Multiple Cognitive Abilities from a Single Cortical Algorithm

Suzanna E. Forwood[1], Rosemary A. Cowell[2], Timothy J. Bussey[1], and Lisa M. Saksida[1]

Abstract

■ One strong claim made by the representational–hierarchical account of cortical function in the ventral visual stream (VVS) is that the VVS is a functional continuum: The basic computations carried out in service of a given cognitive function, such as recognition memory or visual discrimination, might be the same at all points along the VVS. Here, we use a single-layer computational model with a fixed learning mechanism and set of parameters to simulate a variety of cognitive phenomena from different parts of the functional continuum of the VVS: recognition memory, categorization of perceptually related stimuli, perceptual learning of highly similar stimuli, and development of retinotopy and orientation selectivity. The simulation results indicate—consistent with the representational–hierarchical view—that the simple existence of different levels of representational complexity in different parts of the VVS is sufficient to drive the emergence of distinct regions that appear to be specialized for solving a particular task, when a common neurocomputational learning algorithm is assumed across all regions. Thus, our data suggest that it is not necessary to invoke computational differences to understand how different cortical regions can appear to be specialized for what are considered to be very different psychological functions. ■

INTRODUCTION

The architecture and computational function of the ventral visual cortex are better understood than those of perhaps any other region of mammalian cortex. After decades of research, many properties of this brain region are well elucidated and generally agreed upon: retinotopy of visual representations in early, posterior regions that disappears in anterior regions (Tootell, Dale, Sereno, & Malach, 1996; Tootell, Switkes, Silverman, & Hamilton, 1988; Hubel & Wiesel, 1962); columnar organization of early visual representations for features such as orientation (Blasdel & Salama, 1986; Hubel & Wiesel, 1959); increasing receptive field size with anterior progression; the emergence of view and position invariant object representations in anterior areas (Tanaka, 2003; Rolls, 1992); a general, hierarchical scheme for visual representations in which simple visual "elements" are coded for in posterior areas and more complex features or whole objects are represented in anterior regions (Bussey & Saksida, 2002; Riesenhuber & Poggio, 1999; Tanaka, Saito, Fukada, & Moriya, 1991; Desimone, Albright, Gross, & Bruce, 1984); the list goes on. However, the issue of how best to characterize the cognitive function of this region remains highly controversial.

The debate over the cognitive contributions of the ventral visual stream (VVS) can be described, in broad terms, as a debate about specialization of function. One strand of the debate concerns category-selective specialization: do there exist regions of ventral visual cortex that are specialized for the processing of certain object categories, such as faces and houses (Op de Beeck, Haushofer, & Kanwisher, 2008; Tsao & Livingstone, 2008; O'Toole, Jiang, Abdi, & Haxby, 2005; Hanson, Matsuka, & Haxby, 2004; Spiridon & Kanwisher, 2002; Kanwisher, McDermott, & Chun, 1997), or are regions instead specialized for domain-general skills such as expertise rather than for object categories (Gauthier & Tarr, 1997), or is the neural code for objects in fact distributed (Haxby et al., 2001)? A second strand of the debate concerns the functions of visual perception and visual memory. The standard view suggests that visual perception and memory are localized to distinct regions within VVS and antero-medial-temporal lobe, with perception a function of posterior areas and memory of anterior areas (Squire & Wixted, 2011; Knowlton & Squire, 1993; Sakai & Miyashita, 1993; Squire & Zola-Morgan, 1991; Mishkin, 1982). An alternative account claims that a given region may contribute to both perception and memory (Cowell, Bussey, & Saksida, 2006, 2010; Lopez-Aranda et al., 2009; Barense et al., 2005; Lee, Barense, & Graham, 2005; Bussey, Saksida, & Murray, 2002, 2003; Buckley & Gaffan, 1998)—indeed, that perceptual and mnemonic tasks may sometimes tap the same neural representations—and that the functional contribution of each brain region is determined not by its location within a

[1]University of Cambridge, [2]University of California, San Diego

cognitive module specialized for a certain function, but by the nature of the stimulus representations it contains (Cowell et al., 2010; Tyler et al., 2004; Bussey & Saksida, 2002).

