist mechanisms selectionists ridicule as pre-Darwinian confusions. What happened to the claim that activity does not regulate the formation of synapses, only their elimination – the core notion of selectionism?

We suspect selectionists find that multiple waves of growth/retraction place them on the horns of a dilemma. Given multiple waves of growth and the presence of neural activity in those systems, why would it be surprising if activity played a role in constructing those circuits?

The reasons why activity might play a specific role in regulating neural outgrowth seem clear enough. The brain's metabolic demands are expensive, and in large-brained creatures they put severe pressures on acquiring sufficient food. A reasonable assumption is that the brain regulates the process of outgrowth to minimize unnecessary resource expenditures, as **Black** & **Greenough** suggest. In other words, activity helps the brain put structure where it needs it, not just everywhere. This seems particularly reasonable in the case of adult plasticity. In a study we cited in the target article, Black et al. (1990) reported a 25% increase in synapses after complex learning. Does this mean that a spontaneously generated surfeit of neural circuits – whose initial construction was insensitive to the environment – was selectivity stabilized as a result of activity? According to the selectionist interpretation, the adult brain would be constructing a massive amount of excess structure all on the chance that it might be stabilized. Although this is a possibility, it seems the more reasonable assumption is that

https://doi.org/10.1017/S0140525X97301587 Published online by Cambridge University Press

activity initiates and guides the construction of specific circuits.

At this point, a selectionist might maintain that the addition of neural structure is due to activity-dependent permissive growth, not activity-dependent guided growth. Although this is a significant departure from the categorical claim that activity does not contribute to the formation of synapses, and makes selectionism something much less than a radically new theory, it still marks a distinction between selectionism and neural constructivism. Indeed, it might be what **Reeke** had in mind when he posited "undirected or broadly but not specifically directed growth." The distinction between selectionism and neural constructivism now comes down to whether activity is permissive or whether it plays a more specific role in constructing neural structure.

R1.4. Directed neural growth. In section 3 of the target article we explored systems and mechanisms that regulate neural construction. Regarding some of these, **Reeke** suggests that the action of a diffusible substance "cannot contain the specific information needed to determine the individual fate of each synapse in the region." Now selectionists seem to be creating the straw man. There is an enormous range of possible theories between undirected, permissive growth and single-synapse specificity. The Hebbian conditions we outlined for directed dendritic growth are clearly more than simply permissive ones. They depend on the right pattern of correlated pre- and post-synaptic activity, and so are extracting statistical structure at specific timescales. These are paradigmatically instructivist.

The Kossel et al. (1995) study cited in the target article demonstrates that these sorts of processes are operating in the dendrites of cells in ocular dominance columns. Numerous studies have shown similar processes to operate in axonal growth. **Purves,** a longtime proponent of directed, constructive development, cites other work (Riddle et al. 1992; 1993; Zheng & Purves 1995) supporting a constructivist approach. **Bolhuis** points to research on the avian system supporting a constructivist model.

Elliott & Shadbolt (and, surprisingly, **Sporns**) point to recent work on the role of neurotrophins as another line of research supporting a constructivist approach (McAllister et al. 1995). Elliott & Shadbolt are right that this is an important source of evidence for constructive, directed growth. Particularly intriguing is the more recent finding of McAllister et al. (1996) that neurotrophin signaling is only effective in eliciting dendritic growth when that dendrite is active. The requirement for conjoint activity suggests that neurotrophins may implement a correlational developmental rule.

R1.5 Neural constructivism is consistent with regressive events. Reeke suggests that genuinely directed growth is a conceptual blunder because it reverses cause and effect. That is, directed growth would have to establish perfect connectivity (the effect) prior to any cause (selection). The constructivist model does not deny the existence of selective stabilization (sect. 3.3; Fig. 3). As we state, if cells do not bear unique molecular addresses, then stochastic sampling mechanisms must be posited. We believe the primary reason for regressive events is not that the brain is following an exuberance/retraction strategy; rather, it is due to a sampling problem, which we called the "needle in the haystack problem" in the target article. Hence, although the algorithms we described in the target article are for directed growth at the level of individual dendritic and axonal segments, the process of outgrowth itself involves stochastic events. An outgrowing fiber within a dense network cannot be *a priori* guaranteed of finding the appropriate target. The functional validation, or invalidation, of those synapses alongside directed growth is therefore a necessary component.

Others, such as **Innocenti,** suggest that with regard to the question of regressive events in development, ours was a biased presentation of the literature. Unfortunately, there was not enough space to present a comprehensive account of regressive events. We decided instead to limit ourselves to an evaluation of some of the canonical papers on regressive events in development, such as Rakic et al. (1986) and the work on ocular dominance columns. In our examination of dendritic development, however, we did cite a number of studies, including one by Innocenti and his colleagues (Vercelli et al. 1992), showing that regressive events play a role there.

Kennedy & Dehay rightly point out that the majority of cortical connectivity derives from corticocortical connections, not thalamocortical connectivity. We did examine in some detail the development of local, horizontal connections (Callaway & Katz 1992), which support the progressive view. As Kennedy & Dehay point out, however, there have been surprisingly few developmental studies of corticocortical connectivity, which is why we did not examine this issue in more detail. From what evidence there is, Kennedy & Dehay conclude that it is consistent with neural constructivism. They are also right that early events derived from internally generated activity structure the brain prior to externally evoked activity, but we disagree that this leads to the conclusion that constructive processes only fine tune cortical connections. As we explored in the target article, the postnatal development of the human brain is both more protracted and more extensive than often supposed. Some dendrites, for example, undergo their majority of growth after the second postnatal year (Schade & van Groenigen 1961).

Reeke also suggests that we have simply equated increases in complexity with increases in size. But far from simply equating these, the target article was an attempt to examine the relationship between structural and representational complexity and to evaluate candidate measures. He is right that there are other functional issues to consider. These issues were considered at length in section 3 of the target article.

Scheibel, a pioneer in the study of the environment's role in shaping the brain, points to interesting results that show just how subtle the interaction between structure and function can be. In one instance (e.g., Mollgard et al. 1971), exposure to enriched environments resulted in a decline in synaptic numbers resulted in larger, more effective synapses. Scheibel also points out that in a study of dendritic structure in Broca's area of left and right hemispheres the difference was not in total dendritic length but in the complexity of various portions of the dendritic tree. We attempted to capture this finer scale sensitivity to activity in the locality and stability conditions (sect. 2.3).

R1.6. Black & Greenough's proposed compromise. Black & Greenough suggest a way out of the selectionist (as overproduction/retraction) and constructivist (as guided

https://doi.org/10.1017/S0140525X97301587 Published online by Cambridge University Press

expansion) choice with their distinction between experience-expectant and experience-dependent processes. According to them, experience-expectant processes are ones used to capture information common to all members of a species, such as early visual experience. Experience-dependent processes are used to capture the idiosyncratic experiences of individuals.

According to **Black & Greenough,** selectionist mechanisms underlie experience-expectant processes whereas constructivist mechanisms underlie experience-dependent processes. **Bolhuis** cites work from the avian system that makes this distinction problematic because the processes overlap and are too intertwined to separate categorically (Bolhuis 1994). As the revised view of ocular dominance column formation demonstrates (and as the neo-selectionist view above reinforces), this distinction breaks down even in paradigmatically experience-expectant systems.

Rather than view experience-expectant and experience-dependent processes as two distinct information-storage strategies, each with their own mechanisms, we suspect that they represent a continuum of the same underlying mechanisms. Rather than evolving distinct mechanisms, it is plausible that experience-dependent storage involves changing some parameters in the mechanisms of experience-expectant storage, such as temporal patterns of expression, the duration of sensitive periods, and the rate and extent of postnatal growth.

Black & Greenough state that they will ignore the target article's technical misunderstandings, only to cite our failure to discriminate between synaptic density and synapse number. We cite **Innocenti** in our defense; he states, "as rightly noticed by Q&S, the assessment of synaptic densities in a changing volume does not provide results unequivocally interpretable in terms of connectivity."

Before we leave the issue of what separates neural constructivism from selectionism, we should mention **Purves's** concern that the target article was couching pretty straightforward ideas in overcomplex terminology. We are sympathetic to this, although we were trying to use the terminology that might be familiar to cognitivists. And, as we have just seen, trying to define selectionism is a perilous semantic adventure. In the end, whether there are really genuine differences seems unclear. The crux of our position is that environmentally derived activity helps guide progressive neural growth by regulating axonal and dendritic outgrowth at the level of individual segments.

R2. Is the right measure axons or dendrites?

Some commentators (**Innocenti; Elliott & Shadbolt; Hurford et al.**) suggest that we place too much emphasis on dendritic development. The target article did emphasize dendritic growth, in part because we believed that activity-dependent axonal growth had already been well-established. In contrast, dendritic development has been almost completely ignored, particularly in computational models of development. Contrary to what Elliott & Shadbolt claim, however, we do not dismiss axonal development as irrelevant. However, within a single paper it is impossible to present a complete account, or cite every possible source of evidence. Therefore, we decided to concentrate on dendrites, which, because of their nonlinear properties, could also be considered the brain's basic computational units. We did state, though, 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" (Note 2). In his view that axons and dendrites adapt to each other, Innocenti echoes a similar interactive view.

Finlay's commentary nicely illustrates the multiple problems – target selection, population matching, and activity-dependent dendritic structuring – that must be solved in what she calls the axo-dendritic interface. As she rightly notes, some of these problems are biological, while some are more cognitive (the latter being the primary subject of the target article). As we stressed in the target article, and as Finlay's commentary so nicely illustrates, understanding how the mind emerges from the developing brain will require explanations across these explanatory levels. There is no such thing as a system's "cognitive architecture," a system devoted solely to information-processing. Instead, the brain must solve multiple demands from homeostasis to language acquisition – all at the same time.

Elliott & Shadbolt suggest that there is a logical gap between sections 3 and 4, in which we go from dendritic growth to learning arguments. They suggest that the work we cited on the learning properties of networks that add units is irrelevant since adding a unit is equivalent to adding a new neuron, not elaborating a dendrite. However, section 3 contained a detailed argument concerning why a local dendritic segment should be considered the brain's basic computational unit, not the entire cell. Because of this equivalence, adding a dendritic segment is tantamount to adding a unit in a conventional connectionist network. For this reason, the "critical step" they believe is absent is presented in detail. The constructivist network research we cite thus has a direct bearing on activity-dependent dendritic development.

Székely notes that although we examined the influence of function on dendritic form, dendritic form also helps determine function. In particular, he reports on the morpho-functional property of brain stem motor neurons. In pointing this out, Székely makes explicit an important point that we left implicit in the target article. Activity-dependent dendritic growth is interesting because form and function are intertwined. By partially regulating form by function during development, activity-dependent growth helps build structures with specific functions.

We agree with nearly everything **Foss** discussed in his insightful commentary. His three conditions that must be satisfied by a complete constructivist account are indeed the core ones. In particular, he is right that the issues involved in equating neuronal growth and representational increase are difficult ones. Much of the target article was aimed at trying to establish this correspondance. He also suggests that we leave the issue of whether this growth is directed as a matter of faith. As we mention in section R1.4, there is increasing evidence that strengthens this claim.

R3. Is neural constructivism a return to *tabula rasa* learning?

Despite our indication to the contrary, some commentators (**Dehaene-Lambertz & Dehaene; Feldman**) suggest we have returned to an extreme empiricism. We cautioned against this extreme interpretation of neural constructivism

https://doi.org/10.1017/S0140525X97301587 Published online by Cambridge University Press

in sections 1.1 and 4.3 of the target article. Neural constructivism is not a return to *tabula rasa* or assumption-free learning. Indeed, one of the central lessons we extracted from learning theory was that true assumption-free learning is impossible. Every learning system, whether engineered by hand or developed by nature, makes implicit assumptions about the world; for each architecture some sequences of input patterns are more easily represented or learned. One way to read the target article is as a proposal for escaping two strong and seemingly conflicting constraints on a theory of development. On the one hand, learning theory tells us that assumption-free learning is impossible: on the other hand, recent neurobiological evidence indicates that detailed, domain-specific knowledge cannot be built into the cortex *a priori*.

Neural constructivism suggests that the way out of this dilemma is with guided representation construction, an approach with attractive learning properties. Its starting point is not the blank slate, but a set of intrinsic constraints (in sect. 4.3 we listed generic initial cortical circuitry, conduction velocities, subcortical organization, learning rates, and hierarchical development; see also Table 1 in **Johnson et al.**). These constraints are augmented with environmentally derived neural activity, which helps regulate the progressive expansion of cortex. As a developmental strategy, this avoids the limiting results of classical learning theory and suggests itself as a reasonable foundation for theories of development.

As **Johnson et al.** point out, Elman et al. (1996) adopts a similar research strategy, one that is proving extremely productive. Johnson et al. add a potential challenge to constructivist approaches: the relative consistency of outcome of cortical specialization. In other words, why do our brains come to have similar types of representations in similar regions of cortex? This paradox may be more apparent than real, though, in that we know from studies of self-organization that even in the presence of noise, structures can reliably emerge again and again without a high degree of prior specification (e.g., Goodwin et al. 1993; Kauffman 1992). The sometimes quite striking individual variability of brain organization further suggests that morphogenetic processes of self-organization produce robust but not tightly constrained structures.

Barton suggests that although we deny we are returning to a *tabula rasa* view, we slip back into one. He suggests that our theory cannot be correct because, although the developing cortex may have a high degree of equipotentiality in theory, in practice this does not matter because the cortex is highly constrained by subcortical structure. The notion of equipotentiality, however, does not demand that it be expressed in every instance. Indeed, as a "potentiality" it remains as a capacity in most cases. However, the developing cortex's high degree of equipotentiality is certainly important in cases of early traumatic injury, when it can underlie reorganization.

