Experience-Dependent Perceptual Categorization in a Behaving Real-World Device

Jeffrey L. Krichmar, James A. Snook, Gerald M. Edelman and Olaf Sporns
The Neurosciences Institute
10640 John Jay Hopkins Drive
San Diego, CA 92121, USA
krichmar@nsi.edu, snook@nsi.edu, edelman@nsi.edu, sporns@nsi.edu

Abstract

The ability of organisms to categorize objects depends on their sensory experience in an environment. We studied the role of frequency and temporal order of stimuli in perceptual categorization on Darwin VI, a neuronal model interfaced with a behaving real-world device. The model consisted of several distinct biologically based networks, representing areas of the central nervous system. Darwin VI was trained to perform a conditioning task, using objects differing in their visual patterns as well as taste. Initially, behaviors were triggered by taste. After conditioning, responses were triggered by pattern vision, and the simulated visual system showed specific and invariant activity patterns for different categories of visual stimuli. We found that visual system activity devoted to an infrequently sampled stimulus class, during early exploration, resulted in a reduction of activity that remained even if the stimulus was sampled at normal rates later. In contrast, frequently sampled stimulus classes resulted in more visual system activity than normal. Presenting exemplars of only two out of three stimulus classes to a conditioned Darwin VI resulted in a significant loss of visual system neurons capable of responding to the deprived stimulus. These studies point to the importance of early sensory experience for the development of perceptual categories.

1. Introduction

Most organisms are capable of discriminating and categorizing a vast number of sensory stimuli, and continue to acquire new perceptual categories throughout their lifetime (Edelman, 1987). In this paper, we explore the hypothesis that perceptual categorization is a highly dynamic process that is strongly dependent upon the actual sequence and content of sensory experience and individual histories of stimulus encounters. We use a model system, Darwin VI, instantiated in the real world as a tool to explore the conjunction of neural and behavioral processes subserving categorization.

Despite its importance for all aspects of cognition and behavior, the biological bases of perceptual categorization are still relatively unexplored. Traditionally, the problem of object recognition and categorization has been cast in the framework of information processing (Barsalou, 1992; Johnson-Laird, 1988). After sensory patterns are acquired and processed, the sensory information is transformed into lists of structural features of objects. These lists are then compared to stored representations of categories. In information processing models, the most similar representation is selected through matching the actual input to a set of stored templates. Retrieving additional stored information draws further inferences about the object. Essentially, classical models of categorization assume that all members of a category are uniquely characterized by a list of singly necessary and jointly sufficient "defining" features. Learning then consists of a procedure to assemble and store such lists allowing the recognition of other exemplars. Supervised as well as unsupervised learning algorithms have been devised, particularly in artificial neural network models, to form adaptive mappings between input patterns and high-level network correlates of categories.

In contrast to the classical, information-processing view of categorization, it has been recognized that category membership is not all-or-none, but is a matter of degree for most natural categories (Mervis and Rosch, 1981). Categories are defined by a "family resemblance" of its members, or a polymorphous set of features (Ryle, 1949), with some members of the category being judged as more typical than others. "Prototype" and "exemplar" theories of categorization have been proposed, but the nature of the elementary features used in categorizing even simple stimuli has remained the subject of debate in cognitive science. This debate is concerned with whether or not such features are created in the course of actual experience (e.g. Schyns et al., 1998). Evidence indicating that the formation of perceptual categories is dependent on the actual sensory experience of individual organisms supports a more dynamic view of how

categories are formed. Experimental studies (Thelen and Smith, 1994; Smith et al., 1999), as well as computer modeling (Edelman et al., 1992; Scheier and Lambrinos, 1996; Pfeifer and Scheier, 1997), suggest that sensorimotor activity and interactions between an organism and its environment are critical in extracting relevant feature dimensions and generating object constancy. For example, bodily activity and self-generated movement of an organism contribute to creating temporal correlations that facilitate the emergence of various kinds of perceptual invariances (Almassy et al., 1998). The central role of sensorimotor activity provides an important rationale for "embodied" models of categorization, i.e. models that include movements of sensors and effectors as key ingredients of category formation (for a review see Pfeifer and Scheier, 1999).

The observed experience-dependence of perceptual categories in humans (Markman, 1989; Eimas, 1994) and other species may have a neural basis in the plasticity and dynamic reorganization of large-scale neuronal networks under conditions of varying sensory input (for a recent review, see Buonomano and Merzenich, 1998). Visual, auditory, somatosensory and motor cortices of primates have been shown to undergo plastic changes in the configuration of cortical maps in relationship to passively induced changes in sensory input patterns, or during extensive behavioral training (Elbert et al., 1995; Nudo et al., 1996). Often, increased exposure to particular stimuli increases the size of cortical areas devoted to these stimuli, while decreased exposure (deprivation) leads to a shrinking or even a disappearance of the corresponding areas. Such plastic changes are observed over a time course of hours to days and can persist throughout adult life (Garraghty and Kaas, 1992; Diamond et al., 1993; Merzenich and deCharms, 1996). While most studies addressing cortical reorganization have been carried out in primary sensory areas, there is evidence that higher cortical areas are subject to similar phenomena. In the inferior temporal cortex (IT), an area that is primarily involved in the recognition of visual objects, recent experiments have demonstrated that prolonged exposure to a training set of visual stimuli leads to the emergence of a larger number of cells that show specific responses to members of the set (Kobatake et al., 1998; see also Rodman, 1994; Sakai and Miyashita, 1991). While it is assumed that the activity of such neurons underlies the organism's capability to categorize visual stimuli, it is less clear how the behavioral activity and the stimulus content of the environment affects an organism's development.

In this paper, we investigate the possible impact of different experiential histories on the formation and maintenance of cortical maps involved in visual object recognition. The use of a biologically based computer simulation interfaced with a behaving real-world device allows us to study the effects of systematic variations in the stimulus content of an environment, as well as of variations in the sequences of stimulus encounters. Darwin VI's nervous system consisted of several distinct networks, simulating areas of the primary visual cortex, IT, motor cortex, as well as a value system. Darwin VI was equipped with a CCD camera and a mechanical gripper capable of picking up and "tasting" small blocks marked by a variety of visual patterns. "Taste" corresponded to the surface conductivity of the blocks and served as an unconditioned stimulus, triggering either appetitive or aversive behavioral responses. Initially, Darwin VI approached and "tasted" all blocks indiscriminately, reflecting its inability to categorize objects visually. This was due to the initial lack of neuronal units in IT with response properties, such as object-selectivity, that could support perceptual categorization. After a number of encounters with stimuli of different classes, and due to synaptic plasticity in sensory afferents to IT and in sensorimotor connections, Darwin VI's appetitive and aversive behaviors were triggered by the conditioned stimulus (i.e. pattern vision) instead of taste (see also Almassy et al., 1998). IT showed specific and invariant activity patterns for several different categories of visual stimuli. We have now found that categorization of different stimulus patterns depended on the order and frequency of their presentation. Our results suggest a number of neural mechanisms underlying the experience-dependence of perceptual categorization.

