Dynamic Plasticity Influences the Emergence of Function in a Simple Cortical Array

JEFF SHRAGER¹ AND MARK H. JOHNSON²

¹University of Pittsburgh, and ²MRC Cognitive Development Unit, London

(Received 1 August 1994; revised and accepted 1 March 1996)

Abstract—In computational experiments with a simplified cortical array we investigated the factors that give rise to the functional organization of the cerebral cortex during brain development. We show that a dynamical spatial modulation of plasticity in the substrate (i.e., a "wave of plasticity") induces higher functional development in the later-developing parts of the cortical array. This result suggests an account of the role that changes in developmental timing may have in the development of different cortical structures. Copyright © 1996 Elsevier Science Ltd

Keywords—Cerebral cortex, Cortical development, Plasticity, Developmental timing, Computational neural models, Hebbian learning, Nerve growth factors.

INTRODUCTION

Our goal in this work is to investigate the factors that give rise to the functional organization of the mammalian cerebral cortex during brain development. Broadly speaking, the adult cerebral cortex can be divided into primary and secondary sensory and motor areas, and "association" areas wherein multiple signals from the sensory areas, and from other association areas, are integrated. Each succeeding large-scale region of cortex may thus be thought of as processing increasing orders of invariants from the stimulus stream, and passing either the invariant information extracted from the stream, or the residual information once the invariant is extracted, forward to other regions of the brain (e.g., Oram & Perrett, 1994; Tanaka et al., 1993). A common conception is that of a cascade of filters, processing and separating stimulus information in series up toward the integration areas where it is combined, and eventually reaching the frontal association areas. Feedback from higher to lower systems may act to select or tune this process (e.g., Mangun & Hillyard, 1991).

One striking aspect of the development of the cerebral cortex is the initial overproduction and subsequent loss of neural connections (synapses), resulting in the relatively sparsely interconnected final functional architecture. In the macaque, our nearest primate model, the period of "exuberance" in synaptogenesis begins before birth and is in furious bloom at birth, reaching its peak in the first few postnatal weeks [Bourgeois (1993) provides a concise review]. In some cases the patterns of synaptic loss in the cortex, or in thalamocortical connections, has been shown to be correlated with experience-dependent activity (Stryker & Harris, 1986; Roe et al., 1990), as may be the periods of cortical plasticity (Harwerth et al., 1986). This process of overproduction of synapses and subsequent stimulus-directed pruning of the dendritic arbor is thought to be key in cortical ontogeny.

The neuroanatomical analyses of Pandya and Yeterian (1990), the electrophysiological observation of Thatcher (1992), the PET observations of Chugani et al. (1987), and the experiments of Harwerth et al. (1986) all suggest a general developmental dynamic in which, grossly speaking, the locus of maximum neural plasticity begins in the primary sensory and motor areas and moves towards the secondary and parietal association areas and finally to the frontal regions. The most direct evidence for such a "wave of
plasticity” is due to Bourgeois (1993), who has demonstrated by direct measurement in the macaque that thinning of the synaptic array is not observable in the frontal regions until around the animal’s puberty. Such a wave of plasticity, combined with the pruning process described above, may enable each region of cortical material to become attuned to pick up the simplest (lowest order, least abstract) regularities available in the input to that region. The initially undifferentiated regions of cortex closest to the direct stimulation (the “earliest” regions) will become attuned to the first order of information from the stimulus signal, in the process giving these early regions form and function. The next spatially adjacent regions of cortex—receiving their input from both subcortical afferents and the earlier cortical regions—will then become attuned to a slightly higher order of regularity, and so on.1

In this paper we use a computational model of a simplified cortical array to explore the way in which the suggested “wave” of plasticity can lead to the differential distribution of function over the cortical surface. We extend a computational model of cortical parcelation due to Kerszberg et al. (1992) to introduce such a wave, and demonstrate that this manipulation induces the development of higher order functionality in latter-developing areas of the simulated cortex.