Barton also suggests that our theory leads to the conclusion that a rat and a mole reared in the same conditions would develop the same basic cortical structures. We were puzzled by this. As we have already indicated, we agree that subcortical organization is an important constraint on cortical development. Indeed, the core of neural constructivism is that structured activity – which is structured by subcortical processing – impinging on the cortex plays a central role in shaping its structural and functional properties.

Kennedy & Dehay point toward evidence that chemical markers distinguish different cortical areas at early stages of development. Further evidence from grafting experiments (Ebrahimi-Gaillard & Roger 1996) supports early differentiation of some subcortical projections. These initial conditions on long-range cortical projections are the types of general architectural constraints that we had in mind as essential in setting up the conditions under which constructive learning takes place.

R4. What happened to modules?

Barton also defends modularity. We suspect that Barton underestimates how strong the nativist modularity commitment is to domain-specific information embedded *a priori* in the recipient cortex. At many points throughout his commentary, Barton appears to support the view that modules are built by constructive processes. He cites Karmiloff-Smith (1992), for example, and nods approvingly at our statement that the cortex is enslaved by the periphery. Karmiloff-Smith's view (1992; see also commentary by **Johnson et al.**) is that modules are developmental outcomes, not developmental prerequisites. Since Barton's comments appear to be in general agreement with the idea that modules emerge as a result of developmental processes, we do not think there is a substantial difference between our views here, although we think strict modularity has problems for other reasons, as we discuss next.

Scholl examines the difference between neural and cognitive modularity. Cognitive modules are defined functionally in terms of such criteria as information encapsulation (Fodor 1983). Although his is a careful analysis, we think Scholl makes too much of this distinction. In particular, we disagree with his functionalist comment that cognitive modularity imposes no constraints on how modules are realized in neural tissue. For example, Scholl recounts the argument that cognitive modules need not be neurally localized. But could a cognitive module be widely distributed across the brain? It depends again on the sense of module. If specialization and information encapsulation are the criteria, then we think the answer is no. In the brain, there is no linearly addressed information store to keep information encapsulated. Instead, multiple patterns of information are distributed among overlapping populations of neural structures. Such patterns will influence one another. In fact, as studies of content addressable memory stores illustrate, it is likely that violation of information encapsulation across many modalities is something the brain makes use of to reconstruct incomplete information.

We also disagree with **Scholl**'s suggestion that too little is known about the neural underpinnings to settle anything regarding cognitive modularity. The modularity thesis owes a lot to Marr's theory of vision (Marr 1982). In this regard, recent neurobiological and neurocomputational work has been decisive in showing that a bottom up theory of vision is no longer tenable (Churchland et al. 1994). Other recent work has shown that auditory and linguistic information interacts at early levels of processing, again contradicting information encapsulation (Saffran et al. 1996).

We were a little puzzled by **Scholl**'s suggestion that to address nativism about cognitive structure, neurobiological evidence would have to demonstrate that bona fide learning was involved. One of the main aims of the target article was to show how processes of learning guide neural develop-

https://doi.org/10.1017/S0140525X97301587 Published online by Cambridge University Press

ment, and how this "constructive learning" gives the developing system special learning properties.

Székely suggests that Broca's area is required for language. Although damage to Broca's area in adults typically results in severe language impairment, children missing the entire left hemisphere can develop normal language skills if the neural damage was early enough to allow for reorganization during development (reviewed in Elman et al. 1996). This striking, large-scale capacity of the developing brain to reorganize is a challenge to theories relying on built-in, domain specific knowledge embedded *a priori* in the cortex.

R5. What does learning theory show?

Blanzieri and **Feldman** address our use of learning theoretic results in support of neural constructivism. As they both rightly point out, the learning results for constructivist neural networks do not primarily address the question of the learning of grammar. Our aim, however, was not to explore work on constructive networks and grammar learning. Instead, since in traditional cognitive science the lessons from grammar acquisition were generalized to the entire mind, we were interested in more general questions regarding learning and development. Although Chomsky's work would have remained highly influential even if it were restricted to more technical aspects of the learnability of grammars, its enormous influence comes from how readily it seemed to generalize to the entire mind. The one point we did want to make was that contrary to general belief, it has not been shown that grammar learning is impossible. We are still far from demonstrating that neural networks could learn syntax under biologically feasible conditions, although some progress is being made, as pointed out by Feldman.

There were a number of different opinions regarding the differences between development and learning in the mature state. **Black & Greenough** advise that the pattern of connections is what matters, and that mathematical models need not distinguish between constructivist and selectionist modes of growth underlying the mature pattern of connections. **Feldman** suggests that we should drop the arbitrary separation of development from learning. **Hurford et al.,** however, criticize us for running these two processes together. Is development different, and if it is, is it important to characterize the learning properties specific to a developing system?

As the target article outlined, the developmental strategies are crucial to understanding the learning properties of cortex. As we considered above (sect. R.1.3), selectionism has a number of problematic learning properties. In contrast, as **Haith** explored in some detail, tying hierarchical representation to activity-driven, expanding computational resources gives the constructivist learner important advantages.

Sporns suggests that we needed to define or quantify complexity in relation to either anatomy or cognitive representations. He admits that this is a hard task, but one that would have put neural constructivism to the test. We avoided introducing technical measures into what was already a theory-laden target article. However, Quartz (1993) has explored a measure known as the Vapnik-Chervonenkis (VC) dimension, which is used in learning theory to quantify the complexity of a class of representations (see Abu-Mostafa 1989); it could also be used to quantify the relation between increases in anatomy and the complexity of cognitive representations.

Raijmakers presents an interesting proposal for refuting a classical learning paradox. Fodor (1980) argued that there is no way for a system to increase its representational power as a function of learning, thus casting doubt on the plausibility of constructivism (see Quartz 1993 for discussion). Quartz (1993) suggested that there is a natural sense (and an appropriate measure) in which constructive neural networks increase their representational power as a function of learning. Raijmakers suggests that this constructivist learning does not refute Fodor's argument because it adds units, and so is a qualitative as opposed to a quantitative change. As an alternative, Raijmakers proposes that a network's dynamical properties, such as a bifurcation or phase transition, could result in increases in representational power without the need for discrete addition of structure. Raijmakers gives the example of Exact ART, a neural network architecture that is related to the constructivist networks we examined (Raijmakers et al. 1996a). However, as Raijmakers notes, the phase transitions in his example are not a function of learning. What is crucial to answer Fodor's challenge is the requirement that increases in representational power be a function of learning rather than that the increase in representational power not be a discrete change to the architecture (such as the addition of new units). It is entirely possible that the phase transitions Raijmakers has in mind are the result of learning. In such a case, however, we think that would represent an alternative response to Fodor's challenge, one that stands beside other responses, such as the neural constructivist one.

R6. What implications does neural constructivism have for nativist linguistics?

According to **Bickerton,** we are the latest in a long line of "C-men," coming on the heels of cognitive linguistics and connectionism, to launch a failed attack on nativist linguistics. Just as we suggested that selectionists have reconstructed their position, so too Bickerton has reconstructed nativist linguistics. Is it the case, as he suggests, that no serious nativist would deny that a vast amount of learning must occur if hard-wired mechanisms are to function effectively? Kenneth Wexler (1996, p. 118) is one serious nativist who thinks otherwise:

> The idea of genetically programmed maturation is so strong in the study of biology that a special term has been defined for exceptions. This term is "plasticity." Plasticity means that there is experience-dependent variation in biological structures or processes. It is considered a major discovery in the study of the brain in neuroscience, for example, when it is demonstrated that a certain process is plastic. The reason this is considered a major discovery is because the general view is one of a biological, genetically based program guiding development.

From the selectionist-linguistics perspective, Piattelli-Palmarini (1989) suggests that the term "learning" no longer has any role to play in cognitive explanation. **Bickerton** also suggests that we have arranged a shotgun marriage between nativist linguistics and selectionism. It is true that they are wholly dissociable. However, nativist linguists (e.g., Lightfoot 1991) interested in neural plausibility look to selectionism because there are no other plausible developmental theories in neurobiology that are consistent with nativism's extreme developmental demands. Wexler's enthusiasm for

https://doi.org/10.1017/S0140525X97301587 Published online by Cambridge University Press

genetically based programs notwithstanding, as we indicated in the target article, there are no available candidate biological theories consistent with genetically programmed maturation.

Bickerton suggests that axonal development does not follow a constructivist theme, but he misrepresents the conclusions we made with regard to axonal development. In the model system – ocular dominance column formation – the evidence we reviewed demonstrated a substantial overall axonal expansion. As **Purves** remarks, the brain gets bigger. Bickerton also suggests that we have overemphasized the influence of Gold's (1967) work on language learnability. Although Bickerton may be right that few linguists understand Gold's paper, the belief that there are "mathematical" results against learning is a monumentally important and widespread one, not just for language, but for the entire mind. As Fodor (1975) put it, nativism was the only game in town.

Bickerton suggests that to have effected a breach the target article would have had to explain seven enormous outstanding research problems, ranging from how creole languages are acquired to how the first human language was acquired. That would have required something substantially longer than what was already a long target article.

Does nativist linguistics really offer satisfying explanations for these problems, as **Bickerton** claimed? It suggests that Williams syndrome subjects (item [c] on his list) have perfect syntax because their "syntax module" is spared. Apes, on the other hand, lack such a module, and therefore cannot acquire language. Chomsky's explanation of how the first human language was acquired is that a language organ suddenly appeared, probably through some process other than natural selection. To our mind, theories like Greenfield's (1991) that stress continuity between language abilities and the hierarchical construction abilities of nonhuman primates are more plausible than positing a novel module, particularly since the human brain contains no new cytoarchitectural regions. As we suggested in the target article, claiming that these abilities are the product of hard-wired modules does not explain them. It only pushes the question back to how such modules could ever develop. Asking this developmental question in the light of current neurobiology, we suggest, reveals that the nativist explanation is no explanation at all.

In fact, **Bickerton** suggests that outside language the evidence for modularity is much weaker. He recommends that language be kept as the only innate module. But if modularity fails for other domains because our expanding knowledge has made it an outmoded explanation, isn't there reason to believe that as we learn more about the neural basis of language, modules will disappear there as well? Perhaps the strongest reason against positing modules for language is a pragmatic one. Substituting faith for knowledge does not lead to a productive research strategy.

After criticizing us for suggesting that selectionist learning is a reduction in initial capacity, **Dehaene-Lambertz & Dehaene** point to linguistic evidence suggestive of environment-dependent loss of initial abilities. Among the findings they cite are those of Patricia Kuhl and her colleagues (Kuhl et al. 1992) on phonetic perception. Infants can perceive non-native contrasts, as can non-humans. Around six months of age, however, children lose this ability and come to perceive only contrasts from their native language. Although there is a clear loss of perceptual ability, loss is not the whole story. Infants form vowel prototypes, or exemplars that influence their speech perception. In particular, in what Kuhl calls the "perceptual magnet effect," a native prototype will "pull" a non-native contrast toward it, with the result that it is perceived as the native one (Iverson & Kuhl 1995). This suggests that the representations of native categories have reorganized and expanded in response to environmental information.

R7. Is the environment richly structured?

Nativists and selectionists agree on one important point: the environment is informationally impoverished. Chomsky's "poverty of the stimulus" arguments are well known (Chomsky 1988). Selectionists (e.g., Edelman 1987; 1993) provide similar arguments, arguing that the lack of external information means that such information cannot be transferred via instruction from the world into the brain. For this reason, Edelman (1987; 1993) dismisses the information processing framework of mainstream cognitive science.

Bickerton suggests that we are raising the ghost of Gibsonian affordances. Like selectionism, J. J. Gibson's (1979) "ecological perception" view dismissed information-processing views, much to the chagrin of mainstream computationalists (Fodor & Pylyshyn 1981; Ullman 1980). Whereas selectionists put all structure inside the head in terms of a rich set of internally generated prerepresentations, Gibson kept it all out in the environment. Ironically, selectionists use the same term as Gibsonians, "resonance," to refer to the simple "registering" relation that goes on between environmental events and internal representations.

We believe that traditional computationalists were too quick to dismiss environmental structure. Poverty of the stimulus arguments depended on prior intuitions about learning. That is, given very weak learning mechanisms (as in Gold's theory), it was not surprising that the environment would be seen as impoverished. More powerful learners, such as neural networks, however, can discover far more structure in the environment, as is now being recognized in statistical studies of sentences and network models of language acquisition (Bates & Elman 1996; Marcus 1995; Seidenberg 1997). The poverty may not have been in the stimulus but in the minds of nativists.

Feldman makes an important point that conceptual learning, particularly through interactions between visual and auditory representations and interaction with the environment, may have a decisive role in language acquisition. Although we will not go into detail here, poverty of the stimulus arguments also obscured the rich structure that a culture holds. Indeed, we believe the gradual transfer of part of the brain's developmental program into a world structured with culture, through a process we call "progressive externalization," is crucially important for our cognitive abilities.

In addition, we are discovering that many instances of visually guided behavior do not require a complex reconstruction of the visual scene. Instead, input from early as well as late stages of visual processing in cortex is extensively routed to subcortical structures to initiate behavioral responses with short latencies (Churchland et al. 1994). This is a different type of computational architecture than that adopted by many researchers in artificial intelligence (including **Feldman**) who assume that planning and action

https://doi.org/10.1017/S0140525X97301587 Published online by Cambridge University Press

selection must be made at higher levels of processing. Feldman offers a scenario in which early development yields "a complete, but primitive, brain – possibly all subcortical." According to him, this primitive brain could then recruit cortical regions. The idea of a completed subcortical brain, though, conflicts with basic facts of developmental neurobiology. It also reflects the outdated idea that subcortical structures are phylogenetically older than cortical ones.