2. Implementation

2.1 Device

The behaving real-world device (NOMAD II) consists of a mobile base equipped with several sensors and effectors, and is capable of communicating with a neural simulation running on remote computers. NOMAD II was constructed with a radio modem and RF transmission of video output to allow untethered exploration, pan and tilt movement for its CCD camera, and object gripping with a one degree-of-freedom manipulator (see Figure 1). The mobile device was constructed on a circular platform (developed by Nomadic Technologies Inc., Mountain View CA) with wheels that permitted independent translational and rotational motion. A miniature CCD camera and sensors embedded in the gripper measuring the surface conductivity of stimuli provided sensory input to the neuronal simulation. Eight infrared (IR) sensors were mounted at 45-degree intervals around the mobile platform. The IR sensors were responsive to the boundaries of the environment and were used to trigger obstacle avoidance reflexes.

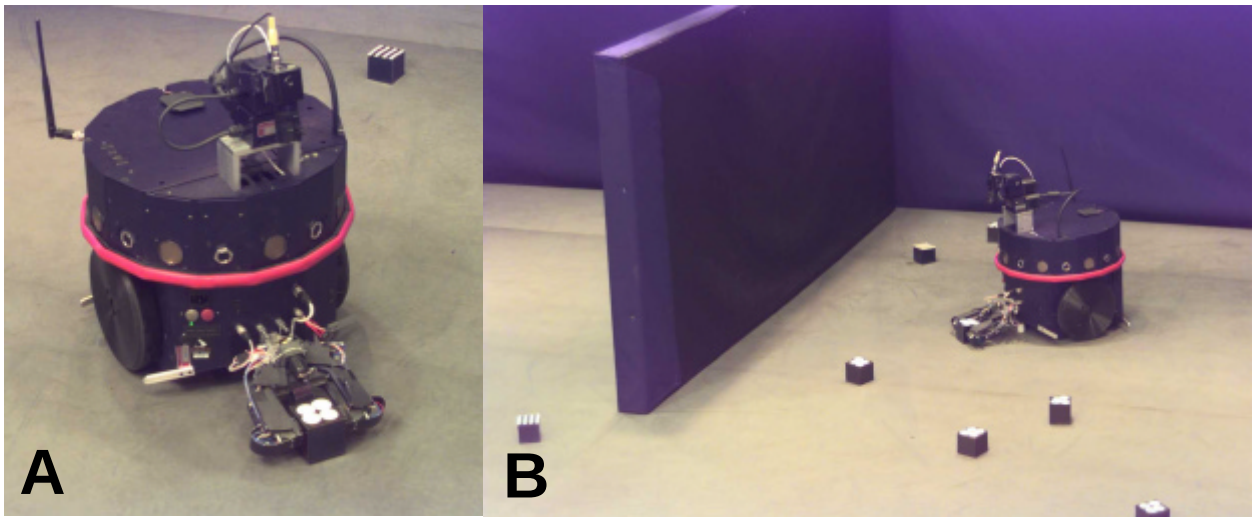

Figure 1. NOMAD II real-world behaving device. A. Close up of NOMAD II picking up an object with its gripper and looking at the object with its CCD camera. B. NOMAD II explored an environment that was partly separated by the divider shown in the photo. Disparate proportions of stimulus objects were on either side of the divider.

Continuous interaction between the simulated nervous system, which ran on a computer server, and the NOMAD II behaving device was achieved by RS-232 communication. The interface between the sensors, effectors on board NOMAD II did not initiate or control any behavior. All behavioral activity other than obstacle avoidance was triggered by signals received from the neuronal simulation.

Input to the visual system consisted of a 64x64 gray level pixel array, derived by spatial averaging of a central 218x218 pixel sub-region of a full-resolution video image sampled by the CCD camera at a maximum rate of 30 frames/sec. When an object, identified as a region of high brightness contrast, was sensed directly in front of the device, the CCD camera tilted downward. When IR sensors located in the gripper signaled an object was located between its two manipulators, the gripper closed, the object was picked up gently, and following contact, conductivity ("taste") was measured across its multiple exposed contacts.

2.2 Environment

The environment consisted of an enclosed area with black cloth-covered walls and a floor covered with opaque black plastic panels, on which we distributed 16 stimulus blocks (6 cm metallic cubes; see Figure 1). The top surfaces of these blocks were covered with removable black and white patterns. All other surfaces of the cubes were featureless and black. All experiments reported in this paper were carried out with multiple exemplars of two basic designs: *blobs* (several white patches 2-3 cm in diameter) and *stripes* (width 0.6 cm, evenly spaced). *Stripes* can be viewed either in a horizontal or vertical orientation, yielding a total of three stimulus classes (*blob*, *horizontal* and *vertical*) of visual patterns to be discriminated. The sides of the stimulus blocks were metallic and could be rendered either conductive or non-conductive. In the experiments described in this paper, all blob-patterned blocks were non-conductive, while all the stripes were conductive. The blocks were distributed randomly throughout segments of the environment and were manipulated only by NOMAD II's gripper.

2.3 Behavior

Basic modes of behavior in Darwin VI were similar to those of Darwin V, previously described in detail (Almassy et al. 1998). Briefly, these behaviors were IR-sensor driven obstacle avoidance, visual exploration, visual approach and tracking, gripping and "tasting", and two main classes of behavioral responses. These responses, called "appetitive" and "aversive", could be activated either by taste (the unconditioned stimulus, US, triggering an unconditioned response, UR), or by vision (the conditioned stimulus, CS, triggering a conditioned response, CR). Appetitive behavior consisted of prolonged gripping and "tasting", of a stimulus block, depositing of the block, and then a 90-degree clockwise turn. Aversive behavior consisted of gripping and "tasting" of a stimulus, releasing the block, and then turning 90-degrees counterclockwise. Conditioned appetitive and aversive responses were triggered if the difference in instantaneous activity between motor areas M_{app} and M_{ave} exceeded a behavioral threshold b (here b = 0.3). The appetitive behavior remained unchanged, irrespective of whether it was elicited directly by taste (US-UR) or by vision (CS-CR). If visual inputs (CS) acted to trigger aversive behavior, the turning response was instantaneous (CR) resulting in immediate removal of the aversive object from the visual field; no "tasting" occurred.

