Cerebral Cortex Advance Access published May 13, 2004

Visual Binding Through Reentrant Connectivity and Dynamic Synchronization in a Brain-based Device

Anil K. Seth, Jeffrey L. McKinstry, Gerald M. Edelman and Jeffrey L. Krichmar

The Neurosciences Institute, 10640 John Jay Hopkins Drive, San Diego, CA 92121, USA

Effective visual object recognition requires mechanisms to bind object features (e.g. color, shape and motion) while distinguishing distinct objects. Synchronously active neuronal circuits among reentrantly connected cortical areas may provide a basis for visual binding. To assess the potential of this mechanism, we have constructed a mobile brain-based device, Darwin VIII, which is guided by simulated analogues of cortical and sub-cortical areas required for visual processing, decision-making, reward and motor responses. These simulated areas are reentrantly connected and each area contains neuronal units representing both the mean activity level and the relative timing of the activity of groups of neurons. Darwin VIII learns to discriminate among multiple objects with shared visual features and associates 'target' objects with innately preferred auditory cues. We observed the co-activation of globally distributed neuronal circuits that corresponded to distinct objects in Darwin VIII's visual field. These circuits, which are constrained by a reentrant neuroanatomy and modulated by behavior and synaptic plasticity, are necessary for successful discrimination. By situating Darwin VIII in a rich real-world environment involving continual changes in the size and location of visual stimuli due to self-generated movement, and by recording its behavioral and neuronal responses in detail, we were able to show that reentrant connectivity and dynamic synchronization provide an effective mechanism for binding the features of visual objects.

Keywords: brain-based device, Darwin VIII, reentrant connectivity, visual binding, visual object recognition

Introduction

Animals effortlessly combine the various attributes of visual stimuli to form coherent perceptual categories and discriminate among multiple objects in a scene. Yet the visual brain is functionally segregated: Separate cortical regions are specialized to respond to features such as shape, color and object motion, and no single region has superordinate control. This poses the so-called binding problem (Treisman, 1996): how do these functionally segregated regions coordinate their activities to link various features of individual objects while distinguishing among different objects? Most proposed mechanisms for solving the binding problem fall into one of two general classes: (i) binding through the influence of attentional processes, executive mechanisms, or superordinate maps (Shadlen and Movshon, 1999; Shafritz *et al.*, 2002); or (ii) binding through the selective synchronization of dynamically formed neuronal groups (Edelman, 1987, 1993; Gray, 1999; Singer, 1999).

The appeal to mechanisms of attention often focuses on parietal or frontal areas, the operations of which are distant from early stages of sensory processing (Shadlen and Movshon, 1999). It has been suggested that these areas bind and select objects in a visual scene by means of an executive mechanism, such as a spotlight of attention that combines visual features at specific locations in space (Treisman, 1998; Shafritz *et al.*, 2002). In this mechanism, strict limits on the number of objects that can be simultaneously bound arise from the limited capacity of the attentional system (Pashler, 1999).

Advocates of neural synchrony, by contrast, suggest that binding is an automatic, dynamic and pre-attentive process arising from low-level neural dynamics. For example, the linkage of neuronal groups by reentry, where reentry refers to the recursive exchange of signals across multiple, parallel and reciprocal connections (Edelman, 1993), can lead to selective synchronization (Tononi *et al.*, 1992; von der Malsburg and Buhmann, 1992; Knoblauch and Palm, 2002a,b). This synchronized activity among neuronal groups can form coherent circuits corresponding to perceptual categories (Sporns *et al.*, 1991). A fundamental question for proponents of neural synchrony is how such emergent functional circuits contribute to an organism's adaptive behavior, especially in situations that require preferential behavior towards one object among many in a scene.

A previous computational model of visual binding based on reentry (Tononi *et al.*, 1992) learned to make simulated saccades to preferred visual objects. This model included nine simulated visual cortical areas, as well as reward and motor systems, and its performance showed that reentrant connections facilitated the recognition and discrimination of multiple visual objects. Despite showing the capabilities of reentrant circuits, the model had several limitations. For example, the stimuli used were taken from a limited set and were of uniform scale. Furthermore, its behavior did not emerge in a rich and noisy environment of the kind confronted by behaving organisms.

In this paper, we address these limitations by embedding a simulated nervous system in a real-world device capable of engaging in rich exploratory and selective behavior. We describe the construction and performance of Darwin VIII, the latest in a series of brain-based devices (Edelman *et al.*, 1992; Almassy *et al.*, 1998; Krichmar *et al.*, 2000; Krichmar and Edelman, 2002). In Darwin VIII, synchronously active neuronal circuits are dynamically formed as the device engages in visually guided behavior that includes a discrimination task. Darwin VIII contains simulated neural areas analogous to the ventral stream of the visual system, areas that influence visual tracking and areas analogous to ascending neuromodulatory reward systems that modulate synaptic plasticity.

To represent the relative timing of neuronal activity in the simulation, the activity of each neuronal unit in Darwin VIII is described by both a firing rate variable and a phase variable. In

line with the proposal that cortical neurons detect temporal coincidences among synaptic inputs (Abeles, 1982; Konig *et al.*, 1996; Azouz and Gray, 2000), our model includes a mechanism that supports synchrony among coactive and connected neuronal units (Tononi *et al.*, 1992). We found that this mechanism provides for the emergence of multiple, differentiated, and globally distributed neuronal circuits that correspond to distinct objects in Darwin VIII's visual field.

As it moves autonomously in the laboratory environment, Darwin VIII approaches and views multiple objects that share visual features. It becomes conditioned to prefer one target object over multiple distracters by association of the target with an innately preferred auditory cue. Darwin VIII demonstrates this preference behaviorally by orienting toward the target. While observing the device's behavior during this task, we recorded the state of every neuronal unit and synaptic connection in its simulated nervous system. As we have noted previously (Krichmar and Edelman, 2002), such data would be impossible to obtain and compare in animal experiments. By enabling behavioral and brain responses to be followed in detail at all levels of control, brain-based devices such as Darwin VIII have heuristic value in interpreting data obtained from behaving animals.

Our results support the hypothesis that visual binding results from the dynamic synchronization of neural activity mediated by reentrant connections among widely dispersed neural areas. The performance of Darwin VIII also suggests that specific timing relations and mean firing rates can act in a complementary fashion to regulate behavior, and that synchrony among groups of neurons — as distinct from synchrony between pairs of individual neurons — may play a significant role in adaptive neural function.

Materials and Methods

We have developed a heuristic in which a physically instantiated neurally organized mobile adaptive device (NOMAD), guided by a simulated nervous system, explores its environment and develops adaptive behaviors through experience-dependent learning. In the present report, we refer to the device and the neural simulation together as Darwin VIII. Darwin VIII is a part of a series of automata in which hypotheses about the nervous system are tested by implementing them in brain-based devices (Reeke *et al.*, 1990; Edelman *et al.*, 1992; Almassy *et al.*, 1998; Krichmar *et al.*, 2000; Krichmar and Edelman, 2002). The NOMAD portion of Darwin VIII consists of a mobile base equipped with a CCD camera for vision (see Appendix A), microphones for hearing (see Appendix B), effectors for movement, and infrared detectors for obstacle avoidance (Fig. 1).

Darwin VIII's behavior is guided by a simulated nervous system (see Appendix *C. Computation*) modeled on the anatomy and physiology of the mammalian nervous system but, obviously, with far fewer neurons and a much less complex architecture. It consists of a number of areas labeled according to the analogous cortical and subcortical brain regions. Each area contains different types of neuronal units, each of which represents a local population of neurons (Edelman, 1987). To distinguish modeled areas from corresponding regions in the mammalian nervous system, the simulated areas are indicated in italics (e.g. *IT*).

Neuroanatomy

In the present experiments, the simulated nervous system contained 28 neural areas, 53 450 neuronal units, and ~1.7 million synaptic connections. It includes a visual system, a tracking system, neural areas that respond to auditory cues, and a value or reward system. Figure 2 shows a high-level diagram of the simulated nervous system including the various neural areas and the arrangement of feedforward and reentrant synaptic connections. Specific parameters relating to each area and to patterns of connectivity are given in Tables 1 and 2.

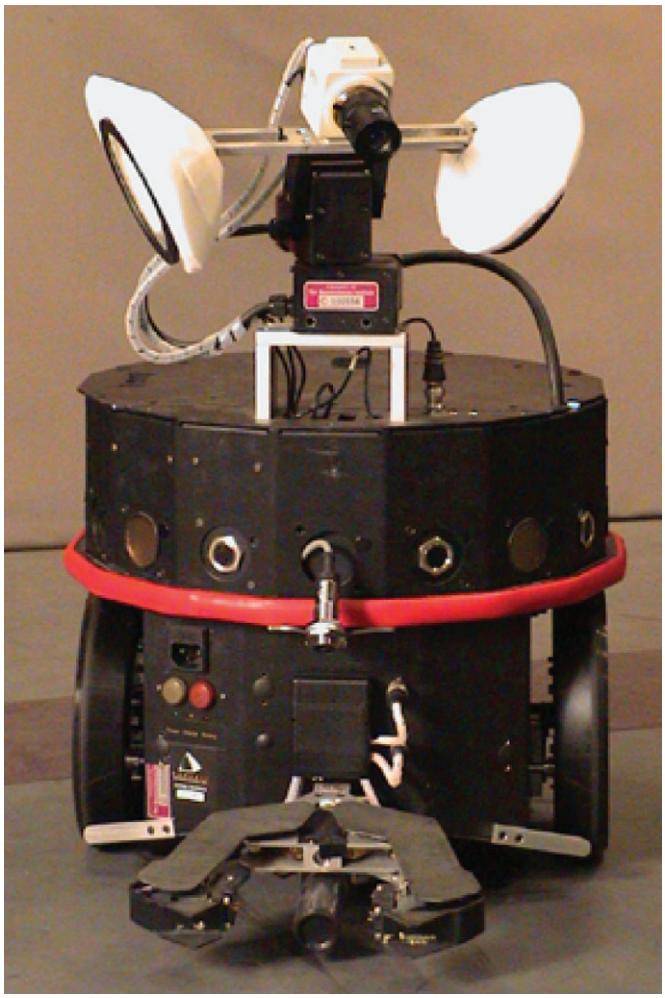

Figure 1. Darwin VIII, a brain-based device. Darwin VIII consists of a mobile base equipped with several sensors and effectors (neurally organized mobile adaptive device, or NOMAD), and a neural simulation running on a remote cluster of computer workstations. NOMAD contains a radio modem to transmit status, IR sensor information, and auditory information to the computer workstations carrying out the neural simulation and to receive motor commands from the simulation. Video output from a CCD camera mounted on Darwin VIII is sent to the workstations via RF transmission. RF input and output to NOMAD allows for untethered exploration. The CCD camera and the two microphones on either side of the camera provide sensory input to the neuronal simulation. An infrared (IR) sensor at the front of platform detects differences in surface reflectivity, triggering reflexive turns. All behavioral activity other than the IR reflexive turn is evoked by signals received from the neural simulation.