R8. Where's the cognitive in development?

A number of commentators were right to point out that the target article weighed heavily on the neural and less so on the cognitive side of development (**Mareschal & Schultz; Estes & Bartsch**). Our main goal in the target article was to explore the neural basis of cognitive development and to relate this to the learning properties of neural constructivism and alternative theories. We think this provides a solid foundation for subsequent research into cognitive development.

Many of the commentaries examine cognitive research that complements the aims of the target article (**Mareschal & Schultz; McCollum**). Mareschal & Schultz, for example, present an interesting series of constructive network simulations on various developmental tasks (e.g., the balance-scale task). It is interesting that the simulations with constructive networks capture the developmental profiles of children better than static feed forward connectionist networks do. This is an important advance in the study of the dynamics of development. A recent book of computer exercises from Plunkett & Elman (1997) should prove valuable in making these research tools more accessible.

Estes & Bartsch suggest that our methodology is backwards. They recommend that description must precede explanation. That is, a clear behavioral description of what happens to the developing child should be the methodological touchstone of development inquiries. Behavioral research is important, but descriptions of development do not take place in a conceptual vacuum. Developmental psychology, like any other scientific endeavor, does not operate without guiding theoretical commitments. Those with nativist inclinations, for example, tend to concentrate on the neonate's capacities rather than on what emerges across development.

The aim of the target article was to make explicit the neural constraints on development and so help define the research questions that can guide behavioral studies. We agree that neither approach has absolute methodical priority over the other. In vision, a long and not very profitable debate ensued after Marr (1982) advocated a top-down approach. Researchers in that tradition attempted to capture the abstract principles of visual processing without using the most obvious constraint on theory – the neural architecture that implements those processes. Others, such as neuroethologists Heiligenberg (1991), working in the electrical fish, and Konishi (1991), studying the barn owl, showed the richness of neural constraints and how dramatically a knowledge of them could transform our understanding of the problems organisms must solve in real environments.

McCollum suggests that the development of bodily-kinesthetic intelligence (walking is her example) involves processes akin to constructive learning, or what she calls second-order learning. She points out that motor development, like cognitive development, typically involves discontinuities, a property we called nonstationarity in the target article. Later stages are not simply refinements of what came before; they involve large-scale reorganization. Although we did not pursue the connection between motor learning and cognitive development in the target article, McCollum's suggestion points to a natural affinity between these processes, one that is not too surprising considering the brain's propensity to conserve basic mechanisms. It also ties in nicely with the view of perception as an actively driven process rather than a passive one (Churchland et al. 1994).

Johnson et al. are right that in constructivism the world does not simply impress itself upon a passive infant. Rather, the infant actively constructs its knowledge by engaging the world and structuring its experiences. They are also right to suggest that we need to match *neural* constructivism with a mechanism for *behavioral* constructivism in the infant. Although we did not discuss it in the target article because of space limitations, we think characterizing the child's contribution to development is essential.

We are sympathetic with **Estes & Bartsch**'s appeal that the behavioral not be lost in the neural. These two sorts of explanation have remained isolated by an explanatory gap, one that we think can be narrowed by computational approaches. As **Mareschal & Schultz** suggest, this is the real promise of neural computational approaches, one that admittedly is just beginning but carries great promise.

Foss also provides an example of rapid learning, a teacher asking a student for an answer, and rightly points out that our paper says very little about its neural underpinnings. He also points out that constructivist learning is slow, while this is an instance of rapid learning. What we call learning in this paper could also be called "representation construction." In contrast to rapid learning, representation construction is slow, as it builds the rich set of representations that are then employed in rapid learning. This process of representation construction followed by rapid learning is evident in the learning curves children display when they acquire various elements of language, for example. Rapid learning is most likely performed by different mechanisms than constructivist ones, likely involving changes in synaptic efficacy rather than representation construction. But Foss is right that we are only beginning to understand how these mechanisms underlie the complex forms of learning he mentions.

R9. Is evolution a progressive increase in representational flexibility?

Because of space limits we mentioned the evolutionary implications of neural constructivism only briefly in the target article. Our suggestion that mammalian evolution should be viewed as a progressive increase in representational flexibility sparked a number of dissenting voices (**Barton; Black & Greenough; Hurford et al.; Purves**).

We did not mean to imply the linear *scala naturae* of global processing power ordering all species, as **Barton** suggests. Rather, our comments were in reaction to evolutionary psychology, which views the human cortex as a collection of specialized, dedicated computers. In Tooby and Cosmides's words (1995, p. 1189):

https://doi.org/10.1017/S0140525X97301587 Published online by Cambridge University Press

The human cognitive architecture is far more likely to resemble a confederation of hundreds or thousands of functionally dedicated computers, designed to solve problems endemic to the Pleistocene, than it is to resemble a single general-purpose computer.

The learning-theoretic arguments we presented were aimed at showing why a learner that allows the specific structure of a problem domain to help build mental representations has a number of advantages over building solutions into the cortex *a priori.* In the case of dynamic environments, which in the case of humans is undeniable, the flexibility that constructive learning gives a developing brain means that it will be capable of building appropriate representations in a much wider range of environments. Guessing *a priori* by building in the answer may be faster, but it leaves the guesser out of luck when the environment rapidly changes.

Is human development fast, on a timescale that reflects the unfolding of dedicated circuits – instincts? Issues of metabolic conservation and other resource constraints only go part of the way toward explaining why human development takes so long. According to neural constructivism, a reason human development is so slow is that it maximizes the opportunity of the environment to help shape and build the neural structures underlying thought.

As **Hurford et al.** note, choosing representational flexibility is itself a response to evolutionary pressures. We fully agree with Hurford et al. that evolutionary psychologists take up the challenge to find how environmental structure can influence the evolution of the constructivist learning mechanism itself. But this is a very different research program from one identifying hard-wired circuits.

Both **Hurford et al.** and **Sporns** took exception to our remark that selectionism is incompatible with the claim that evolutionary pressures have picked out specialized circuits. Hurford et al. counter that evolution could pick out the initial set of neural circuits which would then be operated on by selectionist processes. This is an entirely consistent possibility, and one compatible with the claim we were making. Our claim was made in the context of Gazzaniga's (1992) suggestion that the specialized circuits evolutionary psychology posits are compatible with selectionism. Edelman (1987), for example, rightly distinguishes his position from Chomsky's, which he criticizes for its claim that evolution builds in point-to-point mappings in the brain. For the same reason, Piattelli-Palmarini (1989) reacted strongly against Pinker and Bloom's (1990) *BBS* target article arguing that language is the product of natural selection.

R10. Conclusion

During the long hiatus between the completion of the target article and the completion of this response, neuroscience has marched on. Exciting new evidence favoring neural constructivism has emerged that challenges the dogma that new neurons are not created in adult brain. Olfactory receptors and certain neurons in bird brains were known exceptions. Kempermann et al. (1997) have now shown that new neurons are added to the adult mouse hippocampus, which receives inputs from the cortex and is essential for some forms of long-term memory. More of these new neurons survive in mice exposed to an enriched environment, and these mice have better learning and memory compared to mice kept in an impoverished environment. It is likely that new neurons are also formed in adult human brains, perhaps even in the cortex. The differentiation of progenitor cells into neurons is therefore not confined to early development. The implications of these new results for lifelong learning are far reaching.

Our goal in writing the target article was to provide a conceptual framework for integrating the recent advances in neural development with parallel advances in cognitive development and neural computation. Molecular genetics provides powerful tools for studying the mechanisms underlying development at the molecular and cellular levels, but to achieve an integrated view of development, these mechanisms need to be linked with system level descriptions of the nervous system, and ultimately to behavior. This is an exciting enterprise that is just beginning. The authors of the commentaries have given us a glimpse of the diversity of opinion that exists, and we were pleased to have this opportunity to reply to them. We hope that some issues at least have been clarified, if not settled, through this process.

NOTE

1. By "traditional cognitive science" we mean the functionalist based account stemming from Chomsky.

References

Letters *a* and *r* appearing before authors' initials refer to target article and response respectively.

Abu-Mostafa, Y. (1989) The Vapnik-Chervonenkis dimension: Information versus complexity in learning. *Neural Computation* 1:513–25. [rSRQ]

Allman, J. (1985) Evolution of neocortex. In: *Cerebral cortex,* vol. 8A: *Comparative structure and evolution of cerebral cortex,* ed. E. G. Jones & A. Peters. Plenum Press. [DP]

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

Antell, S. E. & Keating, D. P. (1983) Perception of numerical invariance in neonates. *Child Development* 54:695–701. [GD-L]

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

Arimatsu, Y., Miyamoto, M., Nihonmatsu, I., Hirata, K., Uratani, Y., Hatanaka, Y. & Takiguchi-Hayashi, K. (1992) Early regional specification for a molecular neuronal phenotype in the rat cortex. *Proceedings of the National Academy of Sciences USA* 89(19):8879–83. [HK]

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

Bailey, C. H. & Chen, M. (1988a) Long-term sensitization in Aplysia increases the number of presynaptic contacts onto the identified gill motor neuron L7. *Proceedings of the National Academy of Sciences USA* 85:9356–59. [aSRQ]

(1988b) Morphological basis of short-term habituation in Aplysia. *Journal of Neuroscience* 8:2452–59. [aSRQ]

Bailey, C. H. & Kandel, E. R. (1993) Structural changes accompanying memory storage. *Annual Reviews in Physiology* 55:397–426. [OS]

Balazs, R., Hack, N., Jorgensen, O. S. & Cotman, C. W. (1989) N-methyl-D-aspartate promotes the survival of cerebellar granule cells: Pharmacological characterization. *Neuroscience Letters* 101:241–46. [aSRQ]

Baldwin, J. M. (1896) A new factor in evolution. *American Naturalist* 30:441–51. [JH]

Barbe, M. F. & Levitt, P. (1991) The early commitment of fetal neurons to the limbic cortex. *Journal of Neuroscience* 11(2):519–33. [HK]

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

https://doi.org/10.1017/S0140525X97301587 Published online by Cambridge University Press