2.4 Simulated Nervous System

Darwin VI's nervous system consisted of a number of areas representing different brain regions containing multiple neuronal units of different types, representing local populations of neurons or neuronal groups (Edelman, 1987). The activity of a neuronal unit in the simulation expresses the mean activity of a few hundred neurons over a time interval of hundreds of milliseconds. In the instantiation used in the present experiments, the simulated nervous system contained a total of 9,353 neuronal units and 689,250 synaptic connections. There were five major components: a visual system, a taste system, sets of motor neurons capable of triggering behavior, a visual tracking system, and a value system (see Figure 2). In this section we provide a brief anatomical and functional description of these components. Numerical values of relevant simulation parameters are given in Table 1 and Table 2.

Figure 2 schematically illustrates the anatomy and connectivity of the neuronal simulation. The 64x64 gray level pixel image captured by NOMAD II's camera was relayed to a retinal area R and transmitted via topographic connections to a primary visual area VA_P. The spatial arrangement (resembling feature matrices) of these connections resulted in stimulus selectivity within interleaved sub-partitions of VA_P. There were three sub-partitions, selective for blob-like features, or short horizontal or vertical line segments. Topographic, mutually inhibitory connections resulted in sharpening and disambiguation of responses. Responses within VA_P closely followed stimulus onset and projected non-topographically to a secondary visual area, analogous to the inferior temporal cortex (IT). IT contained local excitatory and inhibitory interactions producing firing patterns that were characterized by focal regions of excitation surrounded by inhibition. Responses in IT tended to be longer lasting than in VA_P. IT sent a topographic and modifiable projection to the value system (S_o) and a non-topographic modifiable projection to the motor areas (M_{app} and M_{ave}).

The behavioral motor responses to taste were innate and reflexive. Two groups were capable of triggering two distinct behaviors (i.e. appetitive, M_{app}, and aversive, M_{ave}). The taste system (T_{app} and T_{ave}) consisted of two kinds of sensory units responsive to either the presence or absence of conductivity across the surface of stimulus objects (as measured by sensors in NOMAD II's gripper). The taste system sent information to the motor areas (M_{app} and M_{ave}) and a uniform, non-plastic input to the value system (S_o).

The visual tracking system of Darwin VI controlled navigational movements, in particular the approach to objects identified by brightness contrast with respect to the background. The retinal area R projected to area C ("colliculus"), containing excitatory (C_e) and inhibitory (C_i) units. The pattern of activity within C helped to sharpen C_e responses to visual targets and partially to disambiguate such responses when multiple targets were present. C triggered translational and rotational motion of NOMAD II (via activation of M_{tra} and M_{rot}, respectively), ultimately producing visual approach behavior. Connection strengths in projections from C to M_{tra} and M_{rot} were assigned initial values resembling distributions obtained in earlier work by value-dependent synaptic modification during sensorimotor training (Edelman et al., 1992).

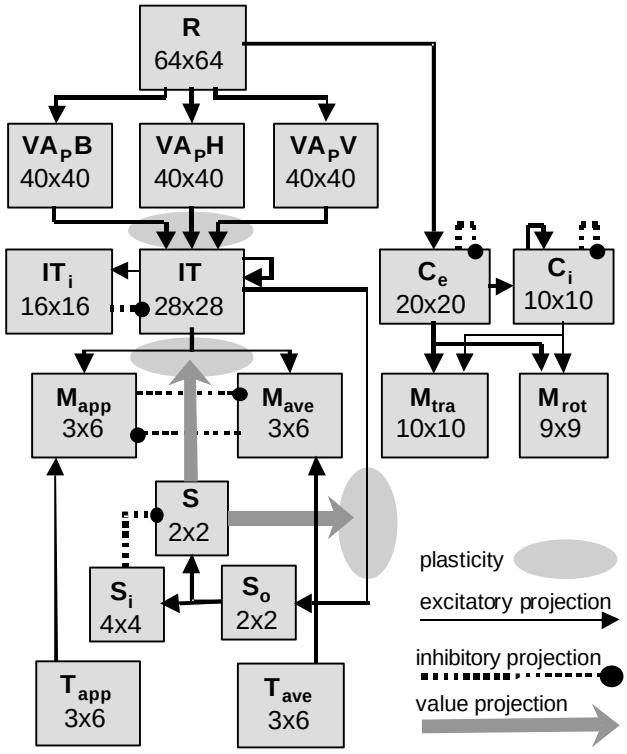

Figure 2. Schematic of the anatomy of Darwin VI's nervous system. The lateral inhibition between the sub-sections of VA_P is omitted for clarity. See text for discussion of areas and Tables 1 and 2 for settings of physiological parameters.

2.5 Cell Activation and Synaptic Rules

After each simulation cycle *t*, which corresponded to approximately 250 ms of real time, the activities $s_i(t)$ of all neuronal units were recomputed. Values for $s_i(t)$ were taken to represent the average firing rate of a local population of neurons. For some connection types (Table 2), strengths of synaptic connections (c_{ij}; from unit *j* to unit *i*) were subject to activity-dependent modification. The inputs to unit *i* could either be voltage-independent or voltage-dependent. The total contribution of input to unit *i* from voltage-independent connections, A^{VI}, is given by:

$$A_i^{VI}(t) = \sum_{l=1}^{M} \sum_{j=1}^{N_l} c_{ij}\, s_j(t)$$

where *M* was the number of different anatomically defined connection types and N_l is the number of connections per type *M* projecting to unit *i*. Negative values for c_{ij} corresponded to inhibitory connections. The total

contribution from voltage-dependent connections, A^{VD}, is given by:

$$A_i^{VD}(t) = \sum_{l=1}^{M} \tanh\left(A_i^{VI}(t) + w\, s_i(t)\right) \sum_{j=1}^{N_l} c_{ij}\, s_j(t)$$

where *w* determined the persistence of unit activity from one cycle to the next. Voltage dependence was simulated so that voltage-dependent connections had no effect unless there was sufficient post-synaptic activity.

The activity level of unit *i* was given by:

$$s_i(t+1) = f\left(\tanh\left(g_i\left(A_i^{VI}(t) + A_i^{VD}(t) + w\, s_i(t)\right)\right)\right) \qquad \text{where}$$

$$f(s) = \begin{cases} 0; & s < \sigma_i \\ s; & \textit{otherwise} \end{cases}$$

and σ_i is a unit specific firing threshold and g_i is a scale factor.