COMPUTATIONAL FRAMEWORK

Numerous computational models of brain development demonstrate the formation of regions in an initially over-wired competitive network, and some show that the statistical structure of afferent activity can bias the process so that regions with one sort of characteristic activity are differentiated from regions with different characteristic activity (e.g., Miller et al., 1989; Durbin & Mitchison, 1990; Grajski & Merzenich, 1990; Kohonen, 1990; Montague et al., 1991; Jacobs & Jordan, 1992; Kerszberg et al., 1992). We take as our starting point the model of Kerszberg et al. (1992), which we shall refer to as the KDC model. This model differs from most others in two important ways. First, in the KDC model the means of dendritic selection and elimination depend upon the activity-dependent diffusion of a neurotrophic substance (also Montague et al., 1991). It is assumed that there is a fixed resource of trophic factor (TF). A Hebbian association rule directs the transportation of TF from interneural space to synapses. Meanwhile, as time progresses, synapses that receive less TF eventually atrophy, whereas synapses that receive more TF eventually reach stability. Thus the period of plasticity depends upon the transportation of TF, and may differ from one part of the network to another in accord with the activity-dependent diffusion of this factor.

A second important feature of the KDC model is that Kerszberg et al. not only described the targeting of afferents onto the cortex, but also were able to describe in fairly general terms the local function served by simulated neural units. These researchers employed a method of analysis by which they were able to assign one of the 16 possible binary logic functions (and one “indeterminate” function) to each neuron in the network both before and after execution of the simulation. This enabled them to describe the progression from the mostly indeterminate functionality of the richly arborized early cortex, to a more specifically functioning, “post-development” cortex.

THE MODEL

Here we provide an overview of the model. A detailed presentation appears in the Appendix.

The model consists of a 30-by-30 square array of units, called “neurons” by Kerszberg et al., which we shall call the “cortical array,” and two one-bit afferent units: A and B. Any unit can take on an activation value of 1 or 0. We refer to taking on a 1 value as “firing,” and taking on a 0 value is “not firing.” Cortical units are either excitatory or inhibitory, and project excitatory or inhibitory links, respectively, to other units in the cortical array. The two afferent units project both excitatory and inhibitory links into the array, which connect to excitatory and/or inhibitory cortical units (Figure 1). When a link connects to a unit, it is said to make a synapse upon that unit, thus connecting a presynaptic unit with a postsynaptic unit. The initial

---

1Fixed, predetermined, intra-cortical connectivity will influence which regions are counted as “spatially adjacent”. Our hypothesis does not discount this possibility, but we hold that at least within large uniform regions of the cortex, such as the primary sensory areas, there is little of this prewiring.

---

FIGURE 1. A diagrammatic representation of the model. A and B are each one bit (0 or 1) and project uniformly and randomly to the 30-by-30 array of excitatory (+) or inhibitory (−) units, which in turn randomly project to one another in accord with principles described in the Appendix. Note that there are exactly 900 units (30 × 30), although only four are shown in the diagram.
wiring of the array is random, and units receive a large number of links (an average of 30 from other cortical units and 10 from the afferents).

Units have a firing threshold and a pool of "trophic factor" (TF). Links between units also have a pool of TF, and take on one of three discrete states: labile (i.e., plastic), stable, or dead. Initially, all links begin in the labile state, and the pools of trophic factor are initialized.

The simulation proceeds in a sequence of "training phases", each of which consists of a series of "settling cycles". Entering into each cycle, every cortical unit has an activation value of 0 or 1 according to the history of activation that has reached it from the afferents or from other units. The afferents take on values in accord with a statistical regimen that is particular to each experiment. These are described in the sections below that discuss the experiments.

On each settling cycle, each cortical unit sums the activation coming to it (from afferents or from other cortical units) along excitatory links, and subtracts the activation from inhibitory links. It then makes a decision about what value (1 or 0) it will take into the next cycle, depending upon whether this sum reaches the unit-specific firing threshold.

After every five settling cycles, training takes place, the afferent values are set to their next value (for the particular experimental regimen), and the cortical unit activations are reset to 0. (Note that the afferents maintain their given values through all five cycles between trainings, thus the term "settling cycles".) In a 500 cycle simulation, there will be 100 training phases—one after every five settling cycles.

Training consists of moving a quantity of trophic factor from the pools associated with the cortical units to the incoming links according to a Hebbian association rule, deciding whether each link will change state (from labile to dead or stable) and updating the firing thresholds of each unit. The Hebbian rule moves TF toward links that connect units that fire in agreement with one another. That is, an active excitatory link (one from a firing excitatory unit, or a specifically excitatory link from a firing afferent) collects TF when the unit upon which it makes a synapse comes above threshold (i.e., fires), and an inhibitory link (one from a firing inhibitory unit, or a specifically inhibitory link from a firing afferent) collects TF when the unit upon which it makes a synapse does not come above threshold (does not fire).