Visual System

The visual system is modeled on the primate occipitotemporal or ventral cortical pathway (in our model *V1* → *V2* → *V4* → *IT*), in which neurons in successive areas have progressively larger receptive fields until, in inferotemporal cortex, receptive fields cover nearly the entire visual field (Ungerleider and Haxby, 1994). Visual images from NOMAD's CCD camera are filtered for color and edges (see Appendix A) and the filtered output directly influences neural activity in area *V1*. *V1* is divided into subregions each having neuronal units that respond preferentially to green (*V1-green*), red (*V1-red*), horizontal line segments (*V1-horizontal*), vertical line segments (*V1-vertical*), 45° lines (*V1-diagonal-right*) and 135° lines (*V1-diagonal-left*). By assuming only these response properties and omitting the detailed microcircuitry of V1 (Raizada and Grossberg, 2003), this model provides a computationally tractable foundation for analyzing higher-

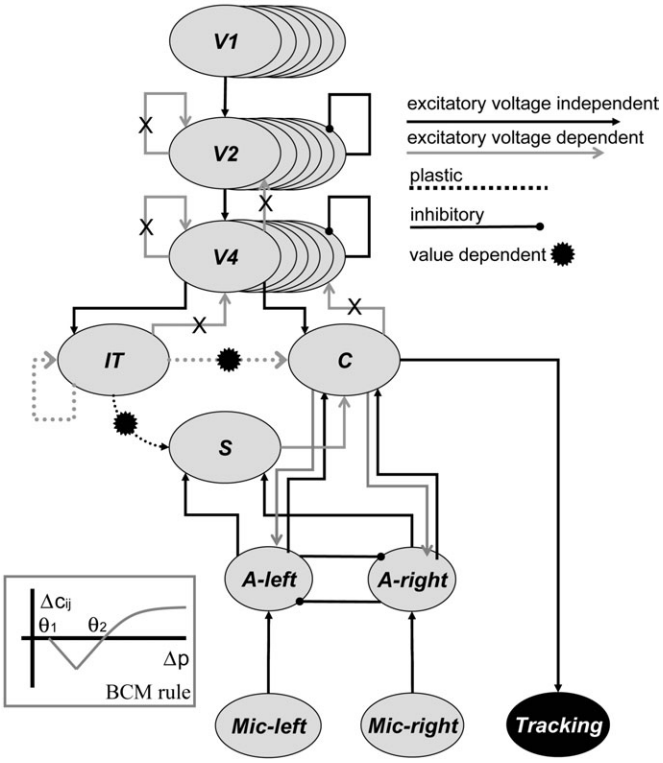

Figure 2. Schematic of the regional and functional neuroanatomy of Darwin VIII. In the version used in the present experiments, the simulated nervous system contained 28 neuronal areas, 53 450 neuronal units and ~1.7 million synaptic connections. The gray ellipses denote different neural areas. Arrows between the areas denote projections from one area to another. Projections marked with an 'X' are removed during lesion experiments. Tracking commands were issued to NOMAD's wheels based on activity in area *C*. Inset shows the form of the BCM rule in which synaptic change (Δc_{ij}) is a function of the phase difference between post- and presynaptic units (Δp) and two thresholds (θ_1 and θ_2). For clarity, specific projections among visual sub-areas have been omitted. See text and Tables 1 and 2 for details.

level interactions within the visual system and between the visual system and other cortical areas.

Subregions of *V1* project topographically to corresponding subregions of *V2*. The receptive fields of neuronal units in *V2* are narrow and correspond closely to pixels from the CCD camera image. *V2* has both excitatory and inhibitory reentrant connections within and among its subregions. Each *V2* subregion projects to a corresponding *V4* subregion topographically but broadly, so that *V4*'s receptive fields are larger than those of *V2*. *V4* subregions project back to the corresponding *V2* subregions with non-topographic reentrant connections. The reentrant connectivity within and among subregions of *V4* is similar to that in *V2*. *V4* projects in turn non-topographically to *IT* so that each neuronal unit in *IT* receives input from three *V4* neuronal units randomly chosen from three different *V4* subregions. Thus, while neuronal units in *IT* respond to a combination of visual inputs, the level of synaptic input into a given *IT* neuronal unit is fairly uniform; this prevents the activity of individual *IT* neuronal units from dominating the overall activity patterns. *IT* neuronal units project to other *IT* neuronal units through plastic connections, and back to *V4* through non-topographic reentrant connections.

Tracking System

The tracking system allows Darwin VIII to orient towards auditory and visual stimuli. The activity of area *C* (analogous to the superior colliculus) dictates where Darwin VIII directs its camera gaze. Tracking in Darwin VIII is achieved by signals to Darwin VIII's wheels based on the vector summation of the activity of the neuronal units in area *C* (Georgopoulos *et al.*, 1986). Each neuronal unit in area *C* has a receptive field which matches its preferred direction, and the area has a topographic arrangement (Lee *et al.*, 1988) such that if activity is predominately on the left side of *C*, signals to Darwin VIII's wheels are issued that evoke a turn towards the left. The auditory areas (*A-left* and *A-right*) have strong excitatory projections to the respective ipsilateral sides of *C* causing Darwin VIII to orient towards a sound source. *V4* projects topographically to *C*, its activity causing Darwin VIII to center its gaze on a visual object. Both *IT* and the value system *S* project to *C*, and plastic connections in the pathways $IT \rightarrow C$ and $IT \rightarrow S$ facilitate target selection by creating a bias in activity, reflecting salient perceptual categories (see *Value System,* below). Prior to experience, because of a lack of bias, Darwin VIII directs its gaze predominately between two objects. After learning to prefer a visual object, changes in the strengths of the plastic connections result in greater activity in those parts of *C* corresponding to the preferred object's position.

Table 1
Neuronal unit parameters

Area	Size	σ–*fire*	σ–*phase*	σ–*vdep*	ω	*g*
V1 (6)	60 × 80	–	–	–	–	–
V2 (6)	30 × 40	0.10	0.45	0.05	0.30	1.0*
V4 (6)	15 × 20	0.20	0.45	0.10	0.50	1.0*
C	15 × 20	0.10	0.10	0.10	0.50	1.0
IT	30 × 30	0.20	0.20	0.10	0.75	1.0
S	4 × 4	0.10	0.00	0.00	0.15	1.0
Mic–right	1 × 1	–	–	–	–	–
Mic–left	1 × 1	–	–	–	–	–
A–left	4 × 4	0.00	0.00	0.10	0.50	1.0
A–right	4 × 4	0.00	0.00	0.10	0.50	1.0

Values of parameters defining properties of neuronal units in Darwin VIII. Area *V1* is an input area and its activity is set based on the camera image (see Appendix A). Areas *V1*, *V2* and *V4* have six sub-areas each, with neuronal units selective for color (red and green), and line orientation (0, 45, 90 and 135°). *Mic-left* and *Mic-right* are input areas and their activity is set based on the microphone input (see Appendix B). The table indicates the number of neuronal units in each area or sub-area (*Size*). Neuronal units in each area apart from areas *VI*, *Mic-left* and *Mic-right* have a specific firing threshold (σ-*fire*), a phase threshold (σ-*phase*), a threshold above which voltage-dependent connections can have an effect (σ-*vdep*), a persistence parameter (ω) and a scaling factor (*g*). Asterisks mark values that are set to 2.0 for Darwin VIII subjects with lesioned reentrant connections (see Table 2).

Auditory System

This system converts inputs from microphones into simulated neuronal unit activity (see Appendix *B. Auditory System and its Inputs*). Neural areas *Mic-left* and *Mic-right* are respectively activated whenever the corresponding microphones detect a sound of sufficient amplitude within a specified frequency range. *Mic-left/Mic-right* project to neuronal units in areas *A-left/A-right* (see Fig. 2). Sound from one side results in activity on the ipsilateral side of the auditory system, which in turn produces activity on the ipsilateral side of *C* causing orientation towards the sound source.

Value System

Activity in the simulated value system (Area *S*, Fig. 2) signals the occurrence of salient sensory events and this activity contributes to the modulation of connection strengths in pathways $IT \rightarrow S$ and $IT \rightarrow C$. Initially, *S* is activated by sounds detected by Darwin VIII's auditory system (see *A-left* $\rightarrow S$ and *A-right* $\rightarrow S$ in Fig. 2). Activity in *S* is analogous to that of ascending neuromodulatory systems in that it is triggered by salient events, influences large regions of the simulated nervous system (see Synaptic Plasticity, below), and persists for several cycles (Aston-Jones and Bloom, 1981; Schultz *et al.*, 1997; Sporns *et al.*, 2000). In addition, due to its projection to the tracking area *C*, area *S* has a direct influence on behavior.

Table 2
Properties of anatomical projections and connection types in Darwin VIII

Projection	Arbor	P	$c_{ij}(0)$	Type	η	θ_1	θ_2	k_1	k_2
V1 → *V2*	[] 0 × 0	1.00	1,2	PI	0.00	0	0	0.00	0.00
V2 → *V2* (intra)	[] 3 × 3	0.75	0.45,0.85	VD	0.00	0	0	0.00	0.00
V2 → *V2* (inter) (X)	[] 2 × 2	0.40	0.5,0.65	VD	0.00	0	0	0.00	0.00
V2 → *V2* (intra)	Θ 18,25	0.10	–0.05,–0.1	VI	0.00	0	0	0.00	0.00
V2 → *V2* (inter)	[] 2 × 2	0.05	–0.05,–0.1	VI	0.00	0	0	0.00	0.00
V2 → *V4*	[] 3 × 3	0.40	0.1,0.12	VI	0.00	0	0	0.00	0.00
V4 → *V2* (X)	[] 1 × 1	0.10	0.25,0.5	VD	0.00	0	0	0.00	0.00
V4 → *V4* (inter) (X)	[] 2 × 2	0.40	1.75,2.75	VD	0.00	0	0	0.00	0.00
V4 → *V4* (intra)	Θ 10,15	0.10	–0.15,–0.25	VI	0.00	0	0	0.00	0.00
V4 → *V4* (inter)	Θ 10,15	0.10	–0.15,–0.25	VI	0.00	0	0	0.00	0.00
V4 → *V4* (inter)	[] 2 × 2	0.03	–0.15,–0.25	VI	0.00	0	0	0.00	0.00
V4 → *C*	[] 3 × 3	1.00	0.002,0.0025	VI	0.00	0	0	0.00	0.00
V4 → *IT*	special	–	0.1,0.15	VI	0.00	0	0	0.00	0.00
IT → *V4* (X)	non-topo	0.01	0.05,0.07	VD	0.00	0	0	0.00	0.00
IT → *IT*	non-topo	0.10	0.14,0.15	VD	0.10	0	0.866	0.90	0.45
IT → *C#*	non-topo	0.10	0.2,0.2	VD	1.00	0	0.707	0.45	0.65
IT → *S#*	non-topo	1.00	0.0005,0.001	VI	0.10	0	0.707	0.45	0.45
C → *V4* (X)	non-topo	0.01	0.05,0.07	VD	0.00	0	0	0.00	0.00
C → *C*	Θ 6,12	0.50	–0.05,–0.15	PI	0.00	0	0	0.00	0.00
C → *Mic-left*	non-topo	1.00	35,35	VD	0.00	0	0	0.00	0.00
C → *Mic-right*	non-topo	1.00	35,35	VD	0.00	0	0	0.00	0.00
S → *C*	non-topo	0.50	0.5,0.5	VD	0.00	0	0	0.00	0.00
S → *S*	non-topo	0.50	0.7,0.8	VD	0.00	0	0	0.00	0.00
A-left → *C*	left only	1.00	0.5,0.5	VD	0.00	0	0	0.00	0.00
A-right → *C*	right only	1.00	0.5,0.5	VD	0.00	0	0	0.00	0.00
A-left → *C*	right only	1.00	–0.15,–0.15	PI	0.00	0	0	0.00	0.00
A-right → *C*	left only	1.00	–0.15,–0.15	PI	0.00	0	0	0.00	0.00
A-left → *S*	non-topo	1.00	35,35	VD	0.00	0	0	0.00	0.00
A-right → *S*	non-topo	1.00	35,35	VD	0.00	0	0	0.00	0.00
A-left ↔ *A-right*	non-topo	1.00	–1,–1	PI	0.00	0	0	0.00	0.00
A-left ↔ *A-right*	non-topo	1.00	–0.5,–0.5	VD	0.00	0	0	0.00	0.00
Mic-left,Mic-right → *A-left,A-right*	non-topo	1.00	5,5	PI	0.00	0	0	0.00	0.00

A presynaptic neuronal unit connects to a postsynaptic neuronal unit with a given probability (P) and given projection shape (Arbor). This arborization shape can be rectangular '[]' with a height and width (*h* × *w*), doughnut shaped 'Θ' with the shape constrained by an inner and outer radius (r1, r2), left-only (right-only) with the presynaptic neuronal unit only projecting to the left (right) side of the postsynaptic area, or non-topographical 'non-topo' where any pairs of presynaptic and postsynaptic neuronal units have a given probability of being connected. The initial connection strengths, $c_{ij}(0)$, are set randomly within the range given by a minimum and maximum value (min, max). A negative value for $c_{ij}(0)$, indicates inhibitory connections. Connections marked with 'intra' denote those within a visual sub-area and connections marked with 'inter' denote those between visual sub-areas. Inhibitory 'inter' projections connect visual sub-areas responding to shape only or to color only (e.g. *V4-red* ↔ *V4-green, V4-horizontal* ↔ *V4-vertical*), excitatory 'inter' projections connect shape sub-areas to color sub-areas (e.g. *V4-red* ↔ *V4-vertical*). Projections marked # are value-dependent. A connection type can be phase-independent/voltage-independent (PI), phase-dependent/voltage-independent (VI) or phase-dependent/voltage-dependent (VD). Non-zero values for η, θ_1, θ_2, k_1 and k_2 signify plastic connections. The connection from *V4* to *IT* was special in that a given neuronal unit in area *IT* was connected to three neuronal units randomly chosen from three different *V4* sub-areas. Projections marked with an 'X' were removed during lesion experiments.