In the model, several types of connections (both voltage-dependent and voltage-independent, see Figure 2) were subject to activity-dependent modification according to a synaptic rule. For some connections, this modification occured independent of value (e.g. IT to M_{app} and M_{ave}), with changes in c_{ij} solely dependent on pre- and post-synaptic activity:

$$\Delta c_{ij}(t) = e\left(c_{ij}(0) - c_{ij}(t)\right) + h\, s_j(t) F\left(s_i(t)\right)$$

where $s_i(t)$ and $s_j(t)$ are activities of post- and pre-synaptic units, respectively, η is a fixed learning rate, ε is a decay constant ($\varepsilon << 1$), and $c_{ij}(0)$ is the initial (t=0) weight of connection c_{ij}. The decay constant ε governed a passive, uniform decay of synaptic weights to their original starting values, implementing a process of "forgetting" that affected all modifiable synaptic weights (see Discussion) and

$$F(t) = \begin{cases} 0; & s < \theta_1 \\ k_1(\theta_1 - s); & \theta_1 \le s < (\theta_1 + \theta_2)/2 \\ k_1(s - \theta_1); & (\theta_1 + \theta_2)/2 \le s < \theta_2 \\ k_2 \tanh\left(r(s - \theta_2)/r\right) & \textit{otherwise} \end{cases}$$

Synaptic changes, which were determined by both pre- and post-synaptic activity and the function *F*, could result in either strengthening or weakening of connections. *F* set limits and rates for synaptic potentiation and depression as a function of post-synaptic activity. It was implemented as a piecewise linear function defined by two thresholds ($0 < \theta_1 < \theta_2 < 1$), two inclinations ($k_1$, k_2) and a saturation parameter r ($r = 6$ throughout) (for details see Sporns et al., 2000).

2.6 Value System

Activation of the value system signaled the occurrence of salient sensory events and leads to a modulation of connection strengths. Area S_o received fixed inputs from the taste system as well as plastic inputs from IT. Area S received excitatory inputs from S_o and inhibitory units from S_i. Area S initiated value-dependent modifications to the IT→S_o, IT→M_{ave} and IT→M_{app}. Area S resembled a diffuse ascending neuromodulatory system in that its units showed uniform phasic responses when activated, and its outputs acted diffusely over multiple pathways by modulating synaptic changes. Value-dependent synaptic changes are computed as:

$$\Delta c_{ij}(t) = e\left(c_{ij}(0) - c_{ij}(t)\right) + h\, s_j(t) F\left(s_i(t)\right) V$$

All parameter designations were identical to the synaptic rule described by Dc_{ij} above, with the addition of the value-dependent term *V*. This term was computed as:

$$V = w_{ik} \sum_{d \in [d_1, d_2]} s_k(t - d)$$

with w_{ik} (mean $w_{ik} = 0.11$) scaling the value input to the afferents of unit s_i, s_k denoting unit activity in area S, with d_1, d_2 denoting upper and lower bound, respectively, of a moving time window over which value inputs to unit s_i are averaged. In the present experiments, $d_1 = 3$ and $d_2 = 9$, providing an averaged value input with a mean time delay of 6 time steps (corresponding to about 1-2 seconds of real time). This formulation of V incorporated three main features of the modeled value signal: (1) its diffuse action, (2) its time-delays (d_1, $d_2 > 0$) and (3) its prolonged (d2 > d1) effect on synaptic plasticity. The neurobiological and functional roles of this simulated value system have been investigated elsewhere (Sporns et al., 2000).

Area	*g*	*s*	*w*
R	1.50	0.20	0.00
VA_p-B	1.33	0.65	0.50
VA_p-H, VA_p-V	1.50	0.50	0.50
IT	1.10	0.04	0.03
IT_i	1.35	0.02	0.15
M_{app}, M_{ave}	2.00	0.10	0.30
T_{app}, T_{ave}	2.00	0.10	0.30
S	2.00	0.05	0.15
S_o	3.00	0.15	0.22
S_i	2.00	0.10	0.22
Ce	1.60	0.10	0.22
Ci	1.00	0.16	0.22
M_{rot}, M_{tra}	2.00	0.05	0.15

Table 1. Values of parameters defining properties of neuronal units in Darwin VI.

Projection	*Prob.*	*Arbor*	$c_{ij}(0)$	h	e	q_1	q_2	k_1	k_2
R→VA_P-B	1.00	O 3x3	0.03	0.0	0.0	0.0	0.0	0.0	0.0
R→VA_P-H,V	1.00	[] 0x4, 4x0	0.04	0.0	0.0	0.0	0.0	0.0	0.0
VA_P-B →VA_P-H,V	1.00	O 2x2	-0.25	0.0	0.0	0.0	0.0	0.0	0.0
VA_P-H ↔VA_P-V	1.00	[] 3x3	-0.0265	0.0	0.0	0.0	0.0	0.0	0.0
VA_P→IT	0.05	[] 30x30	0.022	0.0076	0.0006	0.14	0.29	0.45	0.15
IT→IT_i	1.00	O 5x5	0.015	0.0	0.0	0.0	0.0	0.0	0.0
IT_i →IT	1.00	O 1x1	-0.215	0.0	0.0	0.0	0.0	0.0	0.0
IT →IT	1.00	O 1x1	0.205	0.0	0.0	0.0	0.0	0.0	0.0
IT→IT *	0.70	Θ 13x11	0.0	0.004	0.0004	0.1	0.5	0.45	0.15
IT→M_{app},M_{ave} #	0.40	[] 11x11	0.0016	0.05	0.0005	0.01	0.16	0.1	0.16
IT→S_o #	0.85	O 20x20	0.001	0.02	0.002	0.01	0.18	0.05	0.05
T_{app},T_{ave}→S_o, M_{app},M_{ave}	1.00	O 1x1	0.12	0.0	0.0	0.0	0.0	0.0	0.0
M_{app}↔M_{ave}	1.00	See note 1	-0.12	0.0	0.0	0.0	0.0	0.0	0.0
S_i→S	1.00	O 2x2	-0.285	0.0	0.0	0.0	0.0	0.0	0.0
S_o→S	1.00	O 2x2	0.1	0.0	0.0	0.0	0.0	0.0	0.0
S_o→S_i	0.80	O 2x2	0.055	0.0	0.0	0.0	0.0	0.0	0.0
R→C_e	1.00	O 2x2	0.155	0.0	0.0	0.0	0.0	0.0	0.0
R→C_i	1.00	O 2x2	0.06	0.0	0.0	0.0	0.0	0.0	0.0
C_e →C_e	1.00	See note 2	0.06	0.0	0.0	0.0	0.0	0.0	0.0
C_e →C_i	1.00	[] 1x1	0.105	0.0	0.0	0.0	0.0	0.0	0.0
C_i →C_i	0.80	See note 3	-0.295	0.0	0.0	0.0	0.0	0.0	0.0
C_i →C_i *	0.76	Θ 2x1	0.041	0.0	0.0	0.0	0.0	0.0	0.0
C_e→ M_{rot},M_{tra}	1.00	[] 11x3,3x11	0.07	0.0	0.0	0.0	0.0	0.0	0.0
C_i→ M_{rot},M_{tra} *	1.00	[] 11x3,3x11	0.07	0.0	0.0	0.0	0.0	0.0	0.0