The probability that a link changes state from labile to stable increases as the amount of TF collected by the link increases. The probability that a link changes state from labile to dead increases with the length of time it has remained labile. Dead links do not transmit activation and do not participate in the simulation. Stable links transmit activation, but do not change TF values (i.e., do not undergo training); only labile links change through training. The new threshold of a given cortical unit is determined in accord with the amount of TF collected in the links which target that unit (synapse onto it). (The Appendix gives greater formal detail to all of the processes described above.)

The resulting dynamic of the network is that links which lead to or are associated with appropriate cortical unit firings (i.e., non-firings for inhibitory links, firings for excitatory links) are rewarded and eventually stabilize, and those which are not so associated eventually die from lack of trophic factor. Since the network is initially randomly and heavily wired with both afferent-to-cortex, and intracortical connections, the eventual pattern of wiring, and functionality of the cortical units is determined primarily (in the original KDC model) by the structure of the stimuli in the A and B afferents and the afferent fiber cell targeting.

Functionality of cortical units in the KDC model is determined at any point in the model run by trying out each possible combination of afferent values, allowing settling to take place, and then recording the resulting values of each cortical unit. Since there are two afferents (A and B), each of which can be set to either zero or one, each cortical unit can potentially respond in one of 16 ways, representing the 16 logical functions of two bits: always off, always on, on with A only, on with A and B only, on when A is on but not B, etc. The procedure just described is repeated a number of times, and units that do not consistently produce a single logical function from among the 16 are said to have an indeterminate function. Thus, there are 17 possible functions for each cortical unit: the 16 logical functions, and the indeterminate function.

It is important to note that the functionality of a block of cortical array, by virtue of its connectivity, is of far greater complexity than these 16 logical functions. Combinations of functions, and their time-course dynamics, presumably result in extremely complex functionality of the brain that is not most easily nor most naturally described in boolean terms. However, it is incredibly difficult to determine this very high-order functionality. Therefore, in the present analyses we will speak primarily of the general "order" of functionality in terms of the representation in the cortical array of zeroth, first, or second order logical functions. The zeroth order functions include "false" and "true"; they take no account of the afferent input values. The first-order functions are "A", "B", "¬A", and "¬B", which rely only upon one or the other afferent value. The second-order functions are all others, which are some function of both A and B. This reflects the connectivity and afference upon one unit, but not upon or within a number of cortical units. Determining the contribu-
regimes were similar to those used by Kerszberg et al., we began by replicating their demonstration that the model was able to move from an architecture of mostly indeterminate functions, to one of mostly determinate functions, and that it was able to self-organize into a center-surround sort of functional architecture consisting of tightly grouped mutually excitatory sets of units, interconnected by fewer, mostly inhibitory longer connections.

We manipulated the differential correlations between the afferents inputs by modifying the KDC model so that the A afferent always produced a simple oscillating pattern: 101010.... The B afferent then could take on a correlation with respect to A (given by the parameter: \(P_{AB}\), see Appendix) ranging from 0.5 (randomly correlated) to 1.0 (fully correlated). In the case of \(P_{AB} = 0.5\), B could not be predicted from A, although A remained in its oscillating pattern. This was experimental case C. In the case of \(P_{AB} = 1.0\), if \(A = 1\), \(B = 1\); if \(A = 0\), \(B = 0\) as well. This was experimental case D. We expected that in the fully correlated case we would see a greater proportion of functionality in the array responding to \(A\) or \(B\) (\(A \lor B\)), and correspondingly less functionality responding to afferent A or afferent B alone (\(A, B, \sim A, \sim B\), etc.). That is, the array units should become responsive to the combination of \(A\) or \(B\), but not to both (\(A\) and \(B\)) or either (\(A, B\), etc.), because these afferents carry just the same information.

On the other hand, in the uncorrelated case we would expect to see more functions that represent relations between \(A\) and \(B\) separately, and which respond to \(A\) or respond to \(B\), essentially assigning some of the cortical material to one afferent and some to the other, since they carry different information.

Figure 3 compares the post-training functional density of more complex patterns of connectivity, as are found in the architectures of real brains, is a goal of our future work.

We demonstrate this analytic method by considering the initial state of the cortical array, determined by counting the number of units that are determined to have each of the 17 possible functions (including the indeterminate function). Figure 2 depicts this initial distribution, which is more or less the same for all of the experiments in this paper, depending only upon random number generation. Note the large proportion of indeterminate (??) units, and ones that seem not to be responding to any input (false). Also, there is about an even number of units representing the simplest logical functions: \(A, B, A\) and \(B, A\) or \(B\).