Barlow, H. B. (1988) Neuroscience: A new era? *Nature* 331:571. [DP]
Barone, P., Dehay, C., Berland, M., Bullier, J. & Kennedy, H. (1995) Developmental remodeling of primate visual cortical pathways. *Cerebral Cortex* 1:22–38. [HK]
Barone, P., Dehay, C., Berland, M. & Kennedy, H. (1996) Role of directed growth and target selection in the formation of cortical pathways: Prenatal development of the projection of area V2 to area V4 in the monkey. *Journal of Comparative Neurology* 374:1–20. [HK]
Barron, A. R. (1994) Approximation and estimation bounds for artificial neural networks. *Machine Learning* 14:115–33. [aSRQ, JAF]
Barton, R. A. (submitted) Visual specialisation, ecology and brain size in primates. [RAB]
Bartsch, K. & Wellman, H. M. (1995) *Children talk about the mind*. Oxford University Press. [DE]
Bates, E. A. & Elman, J. L. (1993) Connectionism and the study of change. In: *Brain development and cognition: A reader*, ed. M. H. Johnson. Blackwell. [aSRQ]
(1996) Learning rediscovered [comment]. *Science* 274:1849–50. [rSRQ]
Bates, E. A. & MacWhinney, B. (1987) Competition, variation, and language learning. In: *Mechanisms of language aquisition*, ed. Brian MacWhinney. Lawrence Erlbaum. [aSRQ]
Bateson, P. & Horn, G. (1994) Imprinting and recognition memory: A neural net model. *Animal Behaviour* 48:695–715. [JJB]
Baum, E. B. (1988) Complete representations for learning from examples. In: *Complexity in information theory*, ed. Y. Abu-Mostafa. Springer-Verlag. [aSRQ]
(1989) A proposal for more powerful learning algorithms. *Neural Computation* 1:201–07. [aSRQ]
Becker, L. E., Armstrong, D. L., Chan, F. & Wood, M. M. (1984) Dendritic development in human occipital cortical neurons. *Developmental Brain Research* 13:117–24. [aSRQ]
Bennett, M. R. & Pettigrew, A. G. (1974) The formation of synapses in striated muscle during development. *Journal of Physiology (London)* 241:515–45. [aSRQ]
Bernardo, L. S., Masukawa, L. M. & Prince, D. A. (1982) Electrophysiology of isolated hippocampal pyramidal dendrites. *Journal of Neuroscience* 2:1614–22. [aSRQ]
Bertoncini, J., Bijeljac-Babic, R., Blumstein, S. & Mehler, J. (1987) Discrimination in neonates of very short CVs. *Journal of the Acoustical Society of America* 82:31–37. [GD-L]
Bertoncini, J., Morais, J., Bijeljac-Babic, R., MacAdams, S., Peretz, I. & Mehler, J. (1989) Dichotic perception and laterality in neonates. *Brain and Cognition* 37:591–605. [GD-L]
Bijeljac-Babic, R., Bertoncini, J. & Mehler, J. (1991) How do four-day-old infants categorize multisyllabic utterances? *Developmental Psychology* 29:711–21. [GD-L]
Birinyi, A., Antal, M., Wolf, E. & Székely, G. (1992) The extent of the dendritic tree and the number of synapses in the frog motoneuron. *European Journal of Neurosciences* 4:1003–12. [GS]
Black, J. E. & Greenough, W. T. (1986) Induction of pattern in neural structure by experience: Implications for cognitive development. In: *Advances in developmental psychology*, vol. 4, ed. M. E. Lamb, A. L. Brown & B. Rogoff. Lawrence Erlbaum. [JEB]
Black, J. E., Isaacs, K. R., Anderson B. J., Alcantara, A. A. & Greenough, W. T. (1990) Learning causes synaptogenesis, whereas motor activity causes angiogenesis, in cerebellar cortex of adult rats. *Proceedings of the National Academy of Sciences USA* 87:5568–72. [arSRQ]
Blanzieri, E. & Katenkamp, P. (1996) Learning radial basis function networks on-line. *Proceedings of the 13th International Conference on Machine Learning*. [EB]
Blum, A. & Rivest, R. L. (1988) Training a 3-node neural network is NP-complete. In: *Advances in neural information processing systems*, ed. D. S. Touretzky. Morgan Kaufmann. [aSRQ]
Blumer, A., Ehrenfeucht, A., Haussler, D. & Warmuth, M. (1988) Learnability and the Vapnik–Chervonenkis dimension. UCSC-CRL-87–20. [aSRQ]
Bohme, G. A., Bon, C., Stutzmann, J. M., Doble, A. & Blanchard, J. C. (1991) Possible involvement of nitric oxide in long-term potentiation. *European Journal of Pharmacology* 199:379–81. [aSRQ]
Bolhuis, J. J. (1994) Neurobiological analyses of behavioural mechanisms in development. In: *Causal mechanisms of behavioural development*, ed. J. A. Hogan & J. J. Bolhuis. Cambridge University Press. [rSRQ, JJB]
Boothe, R. G., Greenough, W. T., Lund, J. S. & Wrege, K. (1979) A quantitative investigation of spine and dendrite development of neurons in visual cortex (area 17) of Macaca nemestrina monkeys. *Journal of Comparative Neurology* 186:473–89. [aSRQ]
Borer, H. (1984) *Parametric syntax*. Foris. [DB]
Borges, S. & Berry, M. (1978) The effects of dark-rearing on the development of the visual cortex of the cat. *Brain Research* 180:277–300. [aSRQ]
Bourgeois, J. P., Goldman-Rakic, P. S. & Rakic, P. (1994) Synaptogenesis in the prefrontal cortex of rhesus monkeys. *Cerebral Cortex* 4:78–96. [aSRQ]
Bourgeois, J. P., Jastreboff, P. J. & Rakic, P. (1989) Synaptogenesis in visual cortex of normal and preterm monkeys: Evidence for intrinsic regulation of synaptic overproduction. *Proceedings of the National Academy of Sciences USA* 86:4297–301. [aSRQ]
Brenière, Y., Bril, B. & Fontaine, R. (1989) Analysis of the transition from upright stance to steady locomotion in children with under 200 days of autonomous walking. *Journal of Motor Behavior* 21:20–37. [GM]
Brewer, G. J. & Cotman, C. W. (1989) NMDA receptor regulation of neuronal morphology in cultured hippocampal neurons. *Neuroscience Letters* 99:268–73. [aSRQ]
Brown, R. (1973) *A first language: The early stages*. Harvard University Press. [aSRQ]
Buckingham, D. & Shultz, T. R. (1996) Computational power and realistic cognitive development. *Proceedings of the Eighteenth Annual Conference of the Cognitive Science Society* 507–11. LEA. [DM]
Buell, S. J. & Coleman, P. D. (1981) Quantitative evidence for selective dendritic growth in normal human aging but not in senile dementia. *Brain Research* 214:23–41. [aSRQ]
Bulloch, A. G. & Hauser, G. C. (1990) Sprouting by isolated Helisoma neurons: Enhancement by glutamate. *International Journal of Developmental Neuroscience* 8:391–98. [aSRQ]
Callaway, E. M. & Katz, L. C. (1990) Emergence and refinement of clustered horizontal connections in cat striate cortex. *Journal of Neuroscience* 10:1134–53. [aSRQ]
(1991) Effects of binocular deprivation on the development of clustered horizontal connections in cat striate cortex. *Proceedings of the National Academy of Sciences USA* 88:745–49. [aSRQ]
(1992) Development of axonal arbors of layer 4 spiny neurons in cat striate cortex. *Journal of Neuroscience* 12:570–82. [arSRQ]
Camel, J. E., Withers, G. S. & Greenough, W. T. (1986) Persistence of visual cortex dendritic alterations induced by postweaning exposure to a "superenriched" environment in rats. *Behavioral Neuroscience* 100:810–13. [aSRQ]
Carey, S. (1990) On some relations between the description and the explanation of developmental change. In: *Causes of development*, ed. G. Butterworth & P. Bryant. Lawrence Erlbaum. [DE]
Carpenter, G. A. & Grossberg, S. (1991) *Pattern recognition by self-organizing neural networks*. MIT Press. [SG]
(1993) Normal and amnesic learning, recognition, and memory by a neural model of cortico-hippocampal interactions. *Trends in Neurosciences* 16:131–37. [SG]
Carrasco, R. & Oncina, J., eds. (1994) *Grammatical inference and applications. Lecture notes in artificial intelligence*, vol. 862. Springer Verlag. [JAF]
Case, R. (1985) *Intellectual development: Birth to adulthood*. Academic Press. [DM]
(1992) The role of frontal lobes in the regulation of human development. *Brain and Cognition* 20:51–73. [DM]
Caviness, V. S. & Rakic, P. (1978) Mechanisms of cortical development: A view from mutations in mice. In: *Annual review of neuroscience*, ed. W. M. Cowan, Z. W. Hall & E. R. Kandel. Annual Reviews. [aSRQ]
Chang, F. L. & Greenough, W. T. (1984) Transient and enduring morphological correlates of synaptic efficacy change in the rat hippocampal slice. *Brain Research* 309:35–46. [aSRQ]
Chang, P. L., Isaacs, K. R. & Greenough, W. T. (1991) Synapse formation occurs in association with the induction of long-term potentiation in two-year-old rat hippocampus in vitro. *Neurobiology of Aging* 12:517–22. [aSRQ]
Changeux, J. P. (1985) *Neuronal man*. Pantheon Books. [GD-L]
(1997) Neural Lamarckism. *Trends in Neurosciences* 20:291–92. [OS]
Changeux, J. P. & Danchin, A. (1976) Selective stabilisation of developing synapses as a mechanism for the specification of neuronal networks. *Nature* 264:705–12. [aSRQ, GMI]
Changeux. J. P. & Dehaene, S. (1989) Neuronal models of cognitive functions. *Cognition* 33:63–109. [arSRQ, GD-L]
Chomsky, N. (1965) *Aspects of the theory of syntax*. MIT Press. [aSRQ, DB]
(1980) Rules and representations. *Behavioral and Brain Sciences* 3:1–61. [aSRQ]
(1988) *Language and problems of knowledge: The Managua Lectures*. MIT Press [rSRQ]
Churchland, P. S., Ramachandran, V. S. & Sejnowski, T. J. (1994) A critique of pure vision. In: *Large-scale neuronal theories of the brain*, ed. C. Koch and J. Davis. MIT Press. [rSRQ]
Churchland, P. S. & Sejnowski, T. J. (1988) Perspectives on cognitive neuroscience. *Science* 242:741–45. [aSRQ]
(1992) *The computational brain*. MIT Press. [aSRQ]
Clarke, P. G. H. (1994) Neuronal death in the development of the vertebrate central nervous system. *The Neurosciences* 6:291–97. [GMI]

https://doi.org/10.1017/S0140525X97301587 Published online by Cambridge University Press

Cline, H. T. (1991) Activity-dependent plasticity in the visual systems of frogs and fish. *Trends in Neurosciences* 14:104–11. [aSRQ]

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

Coggeshall, R. E. & Lekan, H. A. (1996) Methods for determining numbers of cells and synapses: A case for more uniform standards of reviews. *Journal of Comparative Neurology* 364:6–15. [aSRQ]

Cohen-Cory, S. & Fraser, S. E. (1995) Effects of brain-derived neurotrophic factor on optic axon branching and remodelling *in vivo*. *Nature* 378:192–96. [TE, OS]

Cohen-Tannoudji, M., Babinet, C. & Wassef, M. (1994) Early determination of a mouse somatosensory cortex marker. *Nature* 368:460–63. [HK]

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

Conlee, J. W. & Parks, T. N. (1983) Late appearance and deprivation-sensitive growth of permanent dendrites in the avian cochlear nucleus (Nucleus Magnocellularis). *Journal of Comparative Neurology* 217:216–26. [aSRQ]

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

Cragg, B. G. (1975) The development of synapses in kitten visual cortex during visual deprivation. *Experimental Neurology* 46:445–51. [aSRQ]

Cramer, K. S. & Sur, M. (1995) Activity-dependent remodeling of connections in the mammalian visual system. *Current Opinion in Neurobiology* 5:106–11. [HK]

Crick, F. (1989) Neural Edelmanism. *Trends in Neurosciences* 7:240–47. [rSRQ, DP]

Crutchfield, J. P. (1994) The calculi of emergence: Computation, dynamics and induction. *Physica* D 75:11–54. [MEJR]

Curtiss, S. (1989) Abnormal language acquisition and the modularity of language. In: *Linguistics: The Cambridge survey*, vol. 2, ed. F. J. Newmeyer. Cambridge University Press. [DB]

Cybenko, G. (1989) Approximation by superpositions of a sigmoid function. *Mathematics of Control, Signals, and Systems* 2:303–14. [aSRQ]

Dailey, M. E. & Smith, S. J. (1996) The dynamics of dendritic structure in developing hippocampal slices. *Journal of Neuroscience* 16:2983–94. [aSRQ]

Dehaene, S. & Changeux, J. P. (1993) Development of elementary numerical abilities: A neuronal model. *Journal of Cognitive Neuroscience* 5:390–407. [GD-L]

Dehaene-Lambertz, G. & Dehaene, S. (1994) Speed and cerebral correlates of syllable discrimination in infants. *Nature* 370:292–95. [GD-L]

Dehay, C., Giroud, P., Berland, M., Killackey, H. & Kennedy, H. (1996a) The contribution of thalamic input to the specification of cytoarchitectonic cortical fields in the primate: Effects of bilateral enucleation in the foetal monkey on the boundaries and dimensions of striate and extrastriate cortex. *Journal of Comparative Neurology* 367:70–89. [HK]

(1996b) Phenotypic characterisation of respecified visual cortex subsequent to prenatal enucleation in the monkey: Development of acetylcholinesterase and cytochrome oxydase patterns. *Journal of Comparative Neurology* 376:386–402. [HK]

Dekaban, A. S. & Sadowsky, D. (1978) Changes in brain weights during the span of human life: Relation of brain weights to body heights and body weights. *Annals of Neurology* 4:345–56. [aSRQ]

Diamond, A. (1988) Abilities and neural mechanisms underlying A*B* performance. *Child Development* 59:523–27. [GD-L]

Diamond, M. C. (1988) *Enriching heredity*. The Free Press. [ABS]

Diamond, M. C., Johnson, R. E. & Ingham, C. A. (1975) Morphological changes in the young, adult and aging rat cerebral cortex, hippocampus and diencephalon. *Behavioral Biology* 14:163–74. [ABS]

Diamond, M. C., Krech, D. & Rosenzweig, M. (1964) The effect of an enriched environment on the histology of the rat cerebral cortex. *Journal of Comparative Neurology* 123:111–20. [ABS]

Diederich, J., ed. (1990) *Artificial neural networks: Concept learning*. IEEE Computer Society Press. [JAF]

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

Durack, J. C. & Katz, L. C. (1996) Development of horizontal projections in layer 2–3 of ferret visual cortex. *Cerebral Cortex* 6:178–83. [aSRQ]

Durbin, R. & Mitchison, G. J. (1990) A dimension reduction framework for understanding cortical maps. *Nature* 343:644–47. [aSRQ]

Durbin, R. & Rumelhart, D.E. (1989) Product units: A computationally powerful and biologically powerful extension to backpropagation networks. *Neural Computation* 1:133–42. [aSRQ]

Easter, S. S., Jr., Purves, D., Rakic, P. & Spitzer, N. C. (1985) The changing view of neural specificity. *Science* 230:507–11. [aSRQ]

Ebrahimi-Gaillard, A. & Roger, M. (1996) Development of spinal cord projections from neocortical transplants heterotopically placed in the neocortex of newborn hosts is highly dependent on the embryonic locus of the origin of the graft. *Journal of Comparative Neurology* 365:129–40. [rSRQ, HK]

Edelman, G. M. (1978) Group selection and phasic reentrant signaling: A theory of higher brain function. In: *The mindful brain: Cortical organization and the group-selective theory of higher brain function*, ed. G. M. Edelman & V. B. Mountcastle. MIT Press. [GNR]

(1987) *Neural Darwinism: The theory of neuronal group selection*. Basic Books. [arSRQ, SG, GD-L, GMI, GNR, OS]

(1993) Neural Darwinism: Selection and reentrant signaling in higher brain function. *Neuron* 10:115–25. [rSRQ, OS]

Edelman, G. M. & Reeke, G. N. Jr. (1982) Selective networks capable of representative transformations, limited generalizations, and associative memory. *Proceedings of the National Academy of Sciences USA* 79:2091–95. [GNR]

Edelman, G. M., Reeke, G. N., Gall, W. E., Tononi, G., Williams, D. & Sporns, O. (1992) Synthetic neural modeling applied to a real-world artifact. *Proceedings of the National Academy of Sciences USA* 89:7267–71. [OS]

Eimas, P. D., Siqueland, E. R., Jusczyk, P. W. & Vigorito, J. (1971) Speech perception in infants. *Science* 171:303–06. [GD-L]

Elliott, T., Howarth, C. I. & Shadbolt, N. R. (1996) Axonal processes and neural plasticity. I: Ocular dominance columns. *Cerebral Cortex* 6:781–88. [TE]