Table 2. Properties of anatomical projections (connection types) in Darwin VI. Arborization shape can be a square [], circular O, surround Θ or a connection mask (see notes below). Values for initial connection strengths, $c_{ij}(0)$, represents means with negative numbers indicating inhibitory connections. Projections marked * are voltage-dependent and # are value-dependent. Non-zero values for h, e,q_1,q_2, k_1 *and* k_2 signify modifiable connection. Note 1: M_{app}↔M_{ave} connection is masked such that each neuronal unit in M_{app} inhibits each neuronal unit in and M_{ave} vice versa. Note 2: C_e →C_e connection is masked such that neuronal units inside the center of the visual field are excited. Note 3: C_i →C_i connection is masked such that neuronal units outside of the center of the visual field are inhibited.

3. Results

All experiments were carried out in the enclosed environment shown in Figure 1B, either with Darwin VI autonomously searching for blocks (Section 3.1) or with blocks placed manually in front of Darwin VI by an experimenter (Section 3.2). Relevant neuronal and behavioral activity was recorded continuously over time and stored for subsequent data analysis.

Neuronal activity in IT provided the basis for visual perceptual categorization in Darwin VI and therefore we focussed our analysis on IT activity patterns. Figure 3 shows examples of such patterns, recorded after visual development was completed with a total of 16 presentations of each stimulus class. As described previously (Almassy et al. 1998; Almassy and Sporns, 1999), many neuronal units in IT were selective for objects of a specific class. In addition, many units showed responses properties that were similar over a large sub-region of the visual field, a property consistent with translation invariance. When compared over multiple experimentals, IT activity patterns showed significant individual variations, while maintaining consistent average numbers of units devoted to one stimulus class.

We performed two separate sets experiments, addressing experience-dependent effects on categorization during early development (Section 3.1) and in the "adult" (Section 3.2). The first set of experiments investigated the effect of variations in stimulus order and frequency on early development. The experiments began with a visually "naïve" Darwin VI exploring an environment that was partially divided into two sub-areas with disparate distributions of stimuli. The second set of experiments investigated the effect of stimulus frequency on a version of Darwin VI that had previous visual experience ("adult" Darwin VI). The experience consisted of exemplars of three stimulus classes presented in equal proportion until categorization was achieved with high accuracy. This "adult" Darwin VI was then presented with exemplars of only two out of the three stimuli classes. An experimenter presented blocks to Darwin VI to ensure that fixed proportions of each stimulus pattern were sampled.

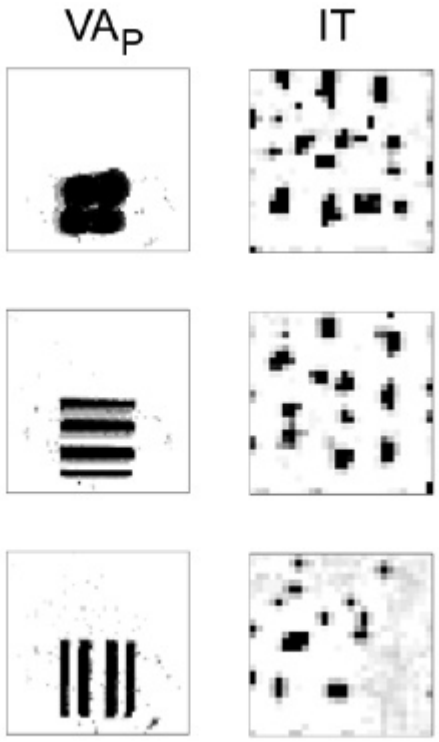

Figure 3: Activity patterns in the visual receiving area R (left column) and in area IT (right column), after the development of visual perceptual categorization is completed. The three stimulus classes (*blobs*, *horizontals* and *verticals*) are represented top to bottom, respectively.

3.2 Effect of Stimulus Frequency and Order on Development

We investigated the effect of stimulus frequency and temporal presentation by partly segregating the enclosure into two equal sized areas with distinct distributions of stimuli (see Figure 1B). One area predominantly contained *blobs*, the other area predominantly contained *stripes*. Although Darwin VI could cross from one side to the other, the divider acted as a partial barrier forcing Darwin VI to spend long time periods sampling primarily one stimulus type. In each of 14 experiments, Darwin VI started with an identical, "naïve" area IT. The activity of the simulated primary visual cortex (VA_P) and IT during stimulus sampling were recorded and stored for analysis.

While the initial state of the "naïve" nervous system at the start of each of the 14 experiments was identical, the "adult" nervous system exhibited systematic differences due to variations in the frequency of stimulus presentation. Each symbol in Figure 4 shows the size of the neuronal response, after Darwin VI explored its environment, to the presentation of a specific stimulus object. The number of neuronal units in IT that were responsive to stimuli of different classes increased significantly with an increase in stimulus presentation frequency (see Figure 4).