It is worthy of note that the large number of units responding with “false” does not mean that these units are disconnected. Rather, they may be connected through afferent and intracortical connections and in such a way that the signals from the afferents contradict one another. So, if a single unit has an inhibitory and excitatory afferent, either directly or indirectly, from each afferent, it will never respond, and so will appear to be dead, thus recording as false. Thus, as we shall see, the number of such unresponsive units can actually decrease as links in the system die. Furthermore, since the training procedure updates unit firing thresholds in accord with the weight of afferent links [Appendix, eqn (A.4)], more units may respond as the simulation proceeds.

**BASELINE REPLIATION (C AND D): CORRELATIONS IN AFFERENT STIMULUS**

For purposes of later comparison, and in order to determine that the parameter values and training conditions were similar to those used by Kerszberg et al., we began by replicating their demonstration that the model was able to move from an architecture of mostly indeterminate functions, to one of mostly determinate functions, and that it was able to self-organize into a center-surround sort of functional architecture consisting of tightly grouped mutually excitatory sets of units, interconnected by fewer, mostly inhibitory longer connections.

We manipulated the differential correlations between the afferents inputs by modifying the KDC model so that the A afferent always produced a simple oscillating pattern: 101010.... The B afferent then could take on a correlation with respect to A (given by the parameter: \(P_{AB}\), see Appendix) ranging from 0.5 (randomly correlated) to 1.0 (fully correlated). In the case of \(P_{AB} = 0.5\), B could not be predicted from A, although A remained in its oscillating pattern. This was experimental case C. In the case of \(P_{AB} = 1.0\), if \(A = 1\), \(B = 1\); if \(A = 0\), \(B = 0\) as well. This was experimental case D. We expected that in the fully correlated case we would see a greater proportion of functionality in the array responding to \(A\) or \(B\) (\(A \lor B\)), and correspondingly less functionality responding to afferent A or afferent B alone (\(A, B, \sim A, \sim B\), etc.). That is, the array units should become responsive to the combination of \(A\) or \(B\), but not to both (\(A\) and \(B\)) or either (\(A, B\), etc.), because these afferents carry just the same information.

On the other hand, in the uncorrelated case we would expect to see more functions that represent relations between \(A\) and \(B\) separately, and which respond to \(A\) or respond to \(B\), essentially assigning some of the cortical material to one afferent and some to the other, since they carry different information.

Figure 3 compares the post-training functional density of more complex patterns of connectivity, as are found in the architectures of real brains, is a goal of our future work.

We demonstrate this analytic method by considering the initial state of the cortical array, determined by counting the number of units that are determined to have each of the 17 possible functions (including the indeterminate function). Figure 2 depicts this initial distribution, which is more or less the same for all of the experiments in this paper, depending only upon random number generation. Note the large proportion of indeterminate (??) units, and ones that seem not to be responding to any input (false). Also, there is about an even number of units representing the simplest logical functions: \(A, B, A\) and \(B, A\) or \(B\).

It is worthy of note that the large number of units responding with “false” does not mean that these units are disconnected. Rather, they may be connected through afferent and intracortical connections and in such a way that the signals from the afferents contradict one another. So, if a single unit has an inhibitory and excitatory afferent, either directly or indirectly, from each afferent, it will never respond, and so will appear to be dead, thus recording as false. Thus, as we shall see, the number of such unresponsive units can actually decrease as links in the system die. Furthermore, since the training procedure updates unit firing thresholds in accord with the weight of afferent links [Appendix, eqn (A.4)], more units may respond as the simulation proceeds.

**BASELINE REPLIATION (C AND D): CORRELATIONS IN AFFERENT STIMULUS**

For purposes of later comparison, and in order to determine that the parameter values and training
density in these cases. We see that, as predicted, the functions that assign cortical material to one afferent and not the other appear in the case of lesser correlation (C), whereas in case D (fully correlated afferent stimulation) a greater portion of cortical material responds to both afferent signals (A or B). Note that the complex functions: A and ~B, and B and ~A, which appear most in the uncorrelated case (C) result from the “pushing apart” of the A and B afferent signals, and thus represent the most specific recognition of difference in signal that may be observed in the array units at this level of analysis.²

EXPERIMENT 1 (I): SPATIAL TARGETING OF AFFERENT PROJECTIONS

The baseline control study demonstrates that our version of the KDC model can produce differential patterns of function over the cortical array in accord with the stimulus structure of the environment. Our goal is to go beyond this to predictably manipulate the level of expressed function and the locations of these functions in the cortical array through a spatially-distributed differential learning rate. In experiment 1, described in this section, we demonstrate that higher-order functions can arise, in the KDC model, in regions of the array that are the analog of association areas; that is, the regions in which the information arising from one afferent projection meets the information arising from the other afferent projection. This provides the first step toward our goal. In the next section (experiment 2) we dynamically induce this differential targeting by introducing a spatially-distributed differential learning rate, and demonstrate that this manipulation enables us to predict that the order of function increases with time and space through learning.