We have advocated an account of object processing that falls into the second of the above camps, arguing in favor of distributed object representations (Cowell, Huber, & Cottrell, 2009; Cowell et al., 2006) and a functional continuum along VVS, in which all processing stages may contribute to perception or memory (or indeed, any object processing function) depending on the representational requirements of the task (Cowell et al., 2010; Bussey & Saksida, 2002). The assumption of a continuous hierarchy of object representations along VVS is central to this explanation of visual cognition, so we have termed this account the "representational–hierarchical view." The existence of a hierarchy of object representations in VVS is widely accepted, forming the basis of many models of object processing (e.g., Riesenhuber & Poggio, 1999; Wallis & Rolls, 1997; Perrett & Oram, 1993). These have successfully used hierarchy to model visual identification of shapes and objects independently of stimulus variability, location, size (Riesenhuber & Poggio, 1999; Fukushima, 1980), 3-D viewing angle (Wallis & Rolls, 1997), and within a cluttered field (Grossberg, 1994). Just as in these models, the representational–hierarchical view of VVS function assumes that simple features reside in posterior regions of VVS, and complex conjunctions of those simple features are housed in more anterior regions. Stimulus representations are hypothesized to reach a maximum of complexity in perirhinal cortex (PRC)—a brain structure situated at the anterior end of the VVS that is known to be critical for judging the familiarity or novelty of objects (Squire, Wixted, & Clark, 2007; Murray, Graham, & Gaffan, 2005; Winters, Forwood, Cowell, Saksida, & Bussey, 2004), as well as for object perception under certain circumstances (Bussey, Saksida, & Murray, 2002; Murray, Bussey, & Saksida, 2001; Buckley & Gaffan, 1998). The complexity of stimulus representations reached in PRC is assumed to correspond to the level of a whole object and confers the functional role of PRC in both object memory and object perception; any task requiring such object-level representations—regardless of the specific cognitive function that it is tapping into—will be affected by damage to PRC. Similarly, any posterior region within VVS may contribute to both perception and memory of visual stimuli according to the level of complexity of the stimulus representations that the region contains: if the task, whether "mnemonic" or "perceptual," is best solved on the basis of simple visual features, then the stage of VVS that will be optimal for its solution is the stage containing simple feature representations (Cowell et al., 2010).

The representational–hierarchical view entails several claims about cortical function in VVS, some of which remain untested, computationally. For example, Bussey and Saksida (2002, 2005) have suggested that if the VVS is truly a functional continuum, the computations carried out in the service of a given cognitive function (say, visual recognition memory or visual discrimination) might be the same at all points along VVS, including PRC. In this case, differences in the contributions to cognition made by each region would simply be due to differences in the stimulus representations contained in each region. Posterior VVS might provide a familiarity signal allowing recognition of simple visual features in the same way that PRC provides a familiarity signal for whole objects. Related to this is the claim by Cowell et al. (2010) that the representational requirements of a task determine which brain region is most critical for the task solution. For example, if a visual discrimination task uses objects but those objects are discriminable on the basis of a simple feature, such as a color, then the task can be solved either using object-level representations or feature-level representations. On the other hand, if a visual discrimination task involves presentation of the same object from different views and requires apprehension that those different views arise from the same object (Lee, Scahill, & Graham, 2008; Buckley, Booth, Rolls, & Gaffan, 2001), then object-level representations will likely be required (because probably no single feature can be used to determine the correspondence of object identity across the different views).

In the present article, we test the viability of these claims. Is it possible that different regions could produce the semblance of distinct functions using the same computational algorithms operating upon different stimulus representations (e.g., perceptual expertise with line orientations as opposed to faces)? Is it true that the representational requirements of a task can determine the relative abilities of different brain regions to solve the task, when all that differs between those regions is the stimulus representations they contain rather than the computations they perform? This latter question has already been tested within the specific domain of visual discrimination learning (Cowell et al., 2010). This study simulates one of the many empirical studies showing a double dissociation of function with the VVS (Iwai & Mishkin, 1968) and interpreted as evidence for distinct functional modules within VVS—anterior structures being for "perception" and posterior structures being for "associative memory." The model it uses to demonstrate this functional dissociation is computationally identical in each of its hierarchical layers, with the sole difference being the complexity of the nature of the stimulus representations. The present article seeks to extend this finding further by establishing whether tasks as diverse as object recognition memory and categorization can be explained in terms of a common cortical learning algorithm responding to differing representational requirements.