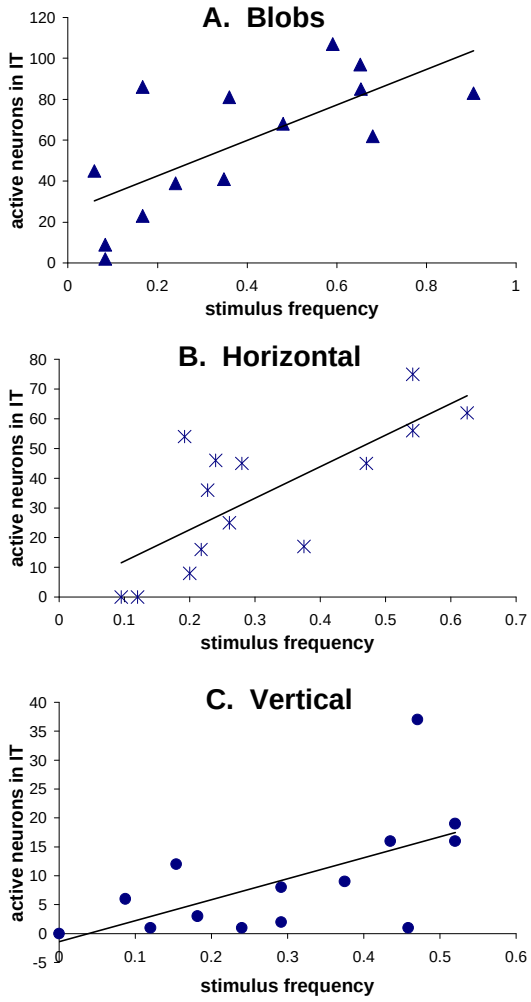

Figure 4. Effect of stimulus frequency on neural activity in IT. Each symbol represents the number of neurons in IT with an activity level $s_i > 0.5$, evaluated at the end of each experiment (N = 14). The stimulus frequency was positively correlated with the number of neurons in IT that have an activity level (s_i) greater than 0.5 (blobs: $r = 0.71$, $p < 0.005$; horizontal stripes: $r = 0.75$, $p < 0.005$; vertical stripes: $r = 0.61$, $p < 0.05$; Pearson's product moment correlation coefficient).

The activity map in area IT exhibited dynamic and competitive changes over an individual experiment. The size of the neural area responsive to objects of a particular class was found to depend on temporal order as well as frequency. Figure 5 shows two representative experiments that illustrate the effect of temporal order on IT's neuronal response to a stimulus class. The panels on the left of Figure 5 show a record of where and in what order the objects were sampled in the experimental arena. The panels on the right side of Figure 5 show the number of active neurons in IT (i.e. $s_i > 0.5$) plotted as a function of the simulation cycle in which a block was sampled. The IT response to each object after the experiment was complete is shown at the right of the figure. If an object class was sampled frequently during early development, the neural area corresponding to that

class expands quickly (e.g., see *blobs* in top row of Figure 5 and *horizontals* in bottom row of Figure 5). At later developmental stages, the extent of neural area occupied by that class was inaccessible for exemplars of other stimulus classes (e.g. *horizontals/verticals* in top row Figure 5, and *blobs/verticals* in bottom row Figure 5). In the experiment depicted in the top row, the number of active neuronal units responsive to *blobs* increased over time, whereas the number of active neurons in IT responding to *stripes* was diminished due to the early allocation of IT neuronal units to *blobs*. In the experiment depicted in the bottom row, the early sampling of *horizontal* objects caused a mapping in IT such that many neurons in IT responded to *horizontals* with a diminished number of neurons responding to *vertical* or *blob* objects.

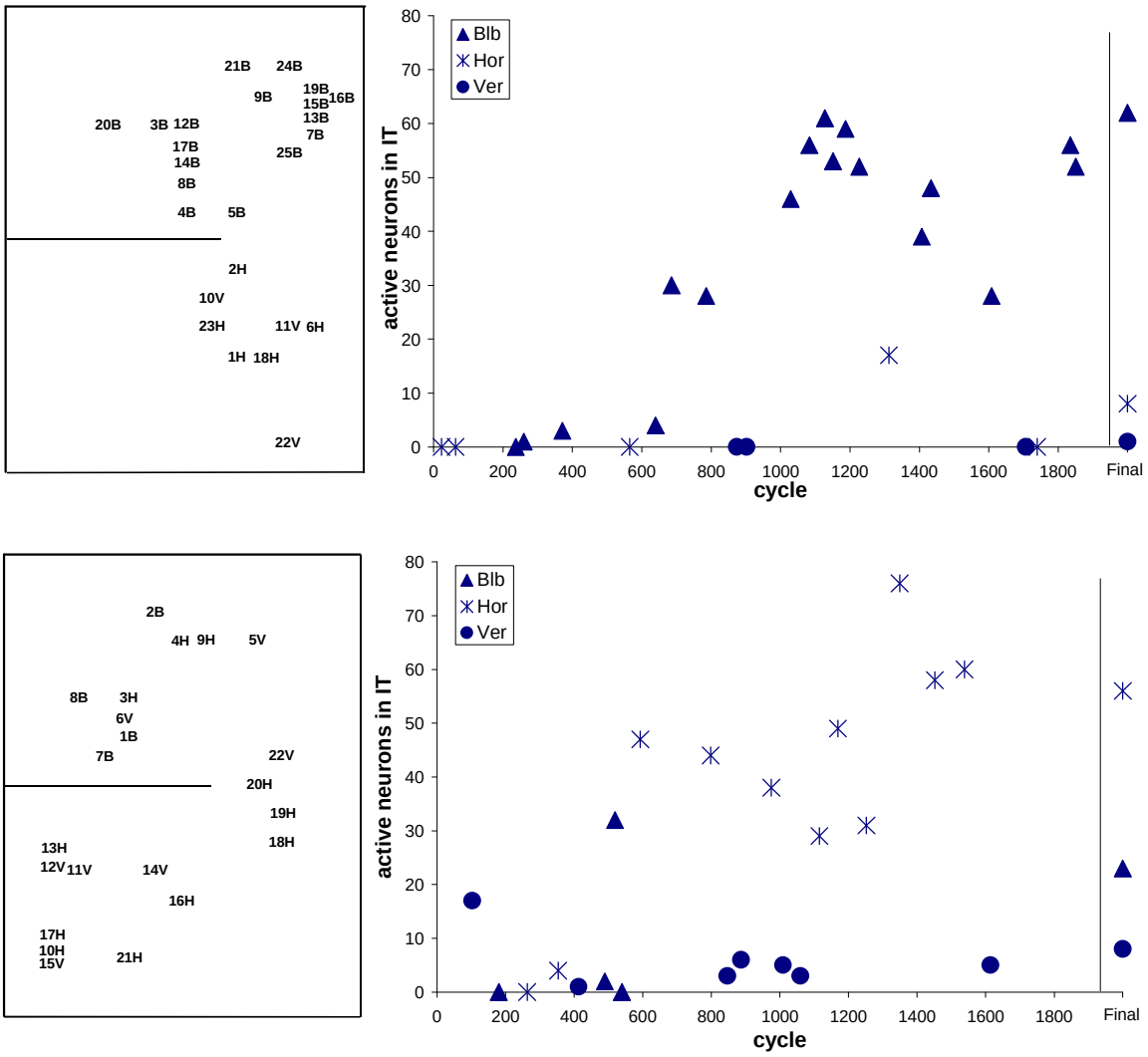

Figure 5: Two representative experiments illustrating the effect of stimulus frequency and temporal order on IT activity during development. The top row shows an experiment in which *blob* objects were sampled early in Darwin VI's development and the bottom row shows an experiment in which *horizontal* objects were sampled early during development. The panels on the left depict the enclosure in which Darwin VI explored. Each number denotes the order, position and type ("B" for *blob*, "H" for *horizontal*, "V" for *vertical*) of a sampled block. The charts on the right show the amount of neural activity in response to an object, measured by the number of neurons in IT with $s_i > 0.5$, as a function of the simulation cycle. The symbols at the right of the chart (marked on the X-axis by Final) denote the IT activity after the experiment was completed.