We modified the KDC model so that each afferent projects only to two-thirds of the cortical array. The left 20 units (all the way from top to bottom—that is, a rectangle of 20 units width and 30 units height) receive input from afferent A, and the right 20 units (ditto) receive inputs from afferent B. Note that this produces an overlap area of the middle third of the array, wherein units receive input from both A and B, as in the previous studies.

In addition, we replaced the spatially uniform inputs of the control study with horizontally-striped inputs wherein each stripe was 3 units in width and they are out of register from one another by 3 units. This topology is depicted in Figure 4.

Figures 5 and 6 depict the results of training with this topology under the standard KDC training regime. In these figures we plot on the dependent axis the function density for each row of the array. First, note the unsurprising result, depicted in Figure 5, that the units to which the A afferent contributed most positive targets are mostly A-responsive units, and that the opposite is the case for the alternating stripes. Second, note that almost all of the functionality is occupied by A and B basic functions. Some of the additional functional density is occupied by the functions A and ~B (correlated

²Because the settling dynamics are enforced by the model to terminate at approximately 500 cycles (see Appendix), we do not here report or consider the details of stabilization rate for these experiments. Also left unreported in the present paper are the details of link lengths, beyond noting that in almost all cases we approximately replicated the general results of a tight bundle of excitatory synapses and occasional longer inhibitory synapses reported by Kerszberg et al. This deserves further detailed treatment but is not a central concern of the present paper.
cortical array. This is done by multiplying the transmission of TF by the amplitude of the slowly moving gaussian wave at each point in the array. Suppose, for instance, that the TF modulation wave (hereafter simply referred to as the “wave”) moves with a rate of 0.25 columns during each training phase, and reaches a peak value of 1.0. During the first training phase the following $\alpha_c$ modulation vector will be utilized:

$$\alpha_c = \{1.0, 0.86, 0.77, 0.66, 0.53, \ldots, 0.0, 0.0\}.$$

That is, TF transmission at location 1 of the array takes place normally ($\alpha_1 = 1.0$) whereas TF transmission at location 2 is reduced to 86% of what would have been moved ($\alpha_2 = 0.86$), etc. On the next training phase, the wave moves to the right a small amount. Since in this example it takes four training phases for the wave to move one full column to the right, the above vector will read approximately as follows at training phase 4 (real cycle 20 with 5 settling cycles between each training phase):

$$\alpha_c = \{0.86, 1.0, 0.86, 0.77, 0.66, \ldots, 0.0, 0.0\}.$$

The progress of the wave thus modulates the transmission of TF, leading to a dynamic plasticity in the cortical array; leftward columns are plastic early on whereas they lose their plasticity early as well. The more rightward columns do not become plastic until later on, but are plastic toward the end of the run when most of the neurons are finally reaching stabilization or death. Note that this manipulation affects only the amount of TF moved during training, not the total amount of TF in the array nor any other simulation parameter.

In certain regimes of this wave, we can expect to observe specific phenomena. First, we expect in general more synapse death because most synapses will receive proportionally less access to TF during a simulation of the same length as those we have already seen. More interestingly, under certain regimes of wave propagation, we expect to observe a tendency toward the development of higher-order functions in later parts of the cortical array.

The original Kerszberg et al. model had a uniform rule for TF transportation [$\alpha_c = 1.0$ in eqn (A.3) of the Appendix, for all $c$], dependent only upon synapse activity, as has been explored in the preceding studies. In order to explore the effect of the “wave plasticity” theory we introduce a dynamic gaussian spatial modulation of the rate of diffusion across the

---

3. Kerszberg notes (personal communication) that we could have adjusted certain parameters in the model in order to retain the cells longer in our wave simulations. We chose not to do this in order to simplify comparison between experiments.
fixed in its function relatively early, as the wave moves on to NEXT. Now, however, NEXT is receiving input that, in addition to the input coming from \(A\) and \(B\) afferents, includes the combined functions fixed by the earlier plasticity in LEFT. Thus, NEXT has, in effect, three afferents: \(A\), \(B\), and \(\text{LEFT}\).