(1997) Axonal processes and neural plasticity. III: Competition for dendrites. *Philosophical Transactions of the Royal Society of London, B* (in press). [TE]

Elman, J. L. (1991) *Incremental learning or the importance of starting small* (Report no. 9101). University of California, San Diego, CA: Center for Research in Language. [GH]

(1993) Learning and development in neural networks: The importance of starting small. *Cognition* 48:71–99. [GH]

Elman, J. L., Bates, E. A., Johnson, M. H., Karmiloff-Smith, A., Parisi, D. & Plunkett, K. (1996) *Rethinking innateness: A connectionist perspective on development*. MIT Press. [arSRQ, JAF, MHJ]

Estes, D., Wellman, H. W. & Woolley, J. (1989) Children's understanding of mental phenomena. In: *Advances in child development and behavior*, ed. H. W. Reese. Academic Press. [DE]

Fahlman, S. E. & Lebiere, C. (1990) The cascade-correlation architecture. In: *Advances in neural information processing systems*, ed. D. S. Touretzky. Morgan Kaufmann. [aSRQ]

Feldman, J. A. & Ballard, D. H. (1982) Connectionist models and their properties. *Cognitive Science* 6:205–54. [aSRQ]

Fiala, B. A., Joyce, J. N. & Greenough, W. T. (1978) Environmental complexity modulates growth of granule cell dendrites in developing but not adult hippocampus of rats. *Experimental Neurology* 59:372–83. [aSRQ]

Finlay, B. L. & Darlington, R. B. (1995) Linked regularities in the development and evolution of mammalian brains. *Science* 268:1578–84. [aSRQ, BLF]

Fischer, K.W. (1980) A theory of cognitive development: The control and construction of hierarchies of skills. *Psychological Review* 87:477-531. [aSRQ, DM]

(1994) Dynamic development of coordination of components in brain and behavior. In: *Human behavior and the developing brain*, ed. G. Dawson & K. W. Fischer. Guilford Press. [DM]

Flood, J. F., Held, R. & Hein, A. (1963) Development and segmentation of visually controlled movement by selective exposure during rearing. *Journal of Comparative Physiological Psychology* 73:181–87. [JEB]

Fodor, J. (1975) *The language of thought*. Harvard University Press. [rSRQ]

(1980) On the impossibility of acquiring "more powerful" structures. In: *Language and learning: The debate between Jean Piaget and Noam Chomsky*, ed. M. Piattelli-Palmarini. Routledge & Kegan Paul. [rSRQ, MEJR]

(1983) *The modularity of mind*.. MIT Press [arSRQ, BJS]

(1992) Précis of *The modularity of mind*. *Behavioral and Brain Sciences* 8:1–42. [BLF]

Fodor, J., Bever, T. & Garrett, M. (1974) *The psychology of language*. McGraw-Hill. [rSRQ]

Fodor, J. & Pylyshyn, Z. (1981) How direct is visual perception? Some reflections on Gibson's "ecological approach." *Cognition* 9:139–96. [rSRQ]

Frasconi, P., Gori, M., Maggini, M. & Soda, G. (1996) Representation of finite state automata in recurrent radial basis function networks. *Machine Learning* 23(1):5–32. [EB]

Frean, M. (1990) The upstart algorithm: A method for constructing and training feedforward neural networks. *Neural Computation* 2:198–209. [aSRQ]

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

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

https://doi.org/10.1017/S0140525X97301587 Published online by Cambridge University Press

Friedlander, M. J., Martin, K. A. C. & Wassenhove-McCarthy, D. (1991) Effects of monocular visual deprivation on geniculocortical innervaton of area 18 in cat. *Journal of Neuroscience* 11:3268–88. [aSRQ]

Friston, K. J., Tononi, G., Reeke, G. N., Jr., Sporns, O. & Edelman, G. M. (1994) Value-dependent selection in the brain: Simulation in a synthetic neural model. *Neuroscience* 59:229–43. [OS]

Fritzke, B. (1994) Growing cell structure – A self-organizing network for un supervised and supervised learning. *Neural Networks* 7(9):1441–60. [EB]

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

Gally J. A., Montague, P. R., Reeke, G. N., Jr. & Edelman, G. M. (1990) The NO hypothesis: Possible effects of a short-lived, rapidly diffusible signal in the development and function of the nervous system. *Proceedings of the National Academy of Sciences USA* 87:3547–51. [aSRQ, GNR]

Gardner, H. (1983) *Frames of mind: The theory of multiple intelligences*. Basic Books. [GM]

Garfield, J. (1994) Modularity. In: *A companion to the philosophy of mind*, ed. S. Guttenplan. Blackwell. [BJS]

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

(1995) On neural circuits and cognition. *Neural Computation* 7:1–13. [aSRQ]

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

Gibson, J. J. (1979) *The ecological approach to visual perception*. Houghton Mifflin. [rSRQ]

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

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

Globus, A., Rosenzweig, M. R., Bennett, E. L. & Diamond, M. C. (1973) Effects of differential experience on dendritic spine counts in rat cerebral cortex. *Journal of Comparative and Physiological Psychology* 82:175–81. [aSRQ]

Globus, A. & Scheibel, A. B. (1967) The effect of visual deprivation on cortical neurons: A Golgi study. *Experimental Neurology* 19:331–245. [aSRQ]

Gold, E. M. (1967) Language identification in the limit. *Information and Control* 10:447–74. [arSRQ, DB]

Goldowsky, B. N. & Newport, E. L. (1993) Modeling the effects of processing limitation on the acquisition of morphology: The less is more hypothesis. In: *Proceedings of the Twenty-fourth Annual Child Language Research Forum*, ed. E. V. Clark. Stanford CA, CSLI, and University of Chicago Press. [GH]

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

Goodhill, G. J. (1992) Correlations, competition and optimality: Modelling the development of topography and ocular dominance. Cognitive Science research paper 226, University of Sussex. [aSRQ]

Goodman, C. & Shatz, C. (1991) The development of ocular dominance columns. In: *Development of the visual system*, ed. D. M.-K. Lam & C. J. Shatz. MIT Press. [aSRQ]

(1993) Developmental mechanisms that generate precise patterns of neuronal connectivity. *Cell* 72:77–98. [HK]

Goodwin, B. C., Kauffman, S. & Murray, J. D. (1993) Is morphogenesis an intrinsically robust process? *Journal of Theoretical Biology* 163:135–44. [rSRQ]

Gottlieb, G. (1976) Conceptions of prenatal development: behavioral embryology. *Psychological Review* 83: 215–34. [aSRQ]

(1991) Experiential canalization of behavioral development: Theory. *Developmental Psychology* 27:4–13. [aSRQ]

Greenfield, P. M. (1991) Language, tools and brain: The ontogeny and phylogeny of hierarchically organized sequential behavior. *Behavioral and Brain Sciences* 14:531–95. [rSRQ]

Greenough, W. T., Black, J. E. & Wallace, C. S. (1987) Experience and brain development. *Child Development* 58:539–59. [aSRQ, JJB]

Greenough, W. T. & Chang, F. L. (1988) Dendritic pattern formation involves both oriented regression and oriented growth in the barrels of mouse somatosensory cortex. *Brain Research* 471:148–52. [aSRQ, JEB]

Greenough, W. T. & Chang, F. L. (1988) Plasticity of synapse structure and pattern in the cerebral cortex. In: *Cerebral cortex*, vol. 7: *Development and maturation of cerebral cortex*, ed. A. Peters & E. G. Jones. Plenum Press. [OS]

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

Grinnell, A. D. (1995) Dynamics of nerve-muscle interactions in developing and mature neuromuscular junctions. *Physiological Reviews* 75:789–834. [TE]

Grossberg, S. (1975) A neural model of attention, reinforcement, and discrimination learning. *International Review of Neurobiology* 18:263–327. [SG]

(1978a) A theory of human memory: Self-organization and performance of sensory-motor codes, maps, and plans. In: *Progress in theoretical biology*, ed. R. Rosen & F. Snell. Academic Press. [SG]

(1978b) Communication, memory, and development. In: *Progress in theoretical biology*, eds. R. Rosen & F. Snell. Academic Press. [SG]

(1987) *The adaptive brain*, vols. 1 and 2. North Holland. [SG]

(1988) *Neural networks and natural intelligence*. MIT Press. [SG]

(1995) The attentive brain. *American Scientist* 83:438–49. [SG]

(1997) Neural models of reaching. *Behavioral and Brain Sciences* 20(2):310. [SG]

Grossberg, S. & Kuperstein, M. (1986) *Neural dynamics of adaptive sensory-motor control: Ballistic eye movements*. Elsevier. [SG]

Grossberg, S. & Merrill, J. W. L. (1996) The hippocampus and cerebellum in adaptively timed learning, recognition, and movement. *Journal of Cognitive Neuroscience* 8:257–77. [SG]

Grossberg, S., Roberts, K., Aguilar, M. & Bullock, D. (1996) A neural model of multimodal adaptive saccadic eye movement control by superior colliculus. Technical Report CAS/CNS TR-96–029, Boston University. [SG]

Grossberg, S. & Schmajuk, N. (1987) Neural dynamics of Pavlovian conditioning: Conditioned reinforcement, inhibition, and opponent processing. *Psychobiology* 15:195–240. [SG]

Haley, J. E., Wilcox, G. L. & Chapman, P. F. (1992) The role of nitric oxide in hippocampal long-term potentiation. *Neuron* 8:211–16. [aSRQ]

Halford, G. & Simon, T., eds. (1995) *Developing cognitive competence: New approaches to process modeling*. LEA. [DM]

Harris, R. M & Woolsey, T. A. (1981) Dendritic plasticity in mouse barrel cortex following postnatal vibrissa follicle damage. *Journal of Comparative Neurology* 196:357–76. [aSRQ]

Hata, Y. & Stryker, M. P. (1994) Control of thalamocortical afferent rearrangement by postsynaptic activity in developing visual cortex. *Science* 265:1732–35. [TE]

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

Hayes, W. P. & Meyer, R. L. (1988) Optic synapse number but not density is constrained during regeneration onto a surgically halved tectum in goldfish: HRP-EM evidence that optic fibers compete for fixed numbers of postsynaptic sites on the tectum. *Journal of Comparative Neurology* 274: 539–59. [BLF]

Hebb, D. O. (1949) *The organization of behavior: A neuropsychological theory*. Wiley. [aSRQ, JF, ABS]

Heiligenberg, W. (1991) *Neural nets in electric fish*. MIT Press. [rSRQ]

Helmholtz, H. L. F. von (1924) *Helmholtz's treatise on physiological optics*, translated from the Third German Edition by Southall, J. P. C., vols. 1–3. George Banta. [DP]

Herrmann, K. & Shatz, C. J. (1995) Blockade of action potential activity alters initial arborization of thalamic axons within cortical layer 4. *Proceedings of the National Academy of Sciences USA* 92:11244–48. [aSRQ]

Hinton, G. & Nowlan, S. (1987) How learning can guide evolution. *Complex Systems* 1:495–502. [JH]

Hirose, Y., Yamashita, K. & Hijiya, S. (1991) Back-propagation algorithm which varies the number of hidden units. *Neural Networks* 4:61–66. [aSRQ]

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

Holly, J. E. & McCollum, G. (submitted) Timing of secondary vestibular neuron responses to a range of angular head movements. [GM]

Horn, G. (1985) *Memory, imprinting, and the brain*. Clarendon Press. [JJB]

Hornik, K., Stinchcombe, M. & White, H. (1989) Multilayer feedforward networks are universal approximations. *Neural Networks* 2:359–66. [aSRQ]

Hubel, D. & Wiesel, T. (1962) Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *Journal of Physiology (London)* 160:106–54. [aSRQ]

(1963) Receptive fields of cells in striate cortex of very young, visually inexperienced kittens. *Journal of Neurophysiology* 26:994–1002. [aSRQ]

(1965) Binocular interactions in striate cortex of kittens reared with artificial squint. *Journal of Neurophysiology* 28:1041–59. [aSRQ]

(1972) Laminar and columnar distribution of geniculo-cortical fibers in the macaque monkey. *Journal of Comparative Neurology* 146:421–50. [aSRQ]