Neuronal Dynamics

A neuronal unit in Darwin VIII is simulated by a mean firing rate model. The state of each unit is determined by both a mean firing rate variable (*s*) and a phase variable (*p*). The mean firing rate variable of each unit corresponds to the average activity of a group of ~100 real neurons during a time period of ~100 ms. The phase variable, which specifies the relative timing of firing activity, provides temporal specificity without incurring the computational costs associated with modeling of the spiking activity of individual neurons in real-time (see Neuronal Unit Activity and Phase, below).

Synaptic connections between neural units, both within and between neuronal areas, are set to be either voltage-independent or voltage-dependent, either phase-independent or phase-dependent, and either plastic or non-plastic. Voltage-independent connections provide synaptic input regardless of postsynaptic state. Voltage-dependent connections represent the contribution of receptor types (e.g. NMDA receptors) that require postsynaptic depolarization to be activated (Wray and Edelman, 1996). In Darwin VIII, all within-area excitatory connections and all between-area reentrant excitatory connections are voltage-dependent (see Fig. 2 and Table 2). These

connections play a modulatory role in neuronal dynamics (Grossberg, 1999). Phase-dependent connections influence both the activity and the phase of postsynaptic neuronal units, whereas phase-independent connections influence only the activity. All synaptic pathways in Darwin VIII are phase-dependent except those involved in motor output (see Table 2: *A-left/A-right* → *C*, *C* ↔ *C*) or sensory input (see Table 2: *Mic-left/Mic-right* → *A-left/A-right, A-left* ↔ *A-right*, *V1* → *V2*), since signals at these interfaces are defined by magnitude only. Plastic connections are either value-independent or value-dependent, as described below.

Neuronal Unit Activity and Phase

The mean firing rate (s) of each neuronal unit ranges continuously from 0 (quiescent) to 1 (maximal firing). The phase (p) is divided into 32 discrete bins representing the relative timing of activity by an angle ranging from 0 to 2π. The state of a neuronal unit is updated as a function of its current state and contributions from voltage-independent, voltage-dependent, and phase-independent inputs. The voltage-independent input to unit i from unit j is:

$$A_{ij}^{VI}(t) = c_{ij}s_j(t) \quad (1)$$

where $s_j(t)$ is the activity of unit j, and c_{ij} is the connection strength from unit j to unit i. The voltage-independent postsynaptic influence on unit i is calculated by convolving this value into a cosine-tuning curve over all phases:

$$POST_i^{VI} = \sum_{l=1}^{M}\sum_{j=1}^{N_l}\left(A_{ij}^{VI}(t)\sum_{k=1}^{32}\left(\frac{\cos((2\pi/32)(k-p_j(t)))+1}{2}\right)^{tw}\right) \quad (2)$$

where M is the number of different anatomically defined connection types (see Table 2); N_l is the number of connections of type M projecting to unit i; $p_j(t)$ is the phase of neuronal unit j at time t; and tw is the tuning width, which in our experiments is set to 10 so that the width of the tuning curve is relatively sharp (~5 phase bins).

The voltage-dependent input to unit i from unit j is:

$$A_{ij}^{VD}(t) = \Phi(POST_i^{VI}(p_j(t)))c_{ij}s_j(t) \quad (3)$$

where

$$\Phi(x) = \begin{cases} 0; & x < \sigma_i^{vdep} \\ x; & \text{otherwise} \end{cases}$$

where σ_i^{vdep} is a threshold for the postsynaptic activity below which voltage-dependent connections have no effect (see Table 1).

The voltage-dependent postsynaptic influence on unit i is given by:

$$POST_i^{VD} = \sum_{l=1}^{M}\sum_{j=1}^{N_l}\left(A_{ij}^{VD}(t)\sum_{k=1}^{32}\left(\frac{\cos((2\pi/32)(k-p_j(t)))+1}{2}\right)^{tw}\right) \quad (4)$$

The phase-independent activation into unit i from unit j is:

$$A_{ij}^{PI}(t) = c_{ij}s(t) \quad (5)$$

The phase-independent postsynaptic influence on unit i is a uniform distribution based on all the phase-independent inputs divided by the number of phase bins (32).

$$POST_i^{PI} = \sum_{l=1}^{M}\sum_{j=1}^{N_l}\left(\frac{A_{ij}^{PI}(t)}{32}\right) \quad (6)$$

A new phase, $p_i(t+1)$, and activity, $s_i(t+1)$, are chosen based on a distribution created by linearly summing the postsynaptic influences on neuronal unit i (see Fig. 3):

$$POST_i = \sum_{j=1}^{N_{VI}} POST_j^{VI}\sum_{k=1}^{N_{VD}} POST_k^{VD}\sum_{l=1}^{N_{PI}} POST_l^{PI} \quad (7)$$

The phase threshold, σ_i^{phase}, of the neuronal unit is subtracted from the distribution $POST_i$ and a new phase, $p_i(t+1)$, is calculated with a probability proportional to the resulting distribution (Fig. 3; bottom row). If the resulting distribution has an area less than zero (i.e. no inputs are above the phase threshold), a new phase, $p_i(t+1)$, is chosen at random. The new activity for the neuronal unit is the activity level at the newly chosen phase, which is then subjected to the following activation function:

$$s_i(t+1) = \phi(\tanh(g_i(POST_i(p_i(t+1)) + \omega_{s_i}(t)))) \quad (8)$$

where

$$\Phi(x) = \begin{cases} 0; & x < \sigma_i^{fire} \\ x; & \text{otherwise} \end{cases}$$

where ω determines the persistence of unit activity from one cycle to the next, g_i is a scaling factor and σ_i^{fire} is a unit specific firing threshold. Specific parameter values for neuronal units are given in Table 1, and synaptic connections are specified in Table 2.

In this model of a neuronal unit, postsynaptic phase tends to be correlated with the phase of the most strongly active presynaptic inputs (Abeles, 1982; Konig *et al.*, 1996; Azouz and Gray, 2000). We found that this neuronal unit model facilitates the emergence of synchronously active neuronal circuits in both a simple network (see Neuronal Synchrony in a Simple Network Model, below) and in the full Darwin VIII, where such emergence involves additional

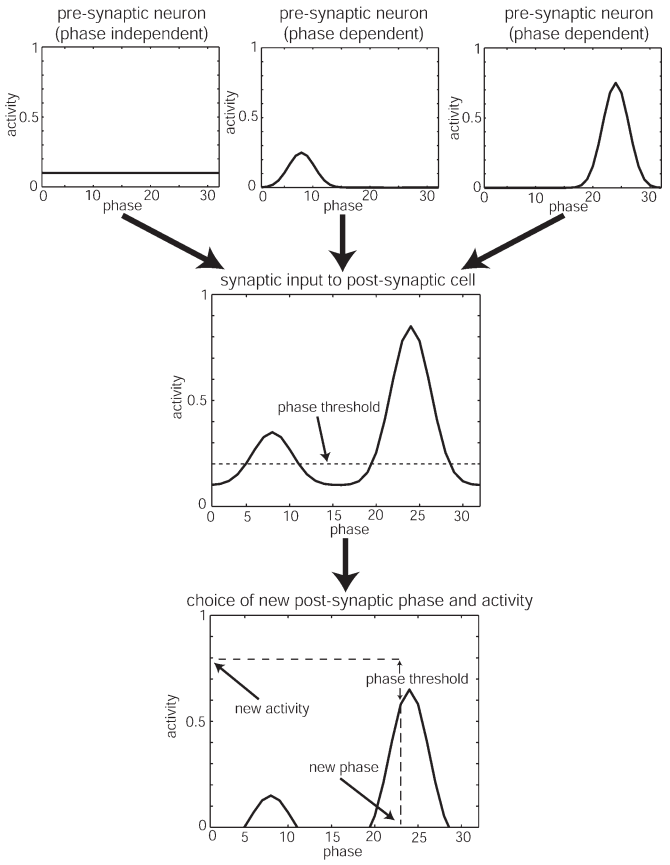

Figure 3. Choosing a new firing rate and phase for a neuronal unit. Phases and activities of presynaptic neuronal units are convolved into either a uniform distribution for phase-independent inputs (see top row left) or a cosine tuning curve distribution for phase-dependent inputs (see top row middle and right). The input distributions are linearly summed to create an overall synaptic input distribution for the postsynaptic neuronal unit (see middle row). The neuronal unit's phase threshold is subtracted from the distribution and a new phase is chosen with a probability proportional to the resulting distribution (see bottom row). The new activity for the neuronal unit is the activity level at the new phase plus the phase threshold.

constraints imposed by reentrant connectivity, plasticity and behavior.

Synaptic Plasticity

Synaptic strengths are subject to modification according to a synaptic rule that depends on the phase and activities of the pre- and postsynaptic neuronal units. Plastic synaptic connections are either value-independent (see $IT \rightarrow IT$ in Fig. 2) or value-dependent (see $IT \rightarrow S$, $IT \rightarrow C$ in Fig. 2). Both of these rules are based on a modified BCM learning rule (Bienenstock *et al.*, 1982) in which thresholds defining the regions of depression and potentiation are a function of the phase difference between the presynaptic and postsynaptic neuronal units (see Fig. 2, inset). Synapses between neuronal units with strongly correlated firing phases are potentiated and synapses between neuronal units with weakly correlated phases are depressed; the magnitude of change is determined as well by pre- and postsynaptic activities. This learning rule is similar to a spike-time dependent plasticity rule (Bi and Poo, 1998; Senn *et al.*, 2001; Song and Abbott, 2001) applied to jittered spike trains where the region of potentiation has a high peak and a thin tail, and the region of depression has a comparatively small peak and fat tail (Izhikevich and Desai, 2003).

Value-independent synaptic changes in c_{ij} are given by:

$$\Delta_{c_{ij}}(t+1) = \eta_{s_i}(t)_{s_j}(t)BCM(\Delta p) \tag{9}$$

where $s_i(t)$ and $s_j(t)$ are activities of post- and presynaptic units, respectively, η is a fixed learning rate and

$$\Delta p = \frac{\cos((2\pi/32)(p_i(t) - p_j(t))) + 1}{2}$$

where $p_i(t)$ and $p_j(t)$ are the phases of post- and presynaptic units ($0.0 \le \Delta p \le 1.0$). A value of Δp near 1.0 indicates that pre- and postsynaptic units have similar phases, a value of Δp near 0.0 indicates that pre- and postsynaptic units are out of phase. The function *BCM* is implemented as a piecewise linear function, taking Δp as input, that is defined by two thresholds (q_1, q_2, in radians), two inclinations (k_1, k_2) and a saturation parameter ρ ($\rho = 6$ throughout):

$$BCM(\Delta p) = \begin{cases} 0; & \Delta p < \theta_1 \\ k_1(\theta_1 - \Delta p); & \theta_1 \le \Delta p < (\theta_1 + \theta_2)/2 \\ k_1(\Delta p - \theta_2); & (\theta_1 + \theta_2)/2 \le \Delta p < \theta_2 \\ k_2 \tanh(\rho(\Delta p - \theta_2))/\rho; & \text{otherwise} \end{cases} \tag{10}$$

Specific parameter settings for fine-scale synaptic connections are given in Table 2.