This account will depend upon a number of parameters of the model, but most especially the wave propagation rate. Therefore, we have explored two regimes of wave propagation. In experiment J the wave propagates at a rate of 0.25 columns per training phase. This brings the wave across the array once over the entirety of the simulation run (generally about 500–550 total cycles, or about 100 training phases, see Appendix). In experiment M we consider a wave propagation rate of three times that of the rate in experiment J. In both of these cases the correlation between \(A\) and \(B\) afferent activity is 0.5, as in study C, above. That study will thus act as a baseline for the present experiments.

Let us begin by considering the percentage overall function density resulting from these wave studies, summarized in Figure 7. As expected, comparing with study C (Figure 3), we see a small amount more false functions, representing synapses that have become or remained functionally disconnected. Otherwise, relatively speaking, the data for experiments J and M look very much the same, and also much the same as case C (Figure 3). Thus, from this chart, with the exception of the slightly larger amount of synapse death, as expected, we appear to have little or no effect of the TF modulation wave, regardless of its rate.

A more sensitive analysis, however, is required in order to make sense of the present study. We thus consider “slices” of the array. Each slice is a small number of columns (two columns in all the cases that we shall see here, though this number was chosen arbitrarily), analyzed as though it was a separate array, and then adjacent slices are compared with one another. This method enables us to produce a plot of the dynamic effect of the wave on the array along the path of its progress.

Let us begin by applying this analysis to the initial state of the matrices, as was represented in Figure 2. We see, in Figure 8, the same data as were presented in Figure 2, though spread out into 15 two-column slices. There is no significant difference between any of the lines labeled “all others” in Figure 8.

We now consider the same slice analysis applied to the non-wave, post-training condition of study C. A slice reanalysis of these same data is presented in Figure 9. For convenience we divide the functions into first-order functions, which depend only upon one

**Experiments J (slow wave, n=3) v. M (fast wave (x3), n=4)**

![Figure 7](image-url)  
**FIGURE 7.** Comparison of overall functional density difference between slow and fast wave cases.

**FIGURE 8.** Slice (2 units wide) functional density of the initial state of the array (as Figure 1), \(n = 7\). Density is computed in pairs of columns of the array. Since the entire array is 30 units in width, there are 15 two-column slices. None of the “all others” functions are statistically different.

**Experiment C, Functional Density (Slice Width 2)**

![Figure 9](image-url)  
**FIGURE 9.** Slice functional density in non-wave case (\(C : P_{AB} = 0.5, n = 5\)). First-order functions are \(A\), \(\sim A\), \(B\), and \(\sim B\). Second-order functions are: \(B\) and \(\sim B\), \(A\) and \(\sim A\), \(A\) or \(B\), \(A\) and \(B\), \(A = B\), \(A = B\), \(A = B\), \(A = B\), and \(A\) v \(B\). The zeroth-order functions, true and false, are not represented here. Most of the missing weight (to 100%) is represented by false. True and Indeterminate contribute essentially no weight. The depicted curves are not significantly different by a pairwise t-test.
variable (A or B) and second-order functions, which depend upon both A and B afferent values. The first-order functions are A, ~A, B, and ~B. The second-order functions are: B and ~A, A and ~B, A xor B, ~A and B, A = B, A > = B, A < = B, and A v B. The zeroth-order functions, true and false, are not represented in Figure 9; most of the missing weight (to 100%) is represented by false. True and indeterminate (?) contribute very little weight. A study of Figure 9 reveals a complex interplay in fairly high density range (about 25%) of the first-order and second-order functions. However, these curves are not significantly different, and therefore do not afford further analysis.

Consider, now, the slice analysis applied to experiment J, with the same parameters as in study C, but with a TF modulation wave that propagates at the rate of 0.25 rows per training phase (Figure 10). In contrast to the analysis depicted in Figure 9, the first and second-order function density curves presented here are significantly different from one another, thus affording further analysis in order to assess our theoretical predictions. The density of second-order functions beyond the first slice is greater than the density of first-order functions. Also, with the exception of the tail of the graph, where most synapses are dying off from lack of TF, our hypothesis predicts that these curves will stand in anticorrelation to one another because each subsequent region is taking on either a generally lower or generally higher functionality, displacing the other sort of functionality from the cortical units. Indeed, the first 7 points are anticorrelated. A simple correlation on slices 1-7 has the small negative correlation of -0.339. Since there is a dynamic of dying synapses lending a confusing overall curvature to this result, we can approximately co-vary that out by multiplying the density of false functions by each value here. This corrected correlation is much stronger in the predicted direction, at -0.656 on these first 7 slices. (The influence of this effect of synaptic death and the sense of co-varying according to the false function density may be seen more clearly in Figure 11.)