3.2 Effect of Stimulus Frequency on Late Development

In a second set of experiments we studied the effect of frequency of stimulus presentation on the neural mapping in IT after visual development has reached "adult" levels. To reach these levels of experience, we presented Darwin VI with an equal proportion (8 each) of exemplars of the three object classes until it could discriminate the different objects with high accuracy. We then conducted sets of experiments for each of these four conditions: presenting an additional

eight exemplars of each of the three stimulus classes ("all3"); presenting eight *verticals* and eight *horizontals* ("stripes"); presenting eight *horizontals* and eight *blobs* ("horblob"); or presenting eight *verticals* and eight *blobs* ("verblob"). After all visual training was completed, the weights of the synaptic connections in Darwin VI's nervous system were held static (i.e. learning was turned off), all three stimulus classes were presented, and the neuronal activity of IT was recorded. To ensure that the correct proportion and sequence of blocks was sampled, an experimenter presented the blocks to a behaving Darwin VI. This procedure was performed ten times for each condition (40 experiments altogether).

Figure 6 shows the median number of active neurons (s_i > 0.5) in IT for each experimental condition. The Wilcoxon Signed-Rank test was used to compare the experimental groups ("stripes", "horblob", "verblob") to the control group ("all3"). The number of IT neuronal units responding to a deprived stimulus decreased significantly for each of the experimental groups ($p < 0.05$), whereas units responding to other (over-sampled) stimuli did not change significantly, indicating that IT had become "saturated" with respect to these well-trained stimuli. Because of the dynamics of the rules governing synaptic change and the widespread connectivity in IT, the saturation shown in Figure 6 is plastic and reversible over longer time periods.

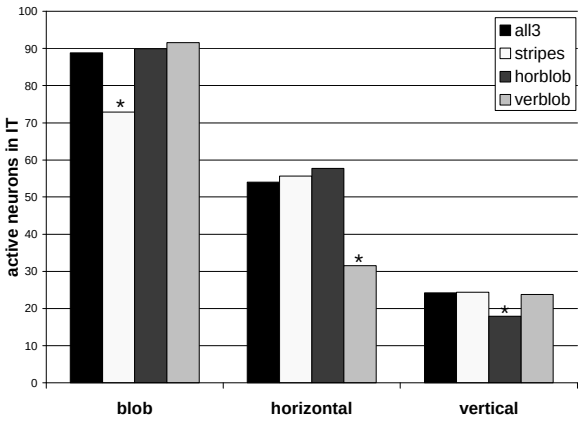

Figure 6: The effect of stimulus frequency after visual development. The figure shows the median number of active neurons in IT after presenting either all three stimulus patterns ("all3"), only *horizontal* and *vertical* ("stripes"), only *horizontal* and *blob* ("horblob"), or *verticals* and *blobs* ("verblob") to a trained Darwin VI (N=10). An asterisk denotes a significant difference between the control group ("all3") and an experimental group ($p < 0.02$ for *blobs* in "stripes"; $p < 0.02$ for *horizontals* in "verblob"; and $p < 0.05$ for *verticals* in "horblob") using the Wilcoxon Rank Sum test of medians. Note that the representation for a deprived stimulus decreased, whereas the over-sampled stimulus remained unchanged.

4. Discussion

In this paper, we studied the role of sensory experience in the development of neural activity underlying perceptual categorization. We employed a neuronal simulation on a real-world behaving device to test the effects of environmental interactions on a developing nervous system. Despite the fact that every experimental run began with an identical device that had an identical nervous system, examination of the simulated nervous system after the experiment showed differences that could only be attributed to experience. Allowing the device to explore an ecologically varied environment or varying the presentation of stimuli brought about these differences.

In particular, we found that both overall frequency and temporal order of stimuli had an impact on the number of neuronal units devoted to specific object classes. Early exposure to examples of a particular stimulus class resulted in rapid development of neuronal units responsive to that class, and produced a suppression of the subsequent development of units devoted to other, less frequently experienced classes. The number of corresponding neuronal units increased significantly with the overall frequency of stimulus encounters (Figure 4). In contrast, after visual development had taken place, the number of neuronal units did not increase further when over-exposed to a stimulus class, while neuronal units that were selective for stimuli that were absent from the experiential history showed a decline (Figure 6). All these effects suggest specific neural mechanisms accompanying perceptual categorization in biological organisms and are in agreement with physiological data (Kobatake et al., 1998).

At the beginning of each experimental run, before any visual experience has taken place, area IT contains neuronal units that are responsive, but non-selective with respect to objects of different classes. Due to the widespread connectivity between the primary visual area and area IT, units in IT receive afferent connections from large portions of the feature selective sub-regions in the primary visual area. Early exposure to exemplars of a particular object class results in synaptic modifications rendering several local groups of units in IT selective for this class. These neuronal groups occupy some portion of the neural map in IT and exert lateral inhibition on other units surrounding them. This cross-inhibitory action puts other units in IT at a competitive disadvantage as other stimuli are encountered. Figure 5 shows representative examples of IT neurons capturing neural space when a specific object class was sampled early and often. This capturing of neuronal space prevented neurons from being devoted to other object classes sampled later. The maps that arise from this competition are dynamic and modify over time in response to the temporal order and frequency of object presentation.