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

Huttenlocher, P. R. (1979) Synaptic density in human frontal cortex-

https://doi.org/10.1017/S0140525X97301587 Published online by Cambridge University Press

developmental changes and effects of aging. *Brain Research* 163:195–205. [arSRQ, GD-L]
(1990) Morphometric study of human cerebral cortex development. *Neuropsychologia* 28:517–27. [aSRQ]
Huttenlocher, P. R. & de Courten, C. (1987) The development of synapses in striate cortex of man. *Human Neurobiology* 6:1–9. [aSRQ]
Huttenlocher P. R., de Courten, C., Garey, L. J. & Van der Loos, H. (1982) Synaptogenesis in human visual cortex: Evidence for synapse elimination during normal development. *Neuroscience Letters* 13:247–52. [aSRQ]
Innocenti, G. M. (1991) The development of projections from cerebral cortex. In: *Progress in sensory physiology*, vol. 12. Springer-Verlag. [GMI]
(1995) Exuberant development of connections, and its possible permissive role in cortical evolution. *TINS* 18:397–402. [GMI]
Innocenti, G. M. & Tettoni, L. (in press) Exuberant growth, specificity, and selection in the differentiation of cortical axons. In: *Normal and abnormal development of the cortex*, ed. A. Galaburda & Y. Christen. Springer-Verlag. [GMI]
Iriki, A., Nozaki, S. & Nakamura, Y. (1988) Feeding behavior in mammals: Corticobulbar projection is reorganized during conversion from sucking to chewing. *Developmental Brain Research* 44:189–96. [GMI]
Iverson, P. & Kuhl, P. K. (1995) Mapping the perceptual magnet effect for speech using signal detection theory and multidimensional scaling. *Journal of the Acoustical Society of America* 97:553–62. [rSRQ]
Jacobs, B., Schall, M. & Scheibel, A. B. (1993) A quantitative dendritic analysis of Wernicke's area in humans. II. Gender, hemispheric, and environmental factors. *Journal of Comparative Neurology* 327:97–111. [aSRQ]
Jaslove, S. W. (1992) The integrative properties of spiny distal dendrites. *Neuroscience* 47:495–519. [aSRQ]
Jerne, N. (1967) Antibodies and learning: Selection versus instruction. In: *The neurosciences: A study program*, ed. G. C. Quarton, T. Melnechuk, F. O. Schmitt. Rockefeller University Press. [aSRQ]
Jernigan, T. L., Archibald, S. L., Berhow, M. T., Sowell, E. R., Foster, D. S. & Hesselink, J. R. (1991) Cerebral structure on MRI, Part I: Localization of age-related changes. *Biological Psychiatry* 29:55–67. [aSRQ]
Johnson, M. (1996) *Developmental cognitive neuroscience*. Blackwell. [JAF]
Johnson, M. H. (1990) Cortical maturation and the development of visual attention in early infancy. *Journal of Cognitive Neuroscience* 2:81–95. [aSRQ]
Johnson, M. H. & Morton, J. (1991) *Biology and cognitive development: The case of face recognition*. Blackwell. [MHJ]
Jones, O. T., Kunze, D. L. & Angelides, K. J. (1989) Localization and mobility of omega-conotoxin-sensitive Ca^{2+} channels in hippocampal CA1 neurons. *Science* 244:1189–93. [aSRQ]
Judd, S. (1988) On the complexity of loading shallow neural networks. *Journal of Complexity* 4:177–92. [aSRQ]
Juraska, J. M. & Fifkova, E. (1979) A Golgi study of the early postnatal development of the visual cortex of the hooded rat. *Journal of Comparative Neurology* 183:247–56. [aSRQ]
Juraska, J. M., Greenough, W. T., Elliott, C., Mack, K. J. & Berkowitz, R. (1980) Plasticity in adult rat visual cortex: An examination of several cell populations after differential rearing. *Behavioral and Neural Biology* 29:157–67. [aSRQ]
Kaas, J. (1995) The reorganization of sensory and motor maps in adult mammals. In: *The cognitive neurosciences*, ed. M. S. Gazzaniga. MIT Press. [HK]
Kadirkamanathan, V. & Niranjan, M. (1993) A function estimation approach to sequential learning with neural networks. *Neural Computation* 5:954–75. [aSRQ]
Kalil, R. E., Dubin, M. W., Scott, G. & Stark, L. A. (1986) Elimination of action potentials blocks the structural development of retinogeniculate synapses. *Nature* 323:156–58. [aSRQ]
Karmiloff-Smith, A. (1992) *Beyond modularity: A developmental perspective on cognitive science*. MIT Press. [arSRQ, RAB, DE, BJS]
Kasper, E. M., Larkman, A. U., Lubke, J. & Blakemore, C. (1994) Pyramidal neurons in layer 5 of the rat visual cortex. II. Development of electrophysiological properties. *Journal of Comparative Neurology* 339:475–94. [aSRQ]
Katz, L. C. & Constantine-Paton, M. (1988) Relationships between segregated afferents and postsynaptic neurones in the optic tectum of three-eyed frogs. *Journal of Neuroscience* 8:3160–80. [aSRQ]
Katz, L. C., Gilbert, C. D. & Wiesel, T. N. (1989) Local circuits and ocular dominance columns in monkey striate cortex. *Journal of Neuroscience* 9:1389–99. [aSRQ]
Katz, L. C. & Shatz, C. J. (1996) Synaptic activity and the construction of cortical circuits. *Science* 274:1133–38. [GD-L]
Katz, M. J., George, E. B. & Gilbert, L. J. (1984) Axonal elongation as a stochastic walk. *Cell Motility* 4:351–70. [aSRQ]
Kauffman, S. (1992) *The origins of order*. Oxford University Press [rSRQ]
Keil, F. (1990) Constraints on constraints: Surveying the epigenetic landscape. *Cognitive Science* 14:135–68. [DM]
Kelso, S. R., Ganong, A. H. & Brown, T. H. (1986) Hebbian synapses in hippocampus. *Proceedings of the National Academy of Sciences USA* 83:5326–30. [aSRQ]
Kempermann, G., Kuhn, H. G. & Gage, F. H. (1997) More hippocampal neurons in adult mice living in an enriched environment. *Nature* 386:493–95. [rSRQ]
Kennedy, H. & Dehay, C. (1993a) Cortical specification of mice and men. *Cerebral Cortex* 3(3):27–35. [HK]
(1993b) The importance of developmental timing in cortical specification. *Perspectives in Developmental Neurobiology* 1:93–100. [HK]
Kennedy, H., Salin, P., Bullier, J. & Horsburgh, G. (1994) Topography of developing thalamic and cortical pathways in the visual-system of the cat. *Journal of Comparative Neurology* 348(2):298–319. [HK]
Killackey, H. (1990) Neocortical expansion: An attempt toward relating phylogeny and ontogeny. *Journal of Cognitive Neuroscience* 2:1–17. [HK]
Kirby, S. & Hurford, J. (1997) The evolution of incremental learning: Language, development and critical periods. Occasional Paper EOPL-97–2, Department of Linguistics, University of Edinburgh, Edinburgh. [JH]
Knight, R. T. & Grabowecky, M. (1995) Escape from linear time: Prefrontal cortex and conscious experience. In: *The cognitive neurosciences*, ed. M. S. Gassaniga. MIT Press. [DE]
Koch, C. & Poggio, T. (1992) Multiplying with synapses and neurons. In: *Single neuron computation*, ed. T. McKenna, J. Davis, S. Zornetzer. Academic Press. [aSRQ]
Koch, C., Poggio, T. & Torre, V. (1982) Retinal ganglion cells: A functional interpretation of dendritic morphology. *Philosophical Transactions of the Royal Society of London, B: Biological Sciences* 298:227–63. [aSRQ]
(1983) Nonlinear interactions in a dendritic tree: Localization, timing, and role in information processing. *Proceedings of the National Academy of Sciences USA* 80:2799–802. [aSRQ]
Koechlin, E., Dehaene, S. & Mehler, J. (submitted) Numerical transformations in five-month-old human infants. [GD-L]
Koester, S. E. & O'Leary, D. D. (1992) Functional classes of cortical projection neurons develop dendritic distinctions by class-specific sculpting of an early common pattern. *Journal of Neuroscience*12:1382–93. [aSRQ]
Konishi, M. (1991) Deciphering the brain's codes. *Neural Computation* 3:1–18. [rSRQ]
Kossel, A., Lowel, S. & Bolz, J. (1995) Relationships between dendritic fields and functional architecture in striate cortex of normal and visually deprived cats. *Journal of Neuroscience*15:3913–26. [arSRQ]
Krech, D., Rosenzweig, M. & Bennett, E. (1960) Effects of environmental complexity and training on brain chemistry. *Journal of Comparative and Physiological Psychology* 53:509–19. [ABS]
Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N. & Lindblom, B. (1992) Linguistic experiences alter phonetic perception in infants by 6 months of age. *Science* 255:606–08. [rSRQ, GD-L]
Lee, K. S., Schottler, F. Oliver, F. & Lynch, G. (1980) Brief bursts of high-frequency stimulation produce two types of structural change in rat hippocampus. *Journal of Neurophysiology* 44:247–58. [aSRQ]
Lenneberg, E. H. (1967) *Biological foundations of language*. Wiley. [aSRQ]
LeVay, S. & Stryker, M. P. (1979). The development of ocular dominance columns in the cat. In: *Society for neuroscience symposium: Aspects of developmental neurobiology*, ed. J. A. Ferrendelli. Society for Neuroscience. [aSRQ]
LeVay, S., Wiesel, T. N. & Hubel, D. H. (1980) The development of ocular dominance columns in normal and visually deprived monkeys. *Journal of Comparative Neurology* 191:1–51. [aSRQ]
Levine, M. (1966) Hypothesis behaviour by humans doing discrimination learning. *Journal of Experimental Psychology* 71:331–36. [JH]
Lichtman, J. W. (1977) The reorganization of synaptic connexions in the rat submandibular ganglion during post-natal development. *Journal of Physiology (London)* 320:121–30. [aSRQ]
Lightfoot, D. (1989) The child's trigger experience: Degree-0 learnability. *Behavioral and Brain Sciences* 12:321–75. [aSRQ]
(1991) *How to set parameters: Arguments from language change*. MIT Press. [arSRQ]
Linsker, R. (1986) From basic network principles to neural architecture: Emergence of orientation-selective cells. *Proceedings of the National Academy of Sciences USA* 83:8390–4. [aSRQ]
Lubke, J. & Albus, K. (1989) The postnatal development of layer VI pyramidal neurons in the cat's striate cortex, as visualized by intracellular Lucifer yellow injections in aldehyde-fixed tissue. *Developmental Brain Research* 45:29–38. [aSRQ]
Lund, J. S. & Holbach, S. M. (1990) Postnatal development of thalamic recipient neurons in the monkey striate cortex: I. Comparison of spine acquisition

https://doi.org/10.1017/S0140525X97301587 Published online by Cambridge University Press

and dendritic growth of layer 4C alpha and beta spiny stellate neurons. *Journal of Comparative Neurology* 309:115–28. [aSRQ]

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

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

Maffei, L. & Galli-Resta, L. (1990) Correlation in the discharges of neighboring rat retrial ganglion cells during prenatal life. *Proceedings of the National Academy of Sciences USA* 87:2861–64. [aSRQ]

Mainen, Z. F., Joerges, J., Huguenard, J. R. & Sejnowski, T. J. (1995) A model of spike initiation in neocortical pyramidal neurons. *Neuron* 15:1427–39. [aSRQ]

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

Marcus, M. (1995) New trends in natural language processing: Statistical natural language processing. *Proceedings of the National Academy of Sciences USA* 92:10052–59. [rSRQ]

Mareschal, D. & Shultz, T. R. (1996) Generative connectionist networks and constructivist cognitive development. *Cognitive Development* 11:571–603. [DM]

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

Marr, D. (1982) *Vision*. W. H. Freeman. [rSRQ]

Matesz, C., Birinyi, A., Kothalawala, D. S. & Székely, G. (1995) Investigation of the dendritic geometry of brain stem motoneurons with different functions using multivariant statistical techniques in the frog. *Neuroscience* 65:1129–44. [GS]

Mathers, L. J. (1979) Postnatal dendritic development in the rabbit visual cortex. *Brain Research* 168:21–29. [aSRQ]

McAllister, A. K., Katz, L. C. & Lo, D. C. (1996) Neurotrophin regulation of cortical dendritic growth requires activity. *Neuron* 17:1057–64. [rSRQ, HK]

McAllister, A. K., Lo, D. C. & Katz, L. C. (1995) Neurotrophins regulate dendritic growth in developing visual cortex. *Neuron* 15:791–803. [rSRQ, TE]

McCabe, B. J. & Horn, G. (1988) Learning and memory: Regional changes in N-methyl-D-aspartate receptors in the chick brain after imprinting. *Proceedings of the National Academy of Sciences USA* 85:2849–53. [JJB]

McCasland, J. S., Bernardo, K. L., Probst, K. & Woolsey, T. A. (1992) Cortical local circuit axons do not mature after early deafferentation. *Proceedings of the National Academy of Sciences USA* 89:1832–36. [aSRQ]

McCollum, G. (1992) Rules of combination that generate climbing fiber tactile receptive fields. *Neuroscience* 50:707–25. [GM]

(1993) Reciprocal inhibition, synergies, and movements. *Journal of Theoretical Biology* 165:291–311. [GM]

(1994) Dissonance: A nervous system analogue to quantum incompatibility. *International Journal of Theoretical Physics* 33:41–52. [GM]

McCollum, G., Holroyd, C. & Castelfranco, A. M. (1995) Forms of early walking. *Journal of Theoretical Biology* 176:373–90. [GM]

Mead, C. (1989) *Analog VLSI and neural systems*. Addison-Wesley. [aSRQ]

Mehler, J. (1985) Language related dispositions in early infancy. In: *Neonate cognition: Beyond the blooming, buzzing confusion*, ed. J. Mehler & R. Fox. Lawrence Erlbaum. [arSRQ]

Mehler, J. & Christophe, A. (1995) Maturation and learning of language during the first year of life. In: *The cognitive neurosciences*, ed. M. S. Gazzaniga. Bradford Books/MIT Press. [GD-L]

Mehler, J., Jusczyk, P., Lambertz, G., Halsted, N., Bertoncini, J. & Amiel-Tison, C. (1988) A precursor of language acquisition in young infants. *Cognition* 29:143–78. [GD-L]

Meissirel, C, Dehay, C. & Kennedy, H. (1993) Transient cortical pathways in the pyramidal tract of the neonatal ferret. *Journal of Comparative Neurology* 338:193–213. [HK]

Meister, M., Wong, R., Baylor, D. & Shatz, C.J. (1991) Synchronous bursts of action potentials in ganglion cells of the developing mammalian retina. *Science* 252:939–43. [aSRQ]

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

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

(1994) Information processing in dendritic trees. *Neural Computation* 6:1031–85. [aSRQ]

Mel, B. W. & Koch, C. (1990) Sigma-Pi learning: On radial basis functions and cortical associative learning. In: *Advances in neural information processing systems*, ed. D. S. Touretzky. Morgan Kaufmann. [aSRQ]

Miller, K. D., Keller, J. B. & Stryker, M. P. (1989) Ocular dominance column development: Analysis and simulation. *Science* 245:605–15. [aSRQ]