Next we consider the latter experiment, M, in which the propagation rate of the wave has been tripled from the baseline rate of the previous experiment. Recall that we did not see any very significant change in the overall functional density by casual inspection of Figure 7. The slice analysis, depicted in Figure 11, reveals a very different picture. Again, the first- and second-order functional density curves are significantly different, but this time the mean values are inverted, so that whereas in the slow-wave case the second-order functions were emphasized, in this fast-wave case it is the first-order functions which are emphasized. (The corrected anti-correlation is somewhat reduced to -0.413). Note also that the absolute values of the second-order function density are much smaller than those of the former case (J, Figure 10). What appears to have happened here is that speeding up the wave (M) has had the effect of spreading out the higher-order

\footnote{It is important to note that although it is generally the case that some functions will always be anticorrelated with others since the total functional density is fixed in a zero-sum game, it is \textit{not} the case that the density of any group of functions will be anticorrelated. Here we have preselected the specific groups of functions to compare, from the numerous ways in which the 15 functions (excluding the FALSE and indeterminate functions) can be subdivided. Note also that until some method is available whereby we can analyze complex functions of a cortical array, we can only gauge complexity in terms of the two levels of abstraction measured here. It is possible that we have obtained the more general result of higher-order functions, but we are unable to see}
functions over the array, as well as slightly increasing the total number of high-order (and indeterminate) functions relative to the slow-wave case (J).

DISCUSSION

We have extended a computational model of cortical development to explore the effect of a "wave" of plasticity in the development of function in a cortical array. We hypothesized that this manipulation would produce a predictable pattern of function (in simple terms) in the array, wherein lower-order functions appear near the edges of the array, where the wave begins, and the density of higher-order functions increases as the wave moves across the array. We found empirical support for our prediction in the comparison of the control case (C) with the wave manipulation (J). This result lends support to our general hypothesis.

These studies are steps toward an account of the way that the differential functional architecture of the cortex might arise in development as an emergent result of the combination of organized stimulus and a neurotrophic dynamic. Our next analytic step will be to develop a formal understanding of more complex functions than are represented by the 16 simple functions analyzed by Kerszberg et al., beyond the higher- v. lower-order analysis that we have introduced. In order to do this we need a method for the analysis of the network of connectivity left by the training process, because more complex functions will result from combinations of the 16 simpler ones. Unfortunately, although general observations can be made of the patterns of connectivity left by the training process, we have yet to discover a good means for such analysis.

Our next empirical step will entail using stimuli that have greater information density along more dimensions, and in which the stimuli lie along those dimensions in organized ways. At the same time we will change the architecture to enable several bands of communication to develop between regions of the cortical array. These extensions will enable us to represent points and relationship between points in the high-dimensional space of complex stimuli, such as those relations found in retinotopic maps of space in mammals (Udin & Fawcett, 1988; Linsker, 1990) and spatial maps of frequency in bats (Suga, 1990). We predict that in the presence of such stimuli the wave manipulation will lead our simulated cortical array to spatially represent increasingly abstract analyses of the stimulus space.

REFERENCES


Stryker, M. P., & Harris, W. A. (1986). Binocular impulse blockade prevents the formation of ocular dominance columns in cat visual cortex. *Journal of Neuroscience*, 6, 2117–2133.


Tanaka, K., Fujita, I., Kobayashi, E., Cheng, K., & Ito, M. (1993). Serial processing of visual object-features in the posterior and
The unit then makes a decision about what activation value (1 or 0) it will take into the next cycle, depending upon whether this sum reaches the unit's threshold, \( T \). We refer to units that take on an activation of 1 as "firing," and those taking on an activation of 0 as "not firing". Simulation proceeds until all links' state is either stable or dead.

Training

Training changes the unit thresholds and link states through activity-dependent movement of TF. In the present experiments, five cycles take place between each change in the afferent values. After the fifth cycle, the network is "trained" by updating the link and unit pools of TF and link states, and the unit threshold values. We therefore refer to the simulation cycles as "settling cycles," and to each group of five settling cycles as one "training phase." In a 500 cycle simulation, there will be 100 training phases—once after every five settling cycles.