Our principal aim is to put all of the cognitive tasks we examine onto a level playing field, computationally, and see whether differences in the input stimuli and representational requirements of different tasks can produce the divergent behaviors associated with those tasks. In previous

Figure 1. Diagram illustrating the model architecture and plasticity. Left: Drawing of the major circuitry, showing the input from the input layer, the lateral connections to one example network unit within the layer, and the output from the layer to a response unit. Right: Learning rules used, illustrating the conditions needed for weight change. See Appendix 1 for a more detailed description of the equations and the definitions of the variables used.

$$w_{ij}(t+1) = w_{ij}(t) + a''_j \cdot \lambda \cdot (a_i - w_{ij}(t))$$

$$w_{jr}(t+1) = w_{jr}(t) + a_j \cdot \alpha \cdot (R - a_r)$$
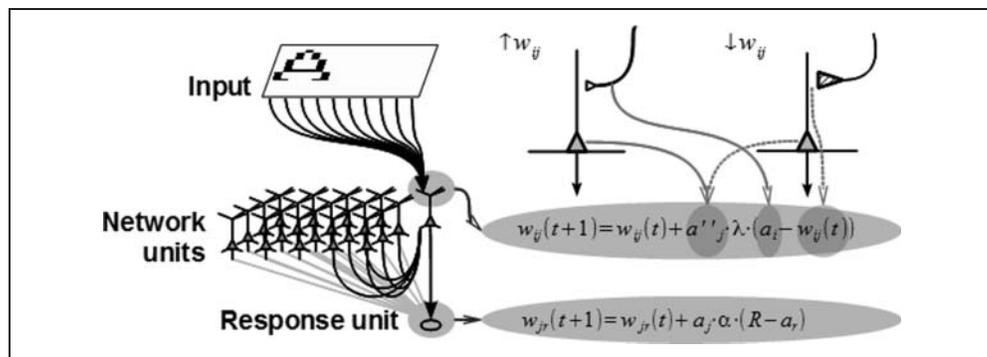
instantiations of the representational–hierarchical view (Cowell et al., 2006, 2010; Bussey & Saksida, 2002), we assumed a hierarchical structure with multiple layers of stimulus representations, in which later layers contained more complex stimulus representations than earlier layers. In the current computational study, we replace that hierarchy with a single layer, using the same learning mechanism and parameters on that layer for all tasks (Figures 1 and 2).

The specific question under investigation in the current article is whether it is possible for different regions to use the same computational algorithm upon different representations to generate a semblance of the distinct computational functions seen in VVS. For each task, the single network layer receives input stimuli at the level of complexity used in the real-world version of the task (e.g., lowercase and uppercase letters for stimulus recognition versus simple lines for the development of orientation selectivity). Each task was run in a separately initiated model, as we are not exploring issues of interference between tasks within the same layer. The use of a single layer, a departure from our previous models (Cowell et al., 2006, 2010), is critical to our present aim; employing a hierarchy with several layers of stimulus representations at different levels of structural complexity and using different layers for different tasks would not be a true test of the hypothesis that a single neurocomputational algorithm operating on different stimulus inputs can produce divergent cognitive functions. By not using a hierarchical model, but instead using a single layer and simply varying the stimulus input in a task-appropriate fashion, we are able to test whether one unifying algorithm can account for the emergence of representations at the appropriate level of complexity for the task. The logical extension of this to the brain, of course, is a series of such layers stacked together, similar to previous models of the VVS (e.g., Riesenhuber & Poggio, 1999; Wallis & Rolls, 1997; Grossberg, 1994; Perrett & Oram, 1993; Fukushima, 1980). However, the current work adds to the existing
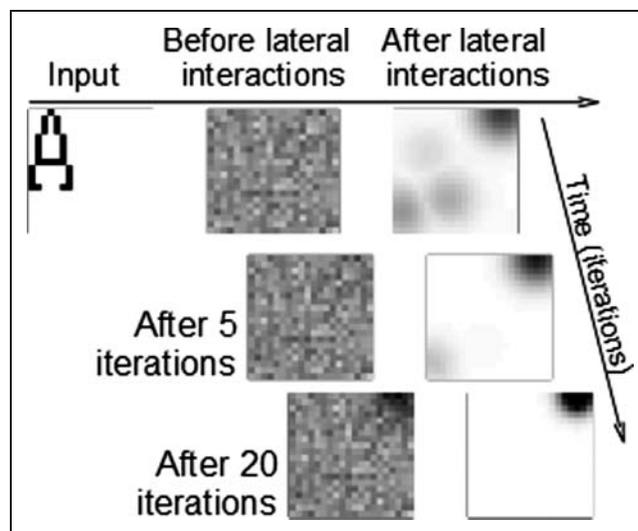
Figure 2. Diagram illustrating the model's response to repeated exposure to one stimulus. In each image, a single unit is represented by a pixel, with the activity of the unit represented by the darkness of the pixel: black = 1 and white = 0. Top row: A stimulus is presented to the model as a specific pattern of activity (left). In the first instance the model is naive, so all the weights between the input units and the network units are at random values producing a random and noisy pattern of activity (middle). After lateral interactions between all the units in the layer, some clusters of units with relatively high propagation layer activity appear to retain local islands of activity. For the other units, the lateral inhibition has reduced their activity to 0, giving the resultant pattern of islands of activity in a sea of inactivity (right). Middle and bottom rows: The pattern of activity after lateral interactions determines which units are able to engage in learning. Those units in the network layer that had moderate activity after one iteration can sustain moderate amounts of learning, so when the same stimulus is presented again, those units are able to generate a stronger activity as a result of their updated weights. The units outside the island of activity will still have noisy and weak activity values before lateral interactions, and are reduced by the lateral interactions, unless other and different stimuli have also been presented that these units learn to represent. The strong activity in the peak units will also serve to inhibit the activity in the other island units, so both "cleaning-up" and strengthening the representation. Thus, with repeated exposure, the peak units show pronounced activity in response to the stimulus and come to signal the presence of that stimulus.