The rule for value-dependent synaptic plasticity differs from the value-independent rule in that an additional term, based on the activity and phase of the value system, modulates the synaptic strength changes. Synaptic connections terminating on neuronal units that are in phase with the value system are potentiated, and connections terminating on units out of phase with the value system are depressed.

The synaptic change for value-dependent synaptic plasticity is given by:

$$\Delta_{c_{ij}}(t+1) = \eta_{s_i}(t)_{s_j}(t)BCM(\Delta p)V(t)BCM_v(\Delta p_v) \tag{11}$$

where $V(t)$ is the mean activity level in the value area S at time t. Note that the BCM_v function is slightly different from the *BCM* function above in that it uses the phase difference between area S and the postsynaptic neuronal unit as input,

$$\Delta p_v = \frac{\cos((2\pi/32)(p_v(t) - p_i(t))) + 1}{2}$$

where $p_v(t)$ is the mean phase in area S. When both *BCM* and BCM_v return a negative number, BCM_v is set to 1 to ensure that the synaptic connection is not potentiated when both the presynaptic neuronal unit and value system are out of phase with the postsynaptic neuronal unit.

Simulation Cycle Computation

During each simulation cycle of Darwin VIII, sensory input is processed, the states of all neuronal units are computed, the connection strengths of all plastic connections are determined, and motor output is generated. In our experiments, execution of each simulation cycle required ~100 milliseconds of real time (see Appendix C).

Experimental Conditions

Experimental Environment

Figure 4*A* shows a diagram of Darwin VIII's environment. A photograph of Darwin VIII in such an environment is given in Figure 4*B*. The environment consisted of an enclosed area with black walls. Various pairs of shapes from a set consisting of a green diamond, a green square, a red diamond, and a red square were hung on two opposite walls. The floor was covered with opaque black plastic panels, and contained a boundary made of reflective construction paper. When this boundary was detected by the infrared detector attached to the front of NOMAD and facing toward the floor, Darwin VIII made one of two reflexive movements: (i) if an object was in its visual field, it backed up, stopped, and then turned ~180°; (ii) if there was no object in its visual field, Darwin VIII turned ~90°, thus orienting away from walls without visual stimuli. Near the boundary of walls containing visual shapes, infrared emitters (IR) on one side of the room were paired with IR sensors containing a speaker on the other side, to create an IR beam (see Fig. 4*A*). If Darwin VIII's movement broke either beam, a tone was emitted. Detection of the tone by Darwin VIII elicited an orientation movement towards the source of the sound.

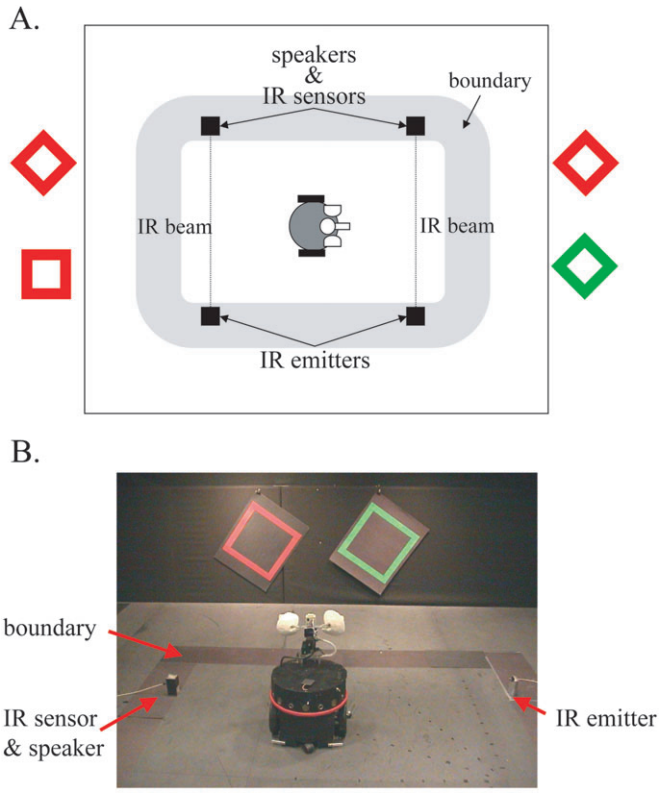

Figure 4. Experimental setup for Darwin VIII. (*A*) Darwin VIII views objects on two of the walls of an arena. The area Darwin VIII explores (90″ by 66″) is constrained by a boundary of reflective construction paper. Detection of this boundary by Darwin VIII's infrared sensor triggers a reflexive turn. When Darwin VIII breaks the beam from the IR emitter to the IR sensor, a tone is emitted from the speaker. (*B*) Darwin VIII photographed in the experimental environment.

Experimental Protocol

Experiments were divided into two stages: training and testing (see Fig. 5). During both stages the activity and phase responses of all neuronal units were recorded for analysis.

In the training stage (Fig. 5*A*), Darwin VIII autonomously explored its enclosure for 10 000 simulation cycles, corresponding to 15–20 min of real time and ~24 approaches to the various pairs of visual shapes. Each pair contained a 'target' shape and a 'distracter' shape. Distracters were deliberately designed to share attributes with the target; for example, when a red diamond was the target, a red diamond–red square pair was hung on one wall and a red diamond–green diamond pair was hung on the other wall. The red diamond was closest to the speaker in both cases. To ensure that the left–right orientation of shapes in the target–distracter pair (e.g. red square on the left, red diamond on the right) did not confound target selection, we exchanged the side of the distracters every sixth viewing of a pair. During the training stage, responses to the speaker caused Darwin VIII to orient towards the target.

During testing (Fig. 5*B*), the speakers were turned off, and Darwin VIII was allowed to explore its enclosure for 15 000 simulation cycles. The first 10 000 cycles involved encounters with the same target and distracters present during the training stage. The final 5 000 cycles involved encounters with the target and the single shape of the set of four that did not share any features with the target (e.g. a pair consisting of a red diamond as target and a green square as distracter).

Training and testing was repeated with three different Darwin VIII 'subjects' using each of the four shapes as a target (a total of 12 training and testing sessions). Each Darwin VIII subject consisted of the same physical device, but each possessed a unique simulated nervous system. This variability among subjects was a consequence of random initialization in both the microscopic details of connectivity between individual neuronal units and the initial connection strengths between those units. The overall connectivity among neuronal units remained similar among different subjects, however, inasmuch as that connectivity was constrained by the synaptic pathways, arborization patterns, and ranges of initial connection strengths (see Fig. 2 and Table 2 for specifics).

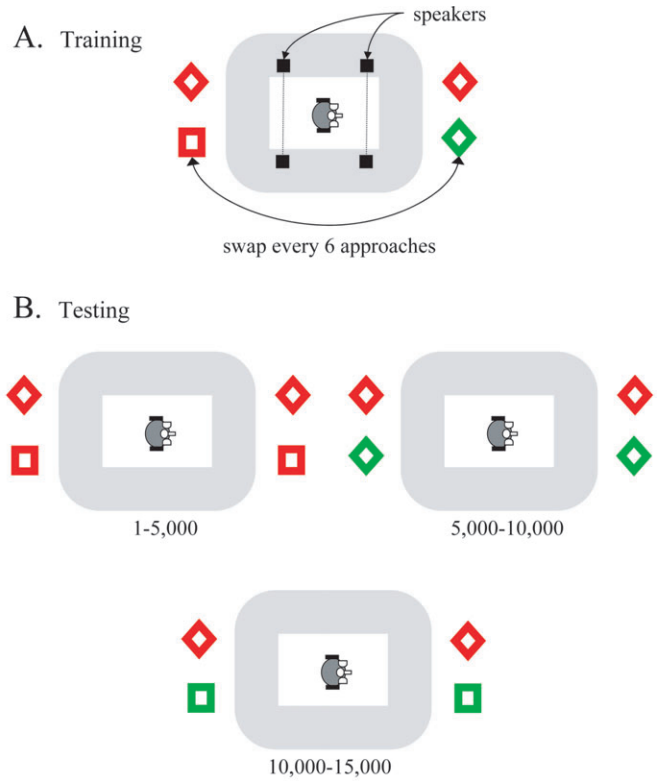

Figure 5. Experimental protocol for Darwin VIII. (*A*) During training, Darwin VIII explores its enclosure for 10 000 simulation cycles, corresponding to roughly 24 approaches to the pairs of visual objects. Responses to sounds emitted by the speaker cause Darwin VIII to orient towards the target, which in this example is a red diamond. The distracters, which are exchanged before every sixth approach to ensure that left–right orientation did not confound target selection, are a green diamond and a red square. (*B*) During testing the speakers are turned off and Darwin VIII explores for 15 000 simulation cycles. While the target is continuously present, the distracters are changed every 5000 cycles.

Results

Neuronal Synchrony in a Simple Network Model

To illustrate how reentrant connections among neuronal units can lead to neuronal synchrony in a mean firing rate model with a phase parameter, we analyzed a simple network model consisting of three neuronal units (Fig. 6*A*). Units *n1* and *n2* receive steady phase-independent input and project via voltage-independent connections to a third neuronal unit *n3*. Units *n1* and *n2* project to each other, and unit *n3* projects back to both *n1* and *n2* via reentrant voltage-dependent connections. All weights were chosen from a uniform random distribution of range 1.4–1.5. Figure 6*B* shows that in this model all neuronal units become synchronized within 10 simulation cycles. By contrast, if reentrant connections are removed (the dotted arrows in Fig. 6*A*) so that only feedforward projections remain, synchrony is not achieved (Fig. 6*C*).

Figure 6*D*,*E* shows the probability distributions from which postsynaptic phases are chosen for each unit (see Fig. 3, bottom). With reentrant connections intact (Fig. 6*D*), distributions for all neurons become peaked at the same phase. With reentrant connections absent ('lesioned' networks, Fig. 6*E*), the probability distributions for *n1* and *n2* remain flat due to their phase-independent input, and the distribution for *n3* varies randomly over time. While for clarity Figure 6 shows only the first 15 simulation cycles, these cycles are representative of network behavior over long durations (we observed both intact and lesioned networks for 10 000 cycles).

To explore whether specific connection strengths are important for network behavior, we repeated the above analysis several times using different random seeds, and we also compared a network in which all weights were set to a mean value (1.45). In all cases we observed networks for 10 000 cycles and noted qualitatively identical results to those shown in Figure 6. We also repeated our analysis for networks in which value-independent plasticity was enabled for the feedforward projections $n1 \rightarrow n3$ and $n2 \rightarrow n3$. As before, we tested networks with randomly selected weights as well as networks with all weights initially set to a mean value (1.45), and again, in both cases we observed synchrony in intact networks and no synchrony in lesioned networks (data not shown). Also, since pre- and postsynaptic units were correlated in activity and phase, plastic connections in the intact networks increased in strength by nearly 100% over 1000 cycles. In lesioned networks, however, because pre- and postsynaptic units were not in phase with each other these connections were depressed to ~10% of their initial values over the same duration.