Mitchison, G. J. & Durbin, R. (1986) Optimal numberings of an N X N array. *S.I.A.M. Journal on Algebraic and Discrete Methods* 7:571–82. [aSRQ]

Mollgard, K., Diamond, M. C., Bennett, E., Rosenzweig, M. & Lindner, B. (1971) Quantitative synaptic changes with differential experience in rat brain. *International Journal of Neuroscience* 2:113–28. [rSRQ, ABS]

Montague, P. R. (1996) The resource consumption principle-attention and memory in volumes of neural tissue. *Proceedings of the National Academy of Sciences USA* 93:3619–23. [aSRQ]

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

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

Montague, P. R. & Sejnowski, T.J. (1994) The predictive brain: Temporal coincidence and temporal order in synaptic learning mechanisms. *Learning and Memory* 1:1–33. [aSRQ]

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

Moore, D., Benenson, J., Reznick, J. S., Peterson, M. & Kagan, J. (1987) Effect of auditory numerical information on infant's looking behavior: Contradictory evidence. *Developmental Psychology* 23:665–70. [GD-L]

Natarajan, B. (1991) *Machine learning: A theoretical approach*. Morgan Kaufmann. [aSRQ]

Nazzi, T., Bertoncini, J. & Mehler, J. (in press) Language discrimination by newborns: Towards an understanding of the role of rhythm. *Journal of Experimental Psychology: Human Perception and Performance*. [GD-L]

Neville, H. (1991) Neurobiology of cognitive and language processing: effects of early experience. In: *Brain maturation and cognitive development*, ed. K. R. Gibson & A. C. Peterson. Aladine de Gruyter Press. [aSRQ]

Newport, E. L. (1990) Maturational constraints on language learning. *Cognitive Science* 14(1):11–28. [GH]

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

Nottebohm, F. (1991) Reassessing the mechanisms and origins of vocal learning in birds. *Trends in Neurosciences* 14:206–11. [JJB]

O'Dell, T. J., Hawkins, R. D., Kandel, E. R. & Arancio O. (1991) Tests of the roles of two diffusible substances in long-term potentiation: Evidence for nitric oxide as a possible early retrograde messenger. *Proceedings of the National Academy of Sciences USA* 88:11285–89. [aSRQ]

O'Kusky, J. & Colonnier, M. (1982a) Postnatal changes in the number of neurons and synapses in the visual cortex (area 17) of the macaque monkey: A stereological analysis in normal and monocularly deprived animals. *Journal of Comparative Neurology* 210:307–15. [aSRQ]

(1982b) Postnatal changes in the number of astrocytes, oligodendrocytes, and microglia in the visual cortex (area 17) of the macaque monkey: A stereological analysis in normal and monocularly deprived animals. *Journal of Comparative Neurology* 210:307–15. [aSRQ]

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

(1992) Development of connectional diversity and specificity in the mammalian brain by the pruning of collateral projections. *Current Opinion in Neurobiology* 2:70–77. [aSRQ, GMI]

O'Leary, D. D. M., Schlaggar, B. L. & Stanfield, B. B. (1992) The specification of sensory cortex: lessons from cortical transplantation. *Expermental Neurology* 115:121–26. [aSRQ]

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

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

(1990) Dynamic changes in optic fiber terminal arbors lead to retinotopic map formation: An in vivo confocal microscopic study. *Neuron* 5:159–71. [aSRQ]

Osherson, D. N., Stob, M. & Weinstein, S. (1986) *Systems that learn*. MIT Press. [aSRQ]

Oudega, M., Varon, S. & Hagg, T. (1994) Distribution of corticospinal motor neurons in the postnatal rat. Quantitative evidence for massive collateral elimination and modest cell death. *Journal of Comparative Neurology* 347:115–26. [HK]

https://doi.org/10.1017/S0140525X97301587 Published online by Cambridge University Press

Pallas, S. L. & Finlay, B. L. (1989) Conservation of receptive field properties of superior colliculus cells after developmental rearrangements of retinal input. *Visual Neuroscience* 2:121–35. [BLF]

Pallas, S. L., Roe, A. W. & Sur, M. (1990) Visual projections induced into the auditory pathway of ferrets. Novel inputs to primary audtory cortex (AI) 59 from the LP/pulvinar complex and the topography of the MGN-AI projection. *Journal of Comparative Neurology* 298:50–68. [aSRQ]

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

Patel, S. N., Rose, S. P. R. & Stewart, M. G. (1988) Training induced dendritic spine density changes are specifically related to memory formation processes in the chick, Gallus domesticus. *Brain Research* 463:168–73. [JJB]

Pearce, I. A., Cambray-Deakin, M. A. & Burgoyne, R. D. (1987) Glutamate acting on NMDA receptors stimulates neurite outgrowth from cerebellar granule cells. *Febs Letters* 223:143–47. [aSRQ]

Peters, A. & Payne, B. R. (1993) Numerical relationships between geniculocortical afferents and pyramidal cell modules in cat primary visual cortex. *Cerebral Cortex* 3:69–78. [HK]

Petit, T. L., LeBoutillier, J. C., Gregorio, A. & Libstug, H. (1988) The pattern of dendritic development in the cerebral cortex of the rat. *Brain Research* 469, 209–19. [aSRQ]

Piaget, J. (1970) *Genetic epistemology*. Columbia University Press. [DM]

(1980) *Adaptation and intelligence: Organic selection and phenocopy*. University of Chicago Press. [DM]

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

Pinker, S. (1979) Formal models of language learning. *Cognition* 1:217–83. [aSRQ]

(1984) *Language learnability and language development*. Harvard University Press. [aSRQ, HK]

(1994) *The language instinct*. W. Morrow. [arSRQ]

Pinker, S. & Bloom, P. (1990) Natural language and natural selection. *Behavioral and Brain Sciences* 13:707–84. [rSRQ]

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

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

Plunkett, K. & Elman, J. (1997) *Exercises in rethinking innateness. A handbook for connectionist simulations.* MIT Press [rSRQ]

Plunkett, K. & Sinha, C. (1992) Connectionism and developmental theory. *British Journal of Developmental Psychology*10:209–54. [aSRQ]

Pollack, J. B. (1991) The induction of dynamical recognizers. *Machine Learning* 7:227–52. [MEJR]

Price, D. J. & Blakemore, C. (1985) The postnatal development of the association projection from visual cortical area 17 to area 18 in the cat. *Journal of Neuroscience* 5(9):2443–52. [HK]

Purves, D. (1988a) *Body and brain: A trophic theory of neural connections.* Harvard University Press. [aSRQ, DP]

(1988b) A new theory of brain function (review of G. Edelman's book, *Neural Darwinism*). *Quarterly Review of Biology* 63:202–04. [DP]

(1994) *Neural activity and the growth of the brain*. Cambridge University Press (UK). [aSRQ, TE, DP, OS]

Purves, D. & Hadley, R. D. (1985) Changes in the dendritic branching of adult mammalian neurones revealed by repeated imaging in situ. *Nature* 315:404–06. [aSRQ]

Purves, D., Hadley, R. D. & Voyvodic, J. T. (1986) Dynamic changes in the dendritic geometry of individual neurons visualized over periods of up to three months in the superior cervical ganglion of living mice. *Journal of Neuroscience* 6:1051–60. [aSRQ]

Purves D. & Lichtman, J. W. (1985). *Principles of neural development.* Sinauer. [aSRQ, JH]

Purves, D., Voyvodic, J., Magrassi, L. & Yawo, H. (1987) Nerve terminal remodeling visualized in living mice by repeated examination of the same neuron. *Science* 238:1122–26. [aSRQ]

Purves, D., White, L. E. & Riddle, D. R. (1996) Is neural development Darwinian? *Trends in Neurosciences* 19:460–64. [rSRQ, DP, OS]

Purvis, A. (1995) A composite estimate of primate phylogeny. *Philosophical Transactions of the Royal Society of London, B* 348:405–21. [RAB]

Purvis, A. & Rambaut, A. (1995) Comparative analysis by independent contrasts (CAIC): An Apple Macintosh application for analysing comparative data. *Computer Applications in the Biosciences* 11:247–51. [RAB]

Pylyshyn, Z. (1984) *Computation and cognition: Toward a foundation for cognitive science*. Bradford Books. [arSRQ]

(1985) Plasticity and invariance in cognitive development. In: *Neonate cognition*, ed. J. Mehler & R. Fox. Lawrence Erlbaum. [BJS]

Quartz, S. R. (1993) Nativism, neural networks, and the plausibility of constructivism. *Cognition* 48:123–44. [arSRQ, MEJR]

Quartz, S. R. & Sejnowski, T. J. (1994) Beyond modularity: Neural evidence for constructivist principles in development. *Behavioral and Brain Sciences* 17:725–26. [aSRQ, BJS]

Raijmakers, M. E. J., van der Maas, H. L. J. & Molenaar, P. C. M. (1996) Numerical bifurcation analysis of distance-dependent on-center off-surround shunting neural networks. *Biological Cybernetics* 75:495–507. [MEJR]

Raijmakers, M. E. J., van Koten, S. & Molenaar, P. C. M. (1996) On the validity of simulating stagewise development by means of PDP-networks: Application of catastrophe analysis and an experimental test of rule-like network performance. *Cognitive Science* 20–1:101–36. [MEJR]

Rakic, P. (1982) Early developmental events: Cell lineages, acquisition of neuronal positions, and areal and laminar development. *Neurosciences Research Program Bulletin* 20:439–51. [HK]

(1988) Specification of cerebral cortical areas. *Science* 241:170–76. [HK]

Rakic, P., Bourgeois, J. P., Eckenhoff, M. F., Zecevic, N. & Goldman-Rakic, P. S. (1986) Concurrent overproduction of synapses in diverse regions of the primate cerebral cortex. *Science* 232:232–35. [arSRQ]

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

Rakic, P. & Sidman, R. L. (1973) Weaver mutant mouse cerebellum: defective neuronal migration secondary to abnormality of Bergmann glia. *Proceedings of the National Academy of Sciences USA* 70:240–44. [aSRQ]

Rakic, P., Suner, I. & Williams, R. W. (1991) A novel cytoarchitectonic area induced experimentally within the primate visual cortex. *Proceedings of the National Academy of Sciences USA* 88(6):2083–87. [HK]

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

Ramon y Cajal, S. (1952) *Histologie du systeme nerveux de l'homme et des vertebres*. Reprinted in 2 vols. by Consejo Superior de Investigaciones Cientificas, Madrid. [ABS]

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

Reeke, G. N., Jr., Finkel, L. H., Sporns, O. & Edelman, G. M. (1990) Synthetic neural modeling: A multilevel approach to the analysis of brain complexity. In: *Signal and sense: Local and global order in perceptual maps*, ed. G. M. Edelman, W. E. Gall & W. M. Cowan. Wiley. [GNR, OS]

Regehr, W. G., Connor, J. A. & Tank, D. W. (1989) Optical imaging of calcium accumulation in hippocampal pyramidal cells during synaptic activation. *Nature* 341:533–36. [aSRQ]

Regier, T. (1996) *The human semantic potential*. MIT Press. [JAF]

Riddle, D., Richards, A., Zsuppan, F. & Purves, D. (1992) Growth of the rat somatic sensory cortex and its constituent parts during postnatal development. *Journal of Neuroscience* 12:3509–24. [rSRQ, DP]

Riddle, D. R., Gutierrez, G., Zheng, D., White, L., Richards, A. & Purves, D. (1993) Differential metabolic and electrical activity in the somatic sensory cortex of juvenile and adult rats. *Journal of Neuroscience* 13:4193–4213. [rSRQ, DP]

Roberts, P. D. & McCollum, G. (1996) The stomatogastric nervous system: A formal approach. *Neuroscience* 72:1089–1105. [GM]

Roe, A. W., Pallas, S. L., Hahm, J. & Sur, M. (1990) A map of visual space induced in primary auditory cortex. *Science* 250:818–20. [aSRQ]

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

Roux, W. (1883) Beiträge zur Morphologie der functionellen Anpassung: Nr. 1. Über die Strukter eines hochdifferenzierten bindegewebigen Organ (der Schwanzflosse des Delphins). *Arch. Anat. Physiol. Anat. Abtlg.* 76–160. [aSRQ]

Rowe, M. H. (1991) Functional organization of the retina. In: *Neuroanatomy of visual pathways and their development*, vol. 3, ed. B. Dreher & S. R. Robinson. CRC Press. [BLF]

Rucci, M., Tononi, G. & Edelman, G. M. (1997) Registration of neuronal maps through value-dependent learning: Modeling the alignment of auditory and visual maps in the barn owl's optic tectum. *Journal of Neuroscience* 17:334–52. [OS]

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

Rumelhart, D., McClelland, J. & the PDP research group. (1986). *Parallel distributed processing: Explorations in the microstructure of cognition*. Bradford Books. [aSRQ, SG]

https://doi.org/10.1017/S0140525X97301587 Published online by Cambridge University Press

Saffran, J. R., Aslin, R. N. & Newport, E. L. (1996) Statistical learning by 8-month-old infants. *Science* 274:1926–28. [rSRQ]

Sarter, M., Bernston, G. & Cacioppo, J. (1996) Brain imaging and cognitive neuroscience: Toward strong inference in attributing function to structure. *American Psychologist* 51:13–21. [BJS]

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

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

Scheibel, A. B., Conrad, T., Perdue, S., Tomiyasu, U. & Wechsler, A. (1990) A quantitative study of dendrite complexity in selected areas of the human cerebral cortex. *Brain and Cognition* 12:85–101. [ABS]