literature by isolating one feature of hierarchy—increasing stimulus complexity—and assessing its contribution to the distinct behavioral functions found within the hierarchy of the VVS.

We test the model presented here on its ability to simulate results from a variety of tasks associated with different regions of the VVS, ranging from tasks typically thought of as tapping high-level cognition in anterior VVS to tasks that are associated with low-level vision and development in posterior VVS and primary visual cortex. To provide the most stringent test of our ideas possible, tasks were chosen to represents the broadest range possible of computational occurrences known to be dependent on structures within the VVS. These include (1) recognition memory (cf. Cowell et al., 2006) associated with PRC (Squire et al., 2007; Murray et al., 2005; Winters et al., 2004), (2) categorization of perceptually related stimuli (Posner & Keele, 1968) associated with inferior temporal cortex (Keri, 2003), (3) perceptual learning for discrimination of highly similar visual stimuli (cf. Saksida, 1999) associated with extrastriate cortex (Gilbert, Sigman, & Crist, 2001), (4) the development of retinotopy (Tootell et al., 1988), and (5) the development of orientation selective representations (Bartfeld & Grinvald, 1992; Blasdel & Salama, 1986) both associated with primary visual cortex. The network is able successfully to simulate data across this broad range of tasks, suggesting that the basic computational mechanisms that underlie "low-level" perceptual functions such as development of primary visual cortex and "high-level" cognitive functions such as categorization or recognition memory may be more similar than is usually assumed. Consequently, to explain observed differences in the contributions of different regions within VVS to cognition, it may not be necessary to invoke notions of functional specialization; a more parsimonious account may be offered by assuming shared processing mechanisms operating upon different representational content.

METHODS

Model Overview

The algorithm we use in the present simulations is based on the Kohonen self-organizing feature map (SOFM; Kohonen, 1984): a single-layer model, with no hierarchy or feedback from downstream structures, that is able learn without the need for a "teaching signal." However, because Kohonen's (1984) original SOFM algorithm does not provide a representation of a single-neuron activity—which makes it difficult to model electrophysiological data—we use a model based closely on the SOFM but with a variety of neurally plausible properties. For example, activity calculation for each unit is not based on Cartesian distance but instead uses the product of input weight with input activity and more conventional associa-

tive treatment of these values (Rescorla & Wagner, 1972). In addition, lateral interactions between units are explicitly calculated, unlike in the conventional Kohonen network where they are imposed and a learning rule based on N-methyl-D-aspartate-mediated LTP and LTD is used. These three details have very little impact on the mechanics of the model, meaning that the model that we use here produces much of the same high-level response to stimuli as would be seen in a conventional Kohonen network (Kohonen, 1984) but, at the same time, contains lower-level representations that allow us to model electrophysiological data. A more radical change from the Kohonen network is that the neighborhood size does not decrease as a function of time but rather changes on a trial-by-trial basis as a function of the unit activity response to the current stimulus, in line with recent electrophysiological data (Angelucci et al., 2002). This change has a particular impact on tasks where stimulus familiarity will change on a trial-by-trial basis, such as recognition memory, and is discussed further in the Results to Experiment 1.