The results from this reduced model show that the presence of reentrant connections can facilitate synchronous activity among neural areas, that this synchrony does not depend on specific or differential connection strengths, and that the absence of reentry is not compensated by synaptic plasticity. The full Darwin VIII model has three major differences from

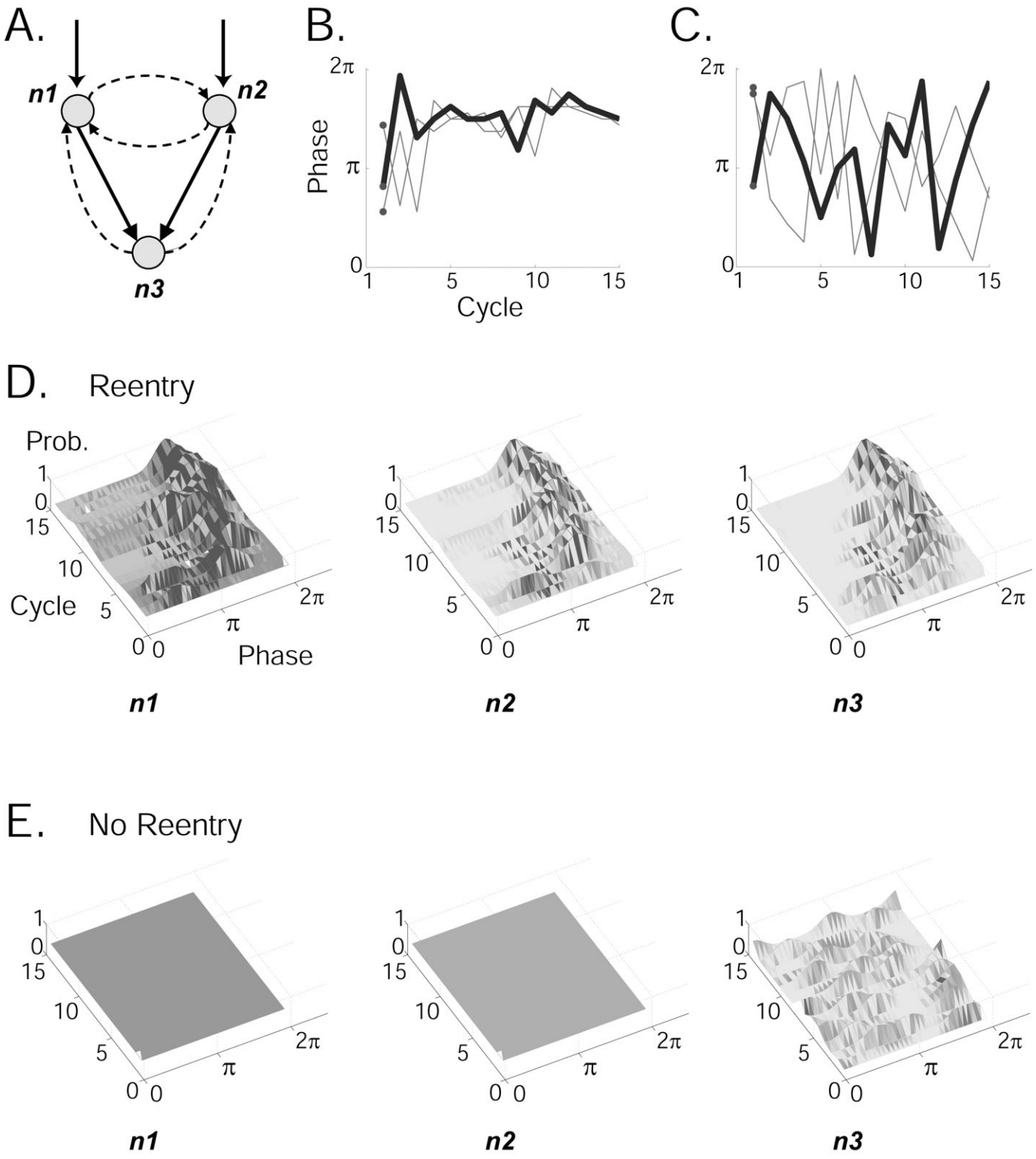

Figure 6. Neuronal synchrony in a simple model. (*A*) A simple network consisting of neuronal units *n1*, *n2* and *n3*. Solid lines represent voltage-independent feedforward connections, dashed lines represent voltage-dependent reentrant connections**.** All connections are phase-dependent. All neuronal units have parameters similar to those used in the visual binding experiments: σ-*fire* = 0.1, σ-*phase* = 0.45, σ-*vdep* = 0.1, ω = 0.3 and g = 1.0 (see Table 1 and Neuronal Unit Activity and Phase). All weights are chosen from a uniform random distribution of range 1.4–1.5. Units *n1* and *n2* receive steady voltage-independent and phase-independent input of magnitude 10.0. (*B*) Phase dynamics over 15 time-steps with reentrant connections intact; the thick black line represents *n3* and the gray lines represent *n1* and *n2*. The network rapidly achieves synchrony. (*C*) Phase dynamics with reentrant connections absent (only feedforward connections remain). The network does not achieve synchrony. (*D*) Probability distributions from which postsynaptic phases are chosen for each neuronal unit. With reentrant connections intact, all distributions become peaked at the same phase. (*E*) Probability distributions from which postsynaptic phases are chosen with reentrant connections absent. Distributions for *n1* and *n2* remain flat, and the distribution for *n3* varies randomly over time.

this reduced model. It has a large-scale reentrant neuroanatomy based on the vertebrate visual cortex, it involves value-dependent and value-independent synaptic plasticity, and it behaves autonomously in a real-world environment.

Target Tracking Behavior

The discrimination performance of each Darwin subject was assessed by how well that subject 'tracked' toward target objects in the absence of auditory cues. This was calculated as the fraction of time for which the target was centered in Darwin VIII's visual field during each approach to a pair of visual objects.

Figure 7*A* shows that all subjects successfully tracked the four different targets over 80% of the time. It should be noted that successful performance on this task is not trivial. Targets and distracters appeared in the visual field at many different scales and at many different positions as Darwin VIII explored its environment. Moreover, because of shared properties, targets cannot be reliably distinguished from distracters on the basis of color or shape alone.

To investigate the importance of the presence of reentrant connections in the model, certain inter-areal reentrant connections were lesioned at different stages of the experimental paradigm. In one case, previously trained subjects were retested after lesioning. In a second, reentrant connections were lesioned in both training and testing stages. Lesions were applied to a subset of inter-areal excitatory reentrant connections (see projections marked with an 'X' in Fig. 2 and in Table 2), which had the effect of transforming the simulated nervous system into a 'feed-forward' model of visual processing. To compensate for the reduction in activity due to these lesions, neuronal unit outputs in areas *V2* and *V4* were amplified (see Table 1). Figure 7B shows that subjects with intact reentrant connections performed significantly better than either lesioned group (Wilcoxon Ranksum test; $P < 0.01$). The decrease in performance observed in the absence of reentry indicates that reentrant connections are essential for behavior above chance in the discrimination task.

Neural Dynamics during Behavior

During Darwin VIII's behavior, circuits comprised of synchronously active neuronal groups were distributed throughout different areas in the simulated nervous system. Multiple objects were distinguishable by the differences in phase between the corresponding active circuits. A snapshot of Darwin VIII's neural responses during a typical behavioral run is given in Figure 8. The figure shows Darwin VIII during an approach to a red diamond target and a green diamond distracter towards the end of a training session. Each pixel in the depicted neural areas represents the activity and phase of a single neuronal unit. The phase is indicated by the color of each pixel and the activity is indicated by brightness of the pixel (black is no activity; very bright is maximum activity). The figure shows two neural circuits which are differentiated by their distinct phases and which were elicited respectively by the red diamond and the green diamond stimuli. As shown in the figure, Darwin VIII has not yet reached the beam that triggers the speaker to emit a tone. The activity of area *S* was nonetheless in phase with the activity in areas *V2* and *V4* corresponding to the target, and was therefore predictive of the target's saliency or value. Area *IT* has two patterns of activity, indicated by the two different phase colors, which reflect two perceptual categories. These patterns were brought about by visual input that is generated during Darwin VIII's movement. Finally, area *C* has more activity on the side that facilitates orientation towards the target (i.e. the red diamond).

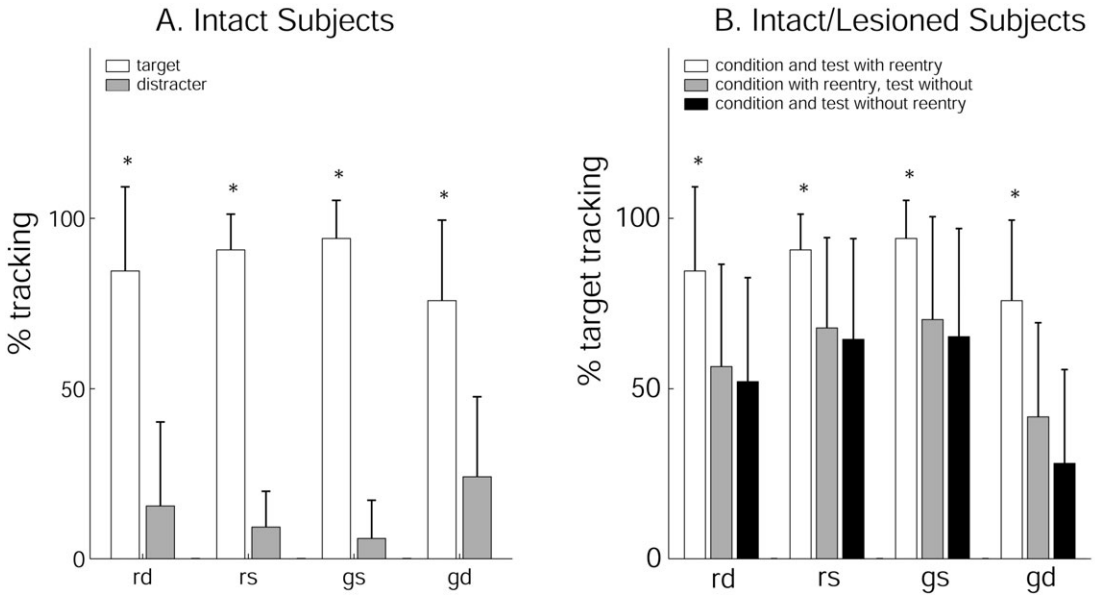

Figure 7. Darwin VIII behavior following conditioning. Three separate Darwin VIII subjects were conditioned to prefer one of 4 target shapes ('rd' = red diamond, 'rs' = red square, 'gs' = green square, 'gd' = green diamond). Activity in *V2* areas was used to assess the percentage of time for which NOMAD's visual field was centered on a particular visual shape. Bars in both graphs represent the mean percentage tracking time with error bars denoting the standard deviation. (*A*) Darwin VIII subjects with intact reentrant connections tracked the targets (white bars) significantly more than the distracters (gray bars) for each target shape, averaging over all approaches (asterisks denote $P < 0.01$ using a paired sample nonparametric sign test). (*B*) Subjects with reentrant connections intact (white bars) tracked targets significantly better than subjects with lesions only during testing (light gray bars), and subjects with lesions during both training and testing (black bars). Asterisks denote $P < 0.01$ using the Wilcoxon ranksum test.

To analyze the dynamics of these neural responses, we examined the phase distributions of active neuronal units during approaches to target–distracter pairs in the testing sessions. Figure 9*A* shows the distribution of neuronal phases in various neural areas during approaches to a red diamond target in the presence of a red square distracter, by an intact subject. The figure shows consistent correlations among phase distributions in *V4R* (red), *V4H* (horizontal), and *V4D* (diagonal). The bimodal distribution in *V4R* reflects the presence of two red shapes in Darwin VIII's environment: one trace correlates reliably with *V4D* and can therefore be associated with the red diamond, the other correlates reliably with *V4H* and can be associated with the red square (not shown is area *V4G* which remained inactive during this period). The phases of the active neuronal units in areas *S*, *IT* and *C* were strongly correlated with the red diamond target, as opposed to the red square distracter, reflecting the synaptic changes brought about by previous conditioning to prefer the red diamond. The global pattern of network activity thus displayed a biased phase distribution in favor of the target.