Scheibel, A. B., Paul, L., Fried, I., Forsythe, A., Tomiyasu, U., Wechsler, A., Kao, A. & Slotnick, J. (1985) Dendritic organization of the anterior speech area. *Experimental Neurology* 87:109–17. [ABS]

Scheich, H. (1996) Auditory imprinting is related to synaptic selection in the chick. Paper presented at Avian Brain and Behaviour Meeting, Tihany, Hungary, August 1996. [JJB]

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

Schlaggar, B. L. & O'Leary, D. D. M. (1991) Potential of visual cortex to develop an array of functional units unque to somatosensory cortex. *Science* 252:1556–60. [aSRQ]

Schuman, E. M. & Madison, D. V. (1991) A requirement for the intercellular messenger nitric oxide in long-term potentiation. *Science* 254:1503–06. [aSRQ]

Segal, G. (1996) The modularity of theory of mind. In: *Theories of theories of mind*, ed. P. Carruthers & P. Smith. Cambridge University Press. [BJS]

Segev, I., Rinzel, J. & Shepherd, G.M. (1995). *The theoretical foundations of dendritic function: Selected papers by Wilfrid Rall with commentaries*. MIT Press. [aSRQ]

Seidenberg, M. S. Language acquisition and use: Learning and applying probabilistic constraints. *Science* 275:1599–603. [rSRQ]

Sejnowski, T. J. & Tesauro, G. (1989) The Hebb rule for synaptic plasticity: Algorithms and implementations. In: *Neural models of plasticity*, ed. J. H. Byrne & W. O. Berry. Academic Press. [aSRQ]

Sengelaub, D. R. (1989) Cell generation, migration, death and growth in neural systems mediating social behavior. In: *Advances in comparative and environmental physiology*, vol. 3, ed. J. Balthazart. Springer-Verlag. [BLF]

Shastri, L. & Ajanagadde, V. (1993) From simple associations to systematic reasoning: A connectionist representation of rules, variables and dynamic bindings using temporal synchrony. *Behavioral and Brain Sciences* 16(3):417–94. [JAF]

Shatz, C. J. (1990) Impulse activity and the patterning of connections during CNS development. *Neuron* 5:745–56. [aSRQ]

(1992) How are specifc connections formed between thalamus and cortex? *Current Opinion in Neurobiology* 2:78–82. [aSRQ]

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

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

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

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

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

Shultz, T. R., Mareschal, D. & Schmidt, W. C. (1994) Modeling cognitive development on balance scale phenomena. *Machine Learning* 16:57–86. [aSRQ, DM]

Shultz, T. R., Schmidt, W. C., Buckingham, D. & Mareschal, D. (1995) Modeling cognitive development with a generative connectionist algorithm. In: *Developing cognitive competence: New approaches to process modeling*, ed. T. J. Simon & G. S. Halford. LEA. [DM]

Shulz, H-D. (1967) *Metrische untersuchungen an den schichten des corpus geniculatum laterale tag- und nachtaktiver primaten*. Doctoral dissertation, Johann Wolfgang Goethe-Universität (Frankfurt). [RAB]

Siegler, R. S. (1989) Mechanisms of cognitive development. *Annual Review of Psychology* 40:353–79. [aSRQ]

Simon, H. A. (1962) The architecture of complexity. *Proceedings of the American Philosophical Society* 106:467–82. [GS]

Simmonds, R. J. & Scheibel, A. B. (1989) The postnatal development of the motor speech area: A preliminary study. *Brain and Language* 37:42–58. [aSRQ, ABS]

Singer, W. (1995) Development and plasticity of cortical processing architectures. *Science* 270:758–64. [HK]

Sperry, R. (1943) Effect of 180° rotation of the retinal fields on visuomotor coordination. *Journal of Experimental Zoology* 92:263–79. [aSRQ]

(1963) Chemoaffinity in the orderly growth of nerve fiber patterns and connections. *Proceedings of the National Academy of Science USA* 50:703–10. [aSRQ, GS]

Sporns, O. (1997a) Biological variability and brain function. In: *Binding the mind: Consciousness and human identity*, ed. J. Cornwell. Oxford University Press. [OS]

(1997b) Variation and selection in neural function. *Trends in Neuroscience* 20:291. [OS]

Sporns, O. & Edelman, G. M. (1993) Solving Bernstein's problem: A proposal for the development of coordinated movement by selection. *Child Development* 64:960–81. [OS]

Sporns, O. & Tononi, G. (1994) *Selectionism and the brain*. Academic Press. [OS]

Stanfield, B. B. & O'Leary, D. D. (1985) Fetal occipital cortical neurones transplanted to the rostral cortex can extend and maintain a pyramidal tract axon. *Nature* 313:135–37. [aSRQ]

Starck, D. (1982) *Vergleichende Anatomie der Virbeltiere*, vol. 3. Springer-Verlag. [GS]

Starkey, P. & Cooper, R. G. J. (1980) Perception of numbers by human infants. *Science* 210:1033–35. [GD-L]

Starkey, P., Spelke, E. S. & Gelman, R. (1983) Detection of intermodal numerical correspondences by human infants. *Science* 222:179–81. [GD-L]

(1990) Numerical abstraction by human infants. *Cognition* 36:97–127. [GD-L]

Stephan, H., Frahm, H. D. & Baron, G. (1981) New and revised data on volumes of brain structures. *Insectivores and Primates. Folia Primatology* 35:1–29. [RAB]

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

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

Sur, M., Garraghty, P. E. & Roe, A. W. (1988) Expermentially induced visual projections into auditory thalamus and cortex. *Science* 242:1437–41. [aSRQ]

Sur, M., Humphrey, A. L. & Sherman, S. M. (1982) Monocular deprivation affects X- and Y-cell retinogeniculate terminations in cats. *Nature* 300:183–85. [aSRQ]

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

Swindale, N. V. (1980) A model for the formation of ocular dominance stripes. *Proceedings of the Royal Society of London, B* 208:243–64. [aSRQ]

Székely, G. & Czéh, G. (1976) Organization of locomotion. In: *Frog neurobiology*, ed. L. Llinás & W. Precht. Springer-Verlag. [GS]

Székely, G. & Matesz, C. (1993) The efferent system of cranial nerve nuclei: A comparative neuromorphological study. In: *Advances in Anatomy Embryology and Cell Biology*, vol. 128. Springer-Verlag. [GS]

Thoenen, H. (1995) Neurotrophins and neuronal plasticity. *Science* 270:593–98. [TE]

Tieman, S. B. & Hirsch, S. (1982) Exposure to lines of only one orientation modifies dendritic morphology of cells in the visual cortex of the cat. *Journal of Comparative Neurology* 211:353–62. [aSRQ]

Tononi, G., Sporns, O. & Edelman, G. M. (1992) Reentry and the problem of integrating multiple cortical areas: Simulation of dynamic integration in the visual system. *Cerebral Cortex* 2:310–35. [OS]

(1994) A measure for brain complexity: Relating functional segregation and integration in the nervous system. *Proceedings of the National Academy of Sciences USA* 91:5033–37. [OS]

(1996) A complexity measure for selective matching of signals by the brain. *Proceedings of the National Academy of Science USA* 93:3422–27. [GNR, OS]

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

(1995) Mapping the evolved functional organization of mind and brain. In: *The cognitive neurosciences*, ed. M. Gazzaniga. MIT Press. [rSRQ]

Trehub, S. E. (1976) The discrimination of foreign speech contrasts by infants and adults. *Child Development* 47:466–72. [GD-L]

https://doi.org/10.1017/S0140525X97301587 Published online by Cambridge University Press

Treiber, F. & Wilcox, S. (1984) Discrimination of number by infants. *Infant Behavior and Development* 7:93–100. [GD-L]

Turkewitz, G. & Kenny, P. A. (1982) Limitations on input as a basis for neural organization and perceptual development: A preliminary theoretical statement. *Developmental Psychobiology* 15:357–68. [GH]

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

Turner, R. S. (1994) *In the eye's mind. Vision and the Helmholtz-Hering controversy*. Princeton University Press. [DP]

Ullman, S. (1980) Against direct perception. *Behavioral and Brain Sciences* 3:373–415. [rSRQ]

Uylings, H. B. M., Kuypers, K., Diamond, M. C. & Veltman, W. A. M. (1978) Effects of differential environments on plasticity of dendrites of cortical pyramidal neurons in adult rats. *Experimental Neurology* 62:658–77. [aSRQ]

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

Valiant, L. G. (1984) A theory of the learnable. *Communications of the ACM* 27:1134–42. [aSRQ]

(1991) A view of computational learning theory. In: *Computation and cognition: Proceedings of the first NEC research symposium*, ed. C. W. Gear. SIAM. [aSRQ]

Valverde, F. (1967) Apical dendritic spines of the visual cortex and light deprivation in the mouse. *Experimental Brain Research* 3:337–52. [aSRQ]

(1968) Structural changes in the area striata of the mouse after enucleation. *Experimental Brain Research* 5:274–92. [aSRQ]

(1971) Rate and extent of recovery from dark rearing in the visual cortex of the mouse. *Brain Research* 33:1–11. [aSRQ]

van der Maas, H. L. J. & Molenaar, P. (1992) Stagewise cognitive development: An application of catastrophe theory. *Psychological Review* 99(3):395–417. [MEJR]

(1996) Catastrophe analysis of discontinuous development. In: *Categorical variables in developmental research*, ed. A. Von Eye & C. C. Clogg. Academic Press. [MEJR]

Varela, F. J., Thompson, E. & Rosch, E. (1991) *The embodied mind*. MIT Press. [GM]

Vercelli, A., Assal, F. & Innocenti, G. M. (1992) Emergence of callosally projecting neurons with stellate morphology in the visual cortex of the kitten. *Experimental Brain Research* 90:346–58. [arSRQ]

Volkmar, F. R. & Greenough, W. T. (1972) Rearing complexity affects branching of dendrites in the visual cortex of the rat. *Science* 176:1445–47. [aSRQ]

Wallace, C. S., Kilman, V. L., Withers, G. S. & Greenough, W. T. (1992) Increases in dendritic length in occipital cortex after 4 days of differential housing in weanling rats. *Behavioral and Neural Biology* 58:64–68. [aSRQ]

Wallhäusser, E. & Scheich, H. (1987) Auditory imprinting leads to differential 2-deoxyglucose uptake and dendritic spine loss in the chick rostral forebrain. *Developmental Brain Research* 31:29–44. [JJB]

Walsh, C. & Cepko, C. L. (1988) Clonally related cortical cells show several migration patterns. *Science* 241:1342–45. [aSRQ]

(1992) Widespread dispersion of neuronal clones across functional regions of the cerebral cortex. *Science* 255:434–40. [aSRQ]

(1993) Clonal dispersion in proliferative layers of developing cerebral cortex. *Nature* 362:632–35. [aSRQ]

Wellman, H. M. & Bartsch, K. (1994) Before belief: Children's early psychological theory. In: *Origins of an understanding of mind*, ed. C. Lewis & P. Mitchell. Lawrence Erlbaum. [DE]

Werker, J. F. & Lalonde, C. E. (1988) Cross-language speech perception: Initial capabilities and developmental change. *Developmental Psychology* 24(5):672–83. [GD-L]

Wexler, K. (1996) The development of inflection in a biologically based theory of language acquisition. In: *Towards a genetics of language*, ed. M. Rice. Lawrence Erlbaum. [rSRQ]

Wexler, K. & Culicover, P. (1980) *Formal principles of language acquisition*. MIT Press. [aSRQ]

White, H. (1990) Connectionist nonparametric regression: Multilayer feedforward networks can learn arbitrary mappings. *Neural Networks* 3:535–49. [aSRQ]

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

Williams, R. W. & Herrup, K. (1988) The control of neuron number. *Annual Review of Neuroscience* 11:423–54. [BLF]

Winfield, D. A. (1981) The postnatal development of synapses in the visual cortex of the cat and the effects of eyelid closure. *Brain Research* 206:166–71. [aSRQ]

Witte, S., Stier, H. & Cline, H. T. (1996) *In vivo* observations of timecourse and distribution of morphological dynamics in *Xenopus* retinotectal axon arbors. *Journal of Neurobiology* 31:219–34. [OS]

Wolff, J. R., Laskawi, R., Spatz, W. B. & Missler, M. (1995) Structural dynamics of synapses and synaptic components. *Behavioral and Brain Research* 66:13–20. [OS]

Wong, R. K., Prince, D. A. & Basbaum, A. I. (1979) Intradendritic recordings from hippocampal neurons. *Proceedings of the National Academy of Sciences USA* 76:986–90. [aSRQ]

Wynn, K. (1992) Addition and subtraction by human infants. *Nature* 358:749–50. [GD-L]

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

Xiong, M. & Finlay, B. L. (1996) What do developmental mapping rules optimize? *Progress in Brain Research* 112:350–61. [BLF]

Xiong, M.-J., Pallas, S. L., Lim, S. & Finlay, B. L. (1994) Regulation of the size of axon arbors of retinal ganglion cells by tectal target availability: Mechanisms of compression and expansion of the retinotectal projection. *Journal of Comparative Neurology* 344:581–97. [BLF]

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

Zheng, D. & Purves, D. (1995) Effects of increased neural activity on brain growth. *Proceedings of the National Academy of Sciences USA* 92:1802–06. [rSRQ, DP]

Zheng, D., LaMantia, A.-S. & Purves, D. (1991) Specialized vascularization of the primate visual cortex. *Journal of Neuroscience* 11:2622–29. [DP]

Ziv, N. E. & Smith, S. J. (1996) Evidence for a role of dendritic filopodia in synaptogenesis and spine formation. *Neuron* 17:91–102. [OS]

https://doi.org/10.1017/S0140525X97301587 Published online by Cambridge University Press