For a minority of the tasks we simulate, associative learning is needed to associate the pattern of activity produced by a self-organizing array of units in response to a given stimulus with an outcome. Previous work has already demonstrated very successfully that simple error correction learning algorithms (Rescorla & Wagner, 1972; Widrow & Hoff, 1960) can learn effectively the associations between a single stimulus and an outcome. This is the case even if stimuli are represented using distributed patterns of activity (Ghirlanda, 2005), and if these distributed patterns change with exposure to stimuli, at the same time that error correction learning is taking place (Saksida, 1999). Critically, this supervised learning does not have any feedback connectivity to the unsupervised self-organizing array of units, and therefore, the additional information provided by the outcome cannot affect learning within the self-organizing array. Thus, we adopt this simple error correction learning algorithm when associative learning is required to solve the task.

The single-layer network architecture is very simple, which lends the model not only parsimony but also clarity. By reducing the number of built-in assumptions, the key aspects of the mechanism in the model that are responsible for the observed simulation results are revealed. Further detailed properties of visual cortex that are known to exist but whose inclusion in the model would obscure the simple mechanism that can account for the simulated findings—such as hierarchical layers of representations or cortical feedback—are purposefully excluded. For full details of the model, see Appendix (also see Figure 1).

Stimuli

Many existing models of cognitive function (e.g., Cowell et al., 2006, 2010; Bussey & Saksida, 2002) approximate

real-world stimuli by representing them with a small array of units, where each unit represents a stimulus dimension such as width, length, or color, and each unit value represents the value for that stimulus in that dimension. However, in the present work, one of our aims was to test whether the stimulus properties hypothesized (i.e., assumed) in existing instantiations of the representational–hierarchical view are indeed possessed by the kinds of stimuli used in the empirical tasks that we have simulated and whether differences in stimuli across different tasks are sufficient to account for the emergence of diverse cognitive functions (such as object recognition memory and perceptual learning). Therefore, in the current model, we use realistic two-dimensional images of visual stimuli—gray-scale representations of lines, shapes, and objects within a 20 × 20 pixel input space—which are, where possible, identical to the stimuli used to collect the original behavioral data.

EXPERIMENT 1: SIMULATION OF STIMULUS RECOGNITION

The study of recognition memory, the ability to judge whether a stimulus has been seen before, has been central to our understanding of memory and amnesia as it is thought to be an example of declarative memory (Squire & Zola-Morgan, 1991), the explicit recall of past events. The critical role of medial-temporal lobe structures in recognition memory was highlighted by the study of temporal lobectomy patients (Scoville & Milner, 1957), and it is now widely acknowledged that neocortical structures, such as the PRC, are essential (Squire et al., 2007; Murray et al., 2005; Winters et al., 2004).

Here, we simulate a preferred looking task that is widely used to assess recognition memory in humans (visual paired comparison; Manns, Stark, & Squire, 2000) and rodents (pontaneous object recognition; Ennaceur & Delacour, 1988). Participants are allowed to study an object and then, after a delay, are shown the studied object along with a new object. Preference for the novel object, an indicator of memory for the familiar object, declines as a function of delay (Forwood, Winters, & Bussey, 2005) and is sensitive to damage to PRC (Winters et al., 2004; Bussey, Muir, & Aggleton, 1999). A self-organizing mechanism (Kohonen, 1984) combined with sharpening of stimulus representations proportional to length of exposure to give measures of stimulus novelty has been used by computational models of recognition memory in PRC (Cowell et al., 2006; Bogacz & Brown, 2003; Norman & O'Reilly, 2003).