To quantify this bias and assess its generality, we compared the proportion of neuronal units in areas *S*, *IT* and *C* associated with the target with the proportion associated with the distracter during testing. Table 3 shows average values of these proportions calculated over all subjects and all four target

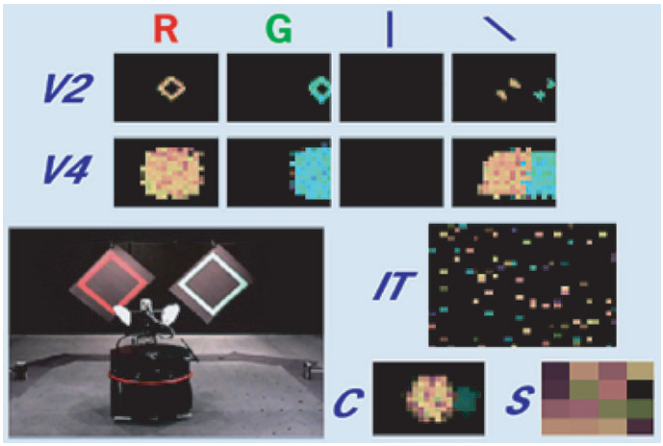

Figure 8. Snapshot of Darwin VIII's neuronal unit activity during a behavioral experiment. Darwin VIII is approaching a red diamond target (left) and a green diamond distracter (right) towards the end of a training session. Darwin VIII has not yet broken the beam that triggers the sound from the speakers located on the left side of the floor. The panels next to Darwin VIII show the activity and phase of selected neural areas [top row: *V2-red*, *V2-green*, *V2-vertical*, *V2-diagonal*; second row: *V4-red*, *V4-green*, *V4-vertical*, *V4-diagonal*; third row (to the right of Darwin VIII): *IT*; fourth row (to the right of Darwin VIII): *C* and *S*]. Each pixel in the selected neural area represents a neuronal unit; the activity is normalized from no activity (black) to maximum activity (bright colors), and the phase is indicated by the color of the pixel (colors were chosen from a pseudocolor map, there is no connection between the color of the stimulus object and the color representing the phases of neuronal responses). The neuronal units responding to the attributes of the red diamond share a common phase (red-orange color), whereas the neuronal units responding to the green diamond share a different phase (blue-green color).

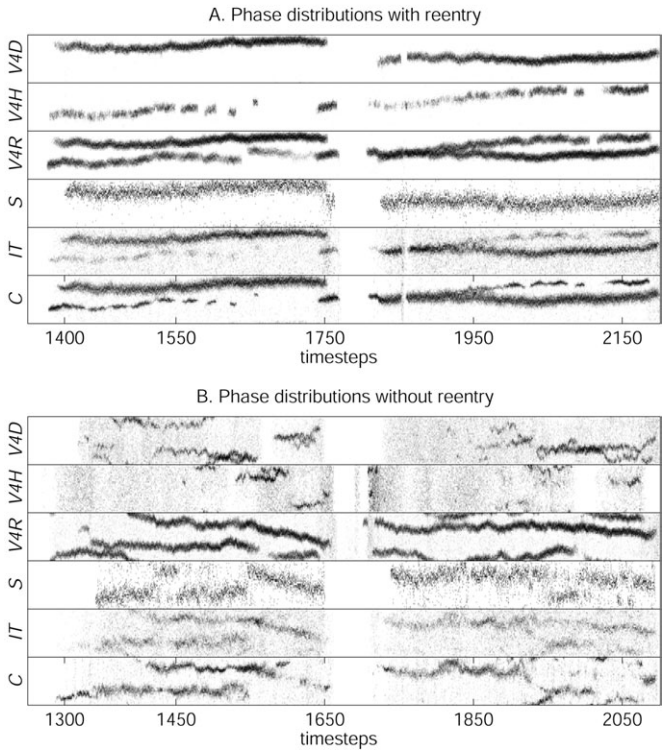

Figure 9. Phase responses of a Darwin VIII subject following conditioning. Each panel shows the distribution of phases in areas *V4D*, *V4H*, *V4R*, *S*, *IT* and *C* during approaches to a target (red diamond) in the presence of a distracter (red square). Shown in the panel for each neural area is the proportion of neuronal units at a particular phase, where phase is shown along the *y*-axis with a range of 0 to 2π, and a gray scale for the marks denotes the proportion of neuronal units at a specific phase (white represents no neuronal units, black represents the maximal proportion of neuronal units). Time-steps are shown along the *x*-axis. The scale was normalized for each neural area shown such that the black marks indicated the maximum proportion for that area. (*A*) Approaches by an intact Darwin VIII subject. (*B*) Approaches by a Darwin VIII subject in which inter-areal reentrant connections were lesioned following conditioning.

Table 3
Neuronal composition and average activity of functional circuits corresponding to target and distracter objects

Area	% units responding to target	% units responding to distracter	Mean firing rate of units responding to target	Mean firing rate of units responding to distracter
S	61.25 (17.78)*	10.34 (4.88)	0.537 (0.089)	0.495 (0.080)
IT	4.29 (0.495)*	2.91 (0.366)	0.579 (0.064)*	0.467 (0.015)
C	19.04 (1.45)*	10.80 (1.93)	0.626 (0.062)*	0.398 (0.032)
V4	14.83 (0.833)*	11.61 (0.819)	0.829 (0.003)*	0.823 (0.002)

Significantly more neuronal units respond in the functional circuits associated with targets than in those associated with the distracters, and the neuronal units associated with the target have significantly higher mean firing rates than neuronal units in circuits associated with distracters (asterisks denote $P < 0.001$ using a two-tailed *t*-test). Neuronal units associated with each shape were identified by correlating their phases with the phases of units in *V2* that responded to specific attributes of the shapes. Mean values were calculated across all subjects and all target shapes; standard errors are given in parentheses.

shapes. A significantly greater proportion of neuronal units were part of functional circuits associated with targets than in circuits associated with distracters. In addition, those units associated with targets had significantly higher firing rates. These observations may be compared with neurophysiological findings in primates which show increased activity in V4 (Reynolds *et al.*, 2000) and IT (Chelazzi *et al.*, 1998), as well as increased synchrony (Fries *et al.*, 2001) in response to a salient object as compared to distracters.

In our model, perceptual categorization and visual object discrimination is enabled by the coherent interaction of local and global processes, as mediated by reentrant connections. Local processes correspond to activity in each neural area, whereas global processes correspond to the distinct, but distributed functional circuits that emerged throughout the simulated nervous system. These interactions are evident in Figure 9*A*, in which activity in each of the local areas strongly reflects the global bias in favor of the red diamond target (see also Table 3).

The Influence of Reentry on Neural Dynamics

Lesioning of reentrant connections interfered significantly with interactions between local and global processes. Even in a very simple network model, removal of reentrant connections can prevent the emergence of neural synchrony (see Fig. 6). On a larger scale, Figure 9*B* shows approaches by the same Darwin subject depicted in Figure 9*A* to the same target/distracter pair, following lesions of inter-areal excitatory reentrant connections. While some individual areas continued to show peaks in their phase distribution (e.g. *V4R*), many did not, and the phase correlations between areas were severely diminished. This occurred not only among the various *V4* areas, but also among *V4* and areas *S*, *IT*, and *C*. The dynamically formed and globally coherent circuits, which were clearly evident in the intact subject, were almost entirely absent in the lesioned subjects. For example, Figure 9*B* shows that activity in area *S* no longer correlates uniquely with a single trace in *V4*, instead, it alternates between two distinct states. The absence of a dominant trace in *IT* and *C* is also evident.

Phase correlations between neural areas were significantly higher for subjects with intact reentrant connections than for subjects in either lesion group ($P << 0.0001$; Wilcoxon sign ranked test). The overall median Spearman's rank correlation coefficient was 0.36 for the intact subjects, 0.21 for the subjects with lesions only during the test stage, and 0.17 for the subjects with lesions in both the training and test stages. It is also notable that subjects with lesions only during testing had significantly higher correlation coefficients than subjects with lesions during both training and testing ($P << 0.0001$; Wilcoxon sign ranked test). This reflects the contribution of reentrant connections to the formation of global circuits during training. All of these findings are consistent with the drop in behavioral performance in the absence of reentrant connections (see Fig. 7).

A representative example of the correlation of phases among neural areas is given in Figure 10 for Subject 1 after conditioning to prefer red diamond targets. The figure is color coded (dark blue denotes no correlation, dark red denotes high correlation), and each colored area shows the correlation coefficient between the mean phases of a given pair of neural areas. Figure 10*A* shows correlation coefficients when reentrant connections were intact. In agreement with the data shown in Figure 9, strong phase correlations were found between areas associated with specific target features (*V4D* and *V4R*), and among these areas and areas *S*, *IT* and *C*. The correlations among neural areas for the same subject with reentrant connections lesioned during testing (Fig. 10*B*) and with reentrant connections lesioned during both conditioning and testing (Fig. 10*C*) were both considerably weaker.

Invariant Object Recognition

Because images of the visual objects varied considerably in size and position as Darwin VIII explored its enclosure, successful discrimination required invariant object recognition. In order to analyze this capacity, we examined the value system, area *S*,

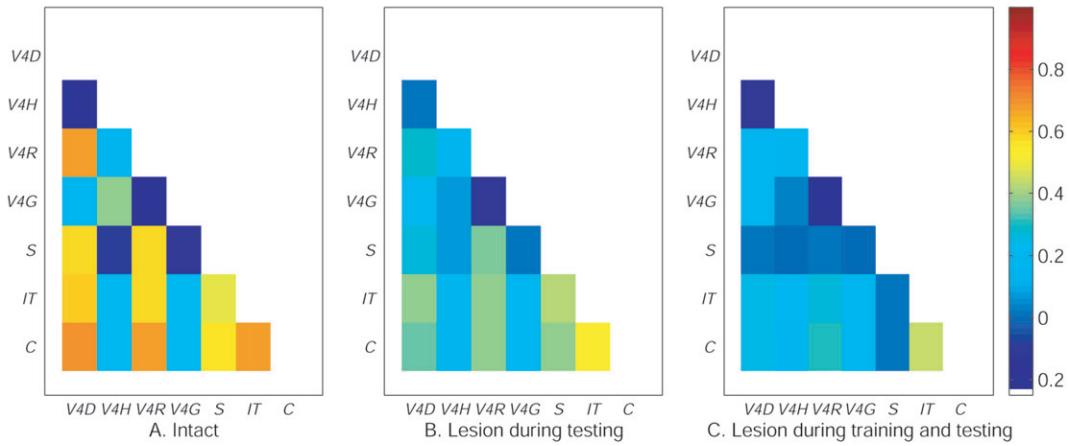

Figure 10. Phase correlations among neural areas for a single subject conditioned to prefer a red diamond target. The color scale ranges from dark blue (no correlation) to dark red (highly correlated). Shown are the Spearman's correlation coefficients between pairs of neural areas, calculated by averaging over all approaches to target/distracter pairs during testing. (*A*) Reentrant connections intact during training and testing. (*B*) Reentrant connections intact during training, but lesioned during testing. (*C*) Reentrant connections lesioned during both training and testing. The correlations between neural areas associated with the target are considerably higher with reentrant connections intact than in either lesion case.

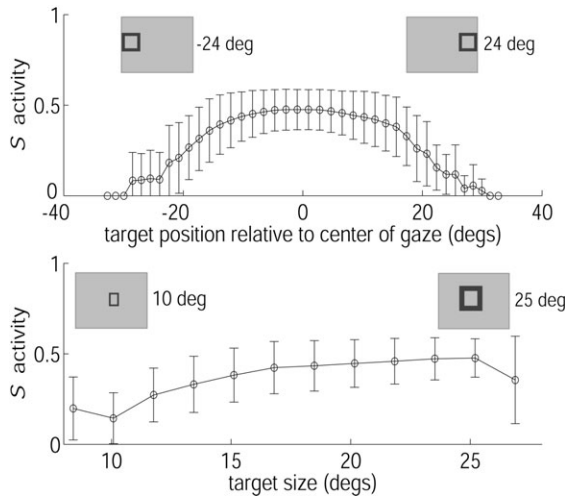

Figure 11. Response of area *S* to target objects in different positions and at different scales. Average values were calculated for all approaches by all Darwin subjects to all targets, error bars indicate standard errors. The top panel shows average responses as a function of target position within the visual field (~±35°). The bottom panel shows average responses as the apparent target size ranged from 8 to 27° of visual angle. The insets in both panels indicate how the target appears in Darwin VIII's field of view at the extreme positions and scales.

which, after conditioning, responded preferentially to target objects over distracters due to plasticity in the pathway *IT*◊*S*. In a typical approach, as Darwin moved from one side of the enclosure to the other, area *S* responded briskly and in phase with neuronal units in areas *V2*, *V4* and *IT* corresponding to attributes of the target. Calculating average values over all Darwin VIII subjects and all target shapes, we found that area *S* responded reliably to target images which appeared within ±20° of the center of Darwin VIII's field of view (the range of the visual field was ~±35°) and as the apparent target size ranged from 8 to 27° of visual angle (see Fig. 11). Thus, Darwin VIII's object recognition was both position and scale invariant.