The afferent values (\( A \) and \( B \)) change after training (that is, every five settling cycles) in accord with a regimen that differs for each experiment. The input from afferent \( A \) is a continuous alternating pattern, as: 1010101... The inputs from afferent \( B \) are generated in a stochastic manner, where the correlation between the \( A \) and \( B \) stimuli are given by \( P_{AB} \), the probability that \( A \) and \( B \) have the same value. With \( P_{AB} \) set to 1.0, \( B \) will give precisely the same input as \( A \). With \( P_{AB} \) set to 0.0, the link values are fully anti-correlated, that is: 180° out of phase; with \( P_{AB} \) set to other values between 0 and 1, \( B \) has the given probability of presenting the same value as \( A \).

Modification of Unit Threshold, and Synapse TF and State

At each training phase (every 5 cycles) a quantity of TF is moved from the pools associated with the post synaptic cortical units (\( \mu_i \)) to pools in the incoming links (synapses: \( \sigma_{ij} \)) according to a Hebbian association rule. The Hebbian rule favors the links that connect units that fire in agreement with one another. \( \delta_{thre} = 1 \) is defined as follows: (1) for excitatory terminals, \( \delta_{thre} = 1 \) whenever the pre- and post-synaptic neurons fire simultaneously, \( \delta_{thre} = 0 \) otherwise; (2) for inhibitory terminals, \( \delta_{thre} = 1 \) whenever the post-synaptic neuron fires but the pre-synaptic (inhibiting) neuron does not fire. Otherwise, \( \delta_{thre} = 0 \).

The concentration of TF in an incoming link, \( \sigma_{ij} \), changes according to the following equations:

\[
\Delta \mu_i = -\delta_{thre} \alpha \mu_i \quad \text{(A.2)}
\]

\[
\Delta \sigma_{ij} = -\sigma \sigma_{ij} \quad \text{(A.3)}
\]

These express the "retrograde diffusion" of TF from the post-synaptic unit (\( \mu_o \)). \( k_5 \) is the diffusion constant. Initially \( \sigma_{ij} = 0 \) for all links between any units \( i \) and \( j \), and \( \mu_i \) is 20.0 for all units \( i \). In the present simulations, the parameters for eqns (A.2) and (A.3) are: \( k_5 = 0.01 \). \( \alpha \) is a scaling factor which is 1.0 (for all \( e \)) for experiments that do not involve trophic waves, but which changes in accord with the spatial position \( e \) in wave experiments J and M (see the relevant sections of the body of the paper).

In words: an excitatory link (one from an excitatory unit, or a specifically excitatory link from an afferent) collects TF when both the pre-synaptic unit, and the unit upon which it makes a synapse fire, whereas an inhibitory link (one from an inhibitory unit, or a specifically inhibitory link from an afferent) collects TF when the pre-synaptic unit fires but the unit upon which it makes a synapse does not fire. That is, when the link has contributed appropriately to the activation of the target unit.

The new threshold of a unit, \( T_i \), is determined in accord with the amount of TF collected in the links which make synapses upon that unit:

\[
\]
\[ T_i = T_{i-1} + T_i \sum_{j \in S_i} W_{ij} \]  \hspace{1cm} (A.4)

where \( T_i = 7.0 \), \( T_{i-1} = 0.2 \), and \( W_{ij} \) is defined in eqn (A.1), above.

All links begin in the labile state. The probability that the state of a link changes from labile to stable is proportional to the amount of TF collected by the link (\( \sigma \)). A synapse is essentially guaranteed to stabilize when its \( \sigma \) reaches 1.0 through a function that peaks rapidly as the amount of TF passes through 1.0. Precisely:

\[ \rho_{\text{stable}} = \frac{1}{1 + \exp(-15(\sigma - 1))}. \]  \hspace{1cm} (A.5)

The probability of a link entering the dead state is proportional to the time it has been labile. Similarly to eqn (A.5), this too is modulated by a function that peaks rapidly as the simulation passes through 500 cycles. Precisely:

\[ \rho_{\text{dead}} = \frac{1}{1 + \exp(-0.02(t - 500))} \]  \hspace{1cm} (A.6)

where \( t \) is the number of cycles through which the simulation has run.

As a result of these relationships, all synapses are highly likely to have either stabilized or died by the end of about 500 simulation cycles (100 training phases).

**Summary**

As a result of these procedures and relationships, links that lead to, or are associated with appropriate cortical unit firings (i.e., non-firings for inhibitory links, firings for excitatory links) are rewarded and eventually stabilize, and those which are not so associated eventually die from lack of TF. The final pattern of wiring and functionality of the cortical units is thus determined primarily by the structure of the stimuli in the \( A \) and \( B \) afferents and the afferent fiber cell targeting.