The current model was run with the default parameters as set out in Table 1. The stimuli and training procedure used in this experiment are detailed in Figure 3. Fifteen simulations were run to replicate multiple subjects, with five being run on each of three delay conditions. Each network corresponds to a single subject in the standard rat spontaneous object recognition task protocol receiving six recognition memory trials in succession. Each recognition memory trial involved exposure to a novel stimulus for a set number of iterations in the sample phase, followed by a delay period. In the choice phase, the now familiar sample stimulus was available alongside a novel stimulus. Between each recognition memory trial and during the delay period within the recognition memory trial, the models were exposed to a fixed set of 14 stimuli to represent neutral familiar stimulus exposure, such as a rat's home cage. For each test session, the learning rate, λ, was set to 0.05 to reduce the amount of learning taking place per iteration so that more gradual changes in the stimulus representation were detectable.

During the choice phase, to simulate stimulus preference, the model was programmed to "switch-if-familiar": for each stimulus presentation, the peak of activity in the stimulus representation on the network layer (a value between 0 and 1) was compared with a randomly generated number between 0 and 1, and if greater, the model would switch to view the other stimulus available on the next trial. The number of times the network viewed the novel stimulus, N_{novel}, and familiar stimulus, $N_{familiar}$, was used

Table 1. The Default Values for the Parameters Used in the SONN Model

Parameter	Symbol in Equations	Value Used
Minimum neighborhood size	N_{min}	2
Maximum neighborhood size	N_{max}	12
Network layer learning rate	λ	0.1
Rw learning rate	α	0.1
Input layer size		array of 20 × 20 units
Network layer size		array of 20 × 20 units[a]

Unless otherwise stated, the given values are used in all simulations.

[a]For Experiment 1, 40 by 40 array was used reflecting the greater surface area in the cortex devoted to primary sensory structures relative to higher structures.

Figure 3. Simulation of recognition memory. Procedure shows the stimuli presented to the network in one of the six repeated sequences that each network was exposed to with new LC letters each time. **In the Choice phase, the simulation is run using a "Switch-if-Familiar" protocol, enabling the model itself to assess familiarity for the current stimulus, and if familiar to switch to the alternative stimulus for the next trial. See Experiment 1 methods for details. Results shows the calculation of a discrimination ratio using the number of trials of the novel and familiar stimuli in the choice phase and average discrimination ratios on the recognition memory task, with five networks tested per delay. A discrimination ratio of 0 indicates no preference for the novel stimulus, a positive score indicates preference for the novel object. Error bars indicate *SEM*.
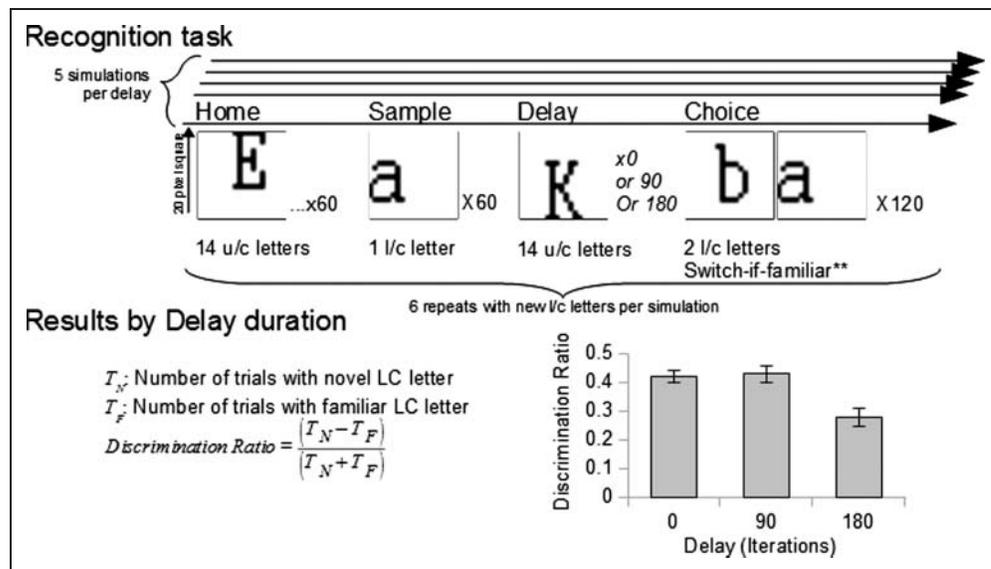
to calculate a discrimination ratio (Winters et al., 2004), with a positive score indicating novelty preference:

$$\text{Discrimination Ratio} = \frac{(N_{\text{Novel}} - N_{\text{Familiar}})}{N_{\text{Total}}}$$

ANOVA was run with Delay as a between-subject factor on the discrimination ratios. If the main effect was significant, the discrimination ratio for each delay was compared with the others using multiple comparisons of means to determine which delays were significantly different from performance at zero delay.