While prolonged exposure of IT to examples of the same stimulus classes leads to saturation in the magnitude and extent of the IT response, the IT response is diminished if no exemplars of the corresponding object class are encountered. This is due to the gradual decay of synaptic weights back to their starting values, which acts as a synaptic mechanism of "forgetting", gradually erasing the effects of learning. This mechanism should be distinguished from "anti-Hebbian" learning (decreasing synaptic weights

if presynaptic and postsynaptic activity are uncorrelated). We chose to implement a mechanism that applied uniformly to all modifiable weights, irrespective of neural activity. Clearly, this represents a very simple choice, but it may suffice to implement the notion that synaptic changes are long lasting, although not permanent. This is a property found in many neural systems exhibiting long term synaptic plasticity. The real biological mechanisms (structural or biophysical) that govern the gradual decay of long-term traces or their selective consolidation are still relatively unknown (for a recent review see McGaugh, 2000). In Darwin VI, this simple mechanism of passive synaptic decay serves to degrade neuronal units that are responsive to previously learned stimuli maintaining the capacity to respond to novel object classes.

In agreement with previous work (Almassy et al., 1998; Sporns et al., 2000), the above results would not be possible without the implementation of a real-world device behaving in its environment and responding to a simulated nervous system. The device never experiences a stimulus in exactly the same way. Light differences, angle of approach or environmental noise cause experiential variations that lead to robust learning. Darwin VI's robust learning and resulting individual differences between experiments could not be achieved by constructing a virtual environment. Our experiments reinforce the importance of testing nervous system theories on embodied platforms (Reeke and Sporns, 1993; Chiel and Beer, 1997; Clancey, 1997; Clark, 1997).

Overall, our studies with Darwin VI also point to the importance of the individual history of stimulus encounters as well as ecological aspects of stimulus composition in the environment in perceptual categorization. Such categorization necessarily involves sensorimotor interactions and self-generated behavior. Further studies of these processes may aid in understanding mechanisms of biological categorization and help towards the construction of more varied and robust artificial categorization architectures.

Acknowledgements. This work was supported by the Neurosciences Research Foundation and the W.M. Keck Foundation.

References

Almassy, N., Edelman, G.M., and Sporns, O. (1998) Behavioral constraints in the development of neuronal properties: A cortical model embedded in a real-world device. *Cerebral Cortex* **8**:346-361.

Almassy, N., and Sporns, O. (1998) Role of behavior in the development of complex neuronal properties. *Proceedings SAB98*, pp. 312-320, MIT Press, Cambridge, MA.

Barsalou, L.W. (1992) *Cognitive Psychology*. Lawrence Erlbaum, Hilldsdale, NJ.

Buonomano D., and Merzenich, M. (1998) Cortical plasticity: From synapses to maps. *Annu. Rev. Neurosci*. 21:149-186.

Chiel, H.J., and Beer, R.D. (1997) The brain has a body: Adaptive behavior emerges from interactions of nervous system, body and environment. Trend Neurosci. 20:553-557.

Clark, A (1997) Being there. Putting brain, body, and world together again. MIT Press, Cambridge, MA.

Clancey, W.J. (1997) *Situated Cognition*. Cambridge University Press, Cambridge.

Diamond, M. E., Armstrong-James, M., and Ebner, F.F. (1993) Experience-dependent plasticity in adult rat barrel cortex. Proc. Natl. Acad. Sci. USA 90, 2082-2086.

Edelman (1987) *Neural Darwinism*. Basic Books, New York, NY.

Edelman, G.M., Reeke, G.N., Gall, W.E., Tononi, G., Williams, D., and Sporns, O. (1992) Synthetic neural modeling applied to a real-world artifact. *Proc. Natl. Acad. Sci. USA* 89:7267-7271.

Elbert, T., Pantev, C., Wienbruch, C., Rockstroh, B., and Taub, E. (1995) Increased cortical representation of the fingers of the left hand in string players. Science 270:305-307.

Eimas, P.D. (1994) Categorization in early infancy and the continuity of development. *Cognition* 50:83-93.

Garraghty, P.E., and Kaas, J.H. (1992) Dynamic features of sensory and motor maps. *Curr. Opin. Neurobiol.* 2:522-527.

Johnson-Laird, P.N. (1988) *The Computer and the Mind.* Harvard University Press, Cambridge, MA.

Kobatake, E., Wang, G., and Tanaka, K. (1998) Effects of shape-discrimination training on the selectivity of inferotemporal cells in adult monkeys. *J. Neurophysiol.* **80**:324-330.

Markman, E.M. (1989) *Categorization and Naming in Children.* MIT Press, Cambridge, MA.

McGaugh, J.L. (2000) Memory – A century of consolidation. *Science* 287:248-249.

Merzenich, M., and deCharms, C. (1996) Neural representations, experience, and change. In *The Mind Brain Continuum*, ed. R. Llinas, P.S. Churchland, pp. 61-81. MIT Press, Cambridge, MA.

Mervis, C.B., and Rosch, E. (1981) Categorization of natural objects. *Annu. Rev. Psychol.* 32:89-115.

Nudo, R., Milliken, G., Jenkins, W., and Merzenich, M. (1996) Use-dependent alterations of movement representations in primary motor cortex of adult squirrel monkeys. *J. Neurosci*. **16**:785-807.

Pfeifer, R., and Scheier, C. (1997) Sensory-motor coordination: The metaphor and beyond. *Robotics Auton. Syst.* **20**:157-178.

Pfeifer, R., and Scheier, C. (1999) *Understanding Intelligence*. MIT Press, Cambridge, MA.

Reeke, G.N., and Sporns, O. (1993) Behaviorally based modeling and computational approaches to neuroscience. *Ann. Rev. Neurosci*. 16:597-623.

Rodman, H. (1994) Development of inferior temporal cortex in the monkey. *Cerebr. Cortex* 5:484-498.

Ryle, G. (1949) *The Concept of Mind*. Hutcheson, London.

Sakai, K., and Miyashita, Y. (1991) Neural organization for the long-term memory of paired associates. *Nature* 354:152-155.

Schyns, P.G., Goldstone, R. L., and Thibaut, J.-P. (1998) The development of features in object concepts. *Behav. Brain Sci.* 21:1-54.

Scheier, C., and Lambrinos, D. (1996) Categorization in a real-world agent using haptic exploration and active perception. *Proc. SAB96*:65-74.

Smith, L.B., Thelen, E., Titzer, R., and McLin, D. (1999) Knowing in the context of acting: The task dynamics of the A-not-B error. *Psychol. Rev*. 106:235-260.

Sporns, O., Almassy, N., and Edelman, G.M. (2000) Plasticity in value systems and its role in adaptive behavior. *Adaptive Behavior* (in press).

Thelen, E. , and Smith, L.B. (1994) A Dynamic Systems Approach to the Development of Cognition and Action. MIT Press, Cambridge, MA.