Value System Activity during Conditioning

As a result of value-dependent synaptic plasticity during conditioning, the visual attributes of target objects became predictive of value. Figure 12 illustrates neural dynamics from a single, intact Darwin VIII subject during both early and late stages of conditioning. As shown in Figure 12 (left), during early conditioning area *S* does not become active until the UCS (unconditioned stimulus; i.e. the tone) is present. The UCS also evokes biases in areas *IT* and *C*, as shown by the rapid abolition of the initially bimodal phase distributions in these areas.

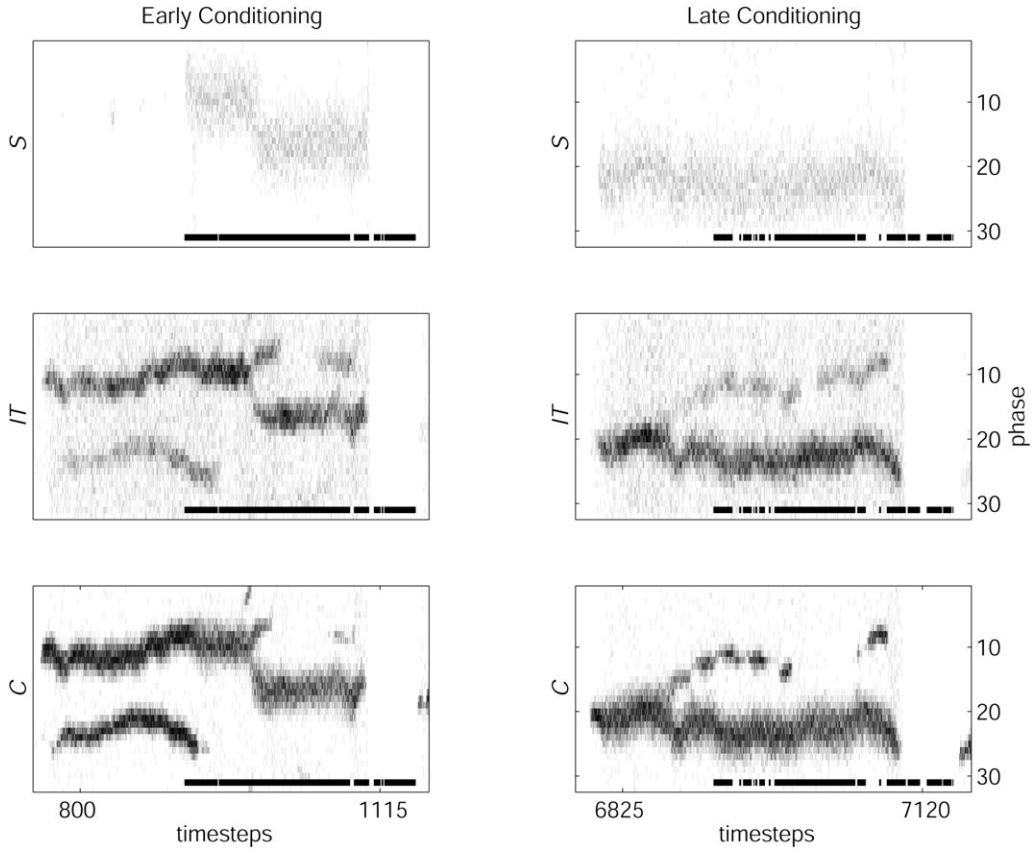

Figure 12. Neural activity during conditioning for a single Darwin VIII subject, showing areas *S*, *IT* and *C* during a single approach to a target shape (red diamond) at an early stage (left panels, time-steps 750–1165), and at a late stage of conditioning (right panels, time-steps 6775–7170). Each panel shows the distribution of neuronal unit phases in the corresponding area over time. As in Figure 9, a gray scale indicates the proportion of neuronal units in each area at a particular phase. The solid line at the bottom of each panel indicates time-steps for which the tone was present. In the early approach (left panels) area *S* is inactive until tone onset, at which point it becomes strongly activated in phase with the upper traces in both *IT* and *C*, which are associated with the target. The lower traces in *IT* and *C*, corresponding to the distracter, become relatively suppressed at the same time. Later in conditioning (right panels), areas *S*, *IT* and *C* are in phase with visual system activity corresponding to the target (lower trace) well before tone onset, and activity associated with the distracter is relatively suppressed well before tone onset.

At a later stage of conditioning, the CS (the conditioned stimulus; i.e. the target visual features) has become associated with value such that activity in *S* now precedes UCS onset (see Fig. 12, right). *S* responds to the target stimulus as soon as the stimulus appears in Darwin VIII's visual field. Activity in *S* then facilitates a bias in areas *IT* and *C*, as shown in Figure 12 by the appearance of a single phase distribution peak in each area well before UCS onset. This shift in the timing of value-related activity, from activity triggered by the auditory UCS in early trials, to activity triggered by the visual CS in later trials, is analogous to the shift in dopaminergic neural activity found in the primate ventral tegmental area during conditioning (Schultz *et al.*, 1997). Value-dependent synaptic plasticity is also similar to 'temporal-difference' learning in that the conditioned stimulus becomes predictive of value (Friston *et al.*, 1994; Montague *et al.*, 1996; Song and Abbott, 2001).

Discussion

We have described a brain-based device, controlled by a simulated nervous system, which bound the visual attributes of distinct stimuli. Binding in Darwin VIII occurred as a result of multilevel interactions involving a reentrant neuroanatomy, the dynamic synchronization of neuronal groups, and the correlations generated by synaptic plasticity and autonomous behavior. We were able to observe Darwin VIII's behavior while simultaneously recording the state of its simulated nervous system at all levels. Specifically, during approaches to visual objects we observed the formation of synchronously active neuronal circuits for each object in Darwin VIII's visual field. These circuits, which were enabled by reentrant connections within and among neural areas, gave rise to motor area activity which in turn evoked discriminatory behavior. Such detailed observations, which are impossible to obtain in living animals, provide insights into the complex, dynamic interactions between brain, body, and behavior that underlie effective visual object recognition.

As with previous brain-based devices (Edelman *et al.*, 1992; Almassy *et al.*, 1998; Krichmar *et al.*, 2000; Krichmar and Edelman, 2002), Darwin VIII had innately specified behavior (i.e. tracking towards auditory or visual stimuli) and innately specified value or salience for certain environmental signals (e.g. positive value of sound). Darwin VIII learned autonomously to associate the value of the sound with the attributes of the visual stimulus closest to the sound source. In every test trial, it successfully oriented towards the target object based on visual attributes alone (see Fig. 7*A*).

The physical embodiment of Darwin VIII was essential for incorporating many of the challenging aspects of this discrimination task, such as variations in the position, scale and luminosity of visual images, sound reflections and slippages during movement. Reliance on elaborate computer simulations risks introducing *a priori* biases in the form of implicit instructions governing interactions between an agent and its environment (Brooks, 1991; Krichmar and Edelman, 2002). By the use of a real-world environment, not only is the risk of introducing such biases avoided, but also the experimenter is freed from the substantial burden of constructing a highly complex simulated environment (Edelman *et al.*, 1992).

The simulated nervous system of Darwin VIII contained cortical areas analogous to the ventral occipito-temporal stream of the visual system (areas *V2*, *V4* and *IT*) and the motor system (area *C*), as well as reward or value systems (area *S*) analogous to diffuse ascending neuromodulatory systems. However, none of these specialized areas or preferential directions of information flow (e.g. 'top-down' or 'bottom-up') were by themselves sufficient for binding the features of visual objects. Rather, visual binding in Darwin VIII was achieved through the interaction of local processes (i.e. activity in each simulated neural area) and global processes (i.e. emergent functional circuits characterized by synchronous activity distributed throughout the simulated nervous system). Reentrant connections among distributed neural areas were found to be essential for the formation of these circuits (see Figs 9, 10, and 12) and for successful performance in a task requiring discrimination between multiple objects with shared features (see Fig. 7). It was striking that reliable discriminations were achieved by Darwin VIII despite the continual changes in the scale and position of stimuli in the visual field which resulted from self-generated movement in a rich real-world environment (see Fig. 11).

The state of each neuronal unit in Darwin VIII was described by both a firing rate variable and a phase variable, where postsynaptic phase tends to be correlated with the phase of the most strongly active presynaptic inputs. This modeling strategy provided the temporal precision needed to represent neural synchrony, without incurring the computational costs associated with modeling of the spiking activity of individual neurons. While representation of precise spike timing is necessary for modeling certain neuronal interactions (Senn *et al.*, 2001; Song and Abbott, 2001), our model suggests that for the purposes of illustrating a possible mechanism for visual binding, such detail is not required. It is also important to emphasize that phase in our model is not intended as a reflection of possible underlying oscillatory activity; specifically, it should not be taken to imply that regular brain oscillations at specific frequencies are an essential component of the neural mechanisms of binding (Gray and Singer, 1989).

Although local regions in Darwin VIII's simulated nervous system had segregated functions based on their input and connectivity, object recognition and discriminative behavior was an emergent property of the whole system, not of any individual area. Darwin VIII's neural responses during an orienting movement toward a target showed this global property in terms of synchronized activity among a dynamic set of neuronal units in different neural areas (see Figures 8 and 9*A*). The simultaneous viewing of two objects clearly evoked two distinct sets of circuits that were distributed throughout the simulated nervous system and distinguished by differences in the relative timing of their activity. When the reentrant connections between neural areas were removed via simulated lesions, coherent interactions among Darwin VIII's neural areas were disrupted (see Figs 9*B* and 10*B*,*C*) resulting in failures in both perceptual categorization and discriminative behavior (see Fig. 7*B*).

Both experience and value shape the global properties of the simulated nervous system. This is clearly shown in Figure 12, where, during early training, area *S* showed no activity and area *C* showed no bias toward the target object until the onset of the auditory cue. Late in the training, area *S* became active well before the auditory cue onset as a result of the value-dependent plastic connections from *IT* to *S*. Activity in area *S* therefore became predictive of the unconditioned stimuli (i.e. the auditory tone). Value-dependent plastic connections from *IT* to *C*

and excitatory connections from *S* to *C* ensured that this shift in the timing of value-related activity resulted in a bias in the activity of area *C* which favored movement toward the target in preference to the distracter. These observations emphasize the role of value systems in modifying the efficacy of distributed neural connections to assure adaptive behavior.

Successful performance in the discrimination task required the complementary action of neural synchrony and experience-dependent changes in neuronal firing rates (see Table 3). Neuronal synchrony, which was indicated by groups of neuronal units sharing a similar phase, was necessary for the formation of multiple global circuits corresponding to each object in view. At the same time, the activity of the neuronal units within these circuits influenced activity levels in areas *V4*, *IT* and *C*, causing Darwin VIII to favor the target over the distracters. These observations suggest that mean firing rate 'codes' (Shadlen and Movshon, 1999) and synchrony-based 'codes' (Gray, 1999; Singer, 1999) need not be considered as mutually exclusive explanations of neuronal function.

A prediction of our model, in which neuronal units represent the activity of small groups of neurons, is that neural synchrony at the group level, rather than zero phase lag among individual neurons, may be sufficient for sensory binding. Although some single-unit recording studies have shown that neurons activated by attended stimuli are more synchronized than neurons activated by unattended stimuli (Steinmetz *et al.*, 2000; Fries *et al.*, 2001), synchronous activity among single units has been difficult to detect in tasks requiring binding (Thiele and Stoner, 2003). Also, microelectrode recordings from primate prefrontal cortex have shown higher levels of correlated firing among local, inhibitory neurons than among excitatory, long-range pyramidal neurons (Constantinidis and Goldman-Rakic, 2002). On the other hand, neuromagnetic recordings of human subjects during binocular rivalry have shown an increase in the intra- and inter-hemispheric coherence of signals associated with a perceptually dominant stimulus, as compared to a stimulus which is not consciously perceived (Srinivasan *et al.*, 1999). However, neuromagnetic signals do not reflect reentrant relations between single neurons; rather, they represent averages across large neuronal populations. This evidence is therefore consistent with our model in suggesting that synchrony can operate at a neuronal group level as well as at the single neuron level.

Higher brain function depends on the cooperative activity of the entire nervous system, reflecting its morphology, its dynamics, and its interactions with the body and the environment. In accord with theoretical views emphasizing the importance of binding through synchrony (Edelman, 1993; Singer, 1999; Engel *et al.*, 2001), we found that visual binding and object discrimination can arise as a result of the constraints reentry and behavior impose on interactions between local processes (activity in particular neural areas) and global processes (synchronously active and broadly distributed neural circuits). This interaction between these processes was essential, and neither specialized areas nor deterministic preferential directions of information flow were alone sufficient to achieve visual binding.

Notes

This work was supported by the W.M. Keck Foundation and the Neurosciences Research Foundation. We thank J.A. Snook, D. Moore and D. Hutson for their contribution to the design of Darwin VIII. We are grateful to Drs E. Izhikevich, J.A. Gally, R. Greenspan and G.N. Reeke, and three anonymous reviewers for detailed criticism.

Address correspondence to Anil K. Seth, W.M. Keck Laboratory, The Neurosciences Institute, 10640 John Jay Hopkins Drive, San Diego, CA 92121, USA. Email: seth@nsi.edu.

Appendix: Specifics of Sensory Input and Computation in Darwin VIII

A. Visual System and its Input

The CCD camera sent 320 × 240 pixel RGB video images, via an RF transmitter, to an ImageNation PXC200 frame grabber attached to one of the workstations running the neural simulation (see Appendix C). The image was spatially averaged to produce an 80 × 60 pixel image. Gabor filters were used to detect edges of vertical, horizontal, and diagonal (45 and 135°) orientations. The output of the Gabor function mapped directly onto the neuronal units of the corresponding *V1* subarea. Color filters (red positive center with a green negative surround, or red negative center with a green positive surround) were applied to the image. The outputs of the color filters were mapped directly onto the neuronal units of *V1-red* and *V1-green*. *V1* neuronal units projected retinotopically to neuronal units in *V2* (see Fig. 2 and Table 2).

B. Auditory System and its Inputs

Microphone input was amplified and filtered in hardware. An RMS (root mean square) chip measured the amplitude of the signal and a comparator chip produced a square waveform which allowed frequency to be measured. Every millisecond, the microcontroller on NOMAD calculated the overall microphone amplitude by averaging the current signal amplitude measurement with the previous three measurements. The microcontroller calculated the frequency of the microphone signal at each time point by inverting the average period of the last eight square waves. *Mic-left* and *Mic-right* responded only to tones between 2.9 and 3.5 kHz having an amplitude of at least 40% of the maximum. The activity of a neuronal unit in *Mic-left* or *Mic-right* was given by

$$s_i^{mic}(t+1) = \tanh(0.9 s_i^{mic}(t) + 0.1 a_i^{mic})$$

where $s_i^{mic}(t)$ is the previous value of a neuronal unit i in *Mic-left* or *Mic-right*, and a_i^{mic} is the current amplitude of the microphone output. This equation served to smooth out spurious noise in the auditory signal.

C. Computation

The simulated nervous system was implemented on a Beowulf cluster of 12 × 1.4 GHz Pentium IV workstations running Message Passing Interface (MPI) parallel software under the Linux operating system. One of the workstations received visual input via RF video transmission from a CCD camera mounted on NOMAD, the mobile base (see Appendix A). The workstation also received auditory (see Appendix B) and infrared detector information, and transmitted motor and actuator commands via an RF modem. A microcontroller (PIC17C756A) onboard NOMAD sampled input and status from the sensors and controlled RS-232 communication between the NOMAD base and the workstations running the neural simulation.

References

Abeles M (1982) Role of the cortical neuron: integrator or coincidence detector? Isr J Med Sci 18:83–92.

Almassy N, Edelman GM, Sporns O (1998) Behavioral constraints in the development of neuronal properties: a cortical model embedded in a real-world device. Cereb Cortex 8:346–361.

Aston-Jones G, Bloom FE (1981) Norepinephrine-containing locus coeruleus neurons in behaving rats exhibit pronounced responses to non-noxious environmental stimuli. J Neurosci 1:887–900.

Azouz R, Gray CM (2000) Dynamic spike threshold reveals a mechanism for synaptic coincidence detection in cortical neurons *in vivo*. Proc Natl Acad Sci USA 97:8110–8115.

Bi GQ, Poo MM (1998) Synaptic modifications in cultured hippocampal neurons: dependence on spike timing, synaptic strength, and postsynaptic cell type. J Neurosci 2:32–48.
Bienenstock EL, Cooper LN, Munro PW (1982) Theory for the development of neuron selectivity: orientation specificity and binocular interaction in visual cortex. J Neurosci 2:32–48.
Brooks R (1991) Intelligence without representation. Artif Intell 47:139–160.
Chelazzi L, Duncan J, Miller EK, Desimone R (1998) Responses of neurons in inferior temporal cortex during memory-guided visual search. J Neurophysiol 80:2918–2940.
Constantinidis C, Goldman-Rakic PS (2002) Correlated discharges among putative pyramidal neurons and interneurons in the primate prefrontal cortex. J Neurophysiol 88:3487–3497.
Edelman GM (1987) Neural Darwinism: the theory of neuronal group selection. New York: Basic Books.
Edelman GM (1993) Neural Darwinism: selection and reentrant signaling in higher brain function. Neuron 10:115–125.
Edelman GM, Reeke GN, Gall WE, Tononi G, Williams D, Sporns O (1992) Synthetic neural modeling applied to a real-world artifact. Proc Natl Acad Sci USA 89:7267–7271.
Engel AK, Fries P, Singer W (2001) Dynamic predictions: oscillations and synchrony in top-down processing. Nat Rev Neurosci 2:704–716.
Fries P, Reynolds JH, Rorie AE, Desimone R (2001) Modulation of oscillatory neuronal synchronization by selective visual attention. Science 291:1560–1563.
Friston KJ, Tononi G, Reeke GN Jr, Sporns O, Edelman GM (1994) Value-dependent selection in the brain: simulation in a synthetic neural model. Neuroscience 59:229–243.
Georgopoulos AP, Schwartz AB, Kettner RE (1986) Neuronal population coding of movement direction. Science 233:1416–1419.
Gray CM (1999) The temporal correlation hypothesis of visual feature integration: still alive and well. Neuron 24:31–47.
Gray CM, Singer W (1989) Stimulus-specific neuronal oscillations in orientation columns of cat visual cortex. Proc Natl Acad Sci USA 86:1698–1702.
Grossberg S (1999) The link between brain learning, attention, and consciousness. Conscious Cogn 8:1–44.
Izhikevich EM, Desai NS (2003) Relating STDP to BCM. Neural Comput 15:1511–1523.
Knoblauch A, Palm G (2002a) Scene segmentation by spike synchronization in reciprocally connected visual areas. I. Local effects of cortical feedback. Biol Cybern 87:151–167.
Knoblauch A, Palm G (2002b) Scene segmentation by spike synchronization in reciprocally connected visual areas. II. Global assemblies and synchronization on larger space and time scales. Biol Cybern 87:168–184.
Konig P, Engel AK, Singer W (1996) Integrator or coincidence detector? The role of the cortical neuron revisited. Trends Neurosci 19:130–137.
Krichmar JL, Edelman GM (2002) Machine psychology: autonomous behavior, perceptual categorization and conditioning in a brain-based device. Cereb Cortex 12:818–830.
Krichmar JL, Snook JA, Edelman GM, Sporns O (2000) Experience-dependent perceptual categorization in a behaving real-world device. In: Animals to animats 6: Proceedings of the Sixth International Conference on the Simulation of Adaptive Behavior. (Wilson SW, ed.), pp. 41–50. Cambridge, MA: MIT Press/Bradford Books.
Lee C, Rohrer WH, Sparks DL (1988) Population coding of saccadic eye movements by neurons in the superior colliculus. Nature 332:357–360.
Montague PR, Dayan P, Sejnowski TJ (1996) A framework for mesencephalic dopamine systems based on predictive Hebbian learning. J Neurosci 16:1936–1947.
Pashler HE (1999) The psychology of attention. Cambridge, MA: MIT Press/Bradford Books.
Raizada RD, Grossberg S (2003) Towards a theory of the laminar architecture of cerebral cortex: computational clues from the visual system. Cereb Cortex 13:100–113.
Reeke GN, Sporns O, Edelman GM (1990) Synthetic neural modeling: the 'Darwin' series of recognition automata. Proc IEEE 78:1498–1530.
Reynolds JH, Pasternak T, Desimone R (2000) Attention increases sensitivity of V4 neurons. Neuron 26:703–714.
Schultz W, Dayan P, Montague PR (1997) A neural substrate of prediction and reward. Science 275:1593–1599.
Senn W, Markram H, Tsodyks M (2001) An algorithm for modifying neurotransmitter release probability based on pre- and postsynaptic spike timing. Neural Comput 13:35–67.
Shadlen MN, Movshon JA (1999) Synchrony unbound: a critical evaluation of the temporal binding hypothesis. Neuron 24:67–77.
Shafritz KM, Gore JC, Marois R (2002) The role of the parietal cortex in feature binding. Proc Natl Acad Sci USA 99:10917–10922.
Singer W (1999) Neuronal synchrony: a versatile code for the definition of relations? Neuron 24:49–65.
Song S, Abbott LF (2001) Cortical development and remapping through spike timing-dependent plasticity. Neuron 32:339–350.
Sporns O, Tononi G, Edelman GM (1991) Modeling perceptual grouping and figure-ground segregation by means of active reentrant connections. Proc Natl Acad Sci USA 88:129–133.
Sporns O, Almassy N, Edelman GM (2000) Plasticity in value systems and its role in adaptive behavior. Adapt Behav 8:129–148.
Srinivasan R, Russell DP, Edelman GM, Tononi G (1999) Increased synchronization of neuromagnetic responses during conscious perception. J Neurosci 19:5435–5448.
Steinmetz PN, Roy A, Fitzgerald PJ, Hsiao SS, Johnson KO, Niebur E (2000) Attention modulates synchronized neuronal firing in primate somatosensory cortex. Nature 404:187–190.
Thiele A, Stoner G (2003) Neuronal synchrony does not correlate with motion coherence in cortical area MT. Nature 421:366–370.
Tononi G, Sporns O, Edelman GM (1992) Reentry and the problem of integrating multiple cortical areas: simulation of dynamic integration in the visual system. Cereb Cortex 2:310–335.
Treisman A (1996) The binding problem. Curr Opin Neurobiol 6:171–178.
Treisman A (1998) Feature binding, attention and object perception. Phil Trans R Soc Lond B Biol Sci 353:1295–1306.
Ungerleider LG, Haxby JV (1994) 'What' and 'where' in the human brain. Curr Opin Neurobiol 4:157–165.
von der Malsburg C, Buhmann J (1992) Sensory segmentation with coupled neural oscillators. Biol Cybern 67:233–242.
Wray J, Edelman GM (1996) A model of color vision based on cortical reentry. Cereb Cortex 6:701–716.