Results and Discussion

ANOVA of the discrimination ratios produced by the model data (Figure 3) showed a significant main effect of Delay ($F(2, 12) = 9.58$, $p < .01$) and a significant intercept ($F(1, 12) = 628.40$, $p < .001$), indicating that the discrimination ratios were significantly different from zero—rates of novel and familiar stimulus exploration were not equal. Multiple comparisons of means (with Tukey adjustment) revealed no significant difference between performance at 0 and 90 iterations delay ($p > .05$), but a significant difference between performance at 0 and 180 iterations delay ($t(4) = 3.59$, $p < .01$), confirming the visible trend in the data that the delay reduced novelty preference in the choice phase. This successful simulation of recog-

nition memory is a product of one basic feature of the model: Familiar stimuli evoke stronger activity patterns than novel stimuli (Figure 2). Unlike in a conventional Kohonen network (Kohonen, 1982) where the neighborhood size and learning rate systematically fall as training progresses, the current model uses a fixed learning rate and a neighborhood size that is driven by the peak unit response to the current stimulus (Angelucci et al., 2002; see Appendix, point 3). Therefore familiar stimuli are capable of evoking stronger single unit activity because prior training has altered the weights of a subset of units to enable them better to represent the stimulus. In turn, this results in a minimal neighborhood size, such that the final activity pattern is spatially limited. By using the peak strength of responding to the current stimulus as a cue to switch exploration to the alternative, a pattern of performance that gradually decays with increasing delay is shown, as seen in animals (Forwood et al., 2005; Eacott, Gaffan, & Murray, 1994; Zola-Morgan, Squire, Amaral, & Suzuki, 1989) and humans (Holdstock, Gutnikov, Gaffan, & Mayes, 2000; Buffalo, Reber, & Squire, 1998). As with healthy animal subjects, this novelty preference is affected by delay periods where intervening stimuli are presented: Intervening stimuli modify the weights that are well tuned to the familiar stimulus, so reducing the strength of the activity pattern to that familiar stimulus after the delay and thus reducing the preference for the novel object (see also Bartko, Cowell, Winters, Bussey, & Saksida, 2010; Cowell et al., 2006).

One prominent hypothesis about how the brain detects novelty in stimuli is that it is related to decreased responding of neurons to repeated stimuli, referred to as "response decrement on stimulus repetition" (Fahy, Riches, & Brown, 1993) or "repetition suppression" (Miller, Li, & Desimone, 1991). Quantification of unit activity in the six sample phases for each of the five simulations for this experiment shows that repetition-induced response decrements are taking place in the majority of units that respond to a stimulus. When presented with a novel stimulus at the beginning of the sample phase, only a small fraction (13.6%) of units respond to the stimulus above a level of 0.01, the rest show minimal or no activity. Of these, when the same stimulus is presented after a further 40 presentations, 8% of units show no change greater than ±10% of their original activity level, 80.1% show a decrease of greater than 10% and only 11.9% show an increase of greater than 10%. Thus, most units that respond selectively to a stimulus show response reduction after repeated exposure to that stimulus, consistent with the data on repetition-sensitive responding (Zhu & Brown, 1995; Fahy et al., 1993; Miller et al., 1991). A minority of units show enhancement, consistent with reports of response increments alongside decrements in the temporal lobe (Table 2 in Zhu & Brown, 1995). The single unit with peak activity to the stimulus, which determines the probability of switching from one stimulus to the other in the model, always increases its activity, by an average of 86%. However, the likelihood of finding the corresponding unit during electrophysiological experiments that tend to sample a few hundred cells in a structure that contains hundreds of thousands of cells is clearly small. Thus, the current model, although it uses increased responding and enhanced specificity of responding to a stimulus as an indicator of stimulus familiarity, shows activity patterns that are consistent with electrophysiological data. It also accounts for the fact that, although the activity of most neurons should be reduced by repeated exposure to a stimulus, a much smaller number of neurons will show enhanced activity (Zhu & Brown, 1995).

Another study has attempted to assess whether the PRC is functionally organized by asking whether neurons that respond to similar stimulus attributes cluster together (Erickson, Jagadeesh, & Desimone, 2000). Data addressing this issue experimentally have been obtained by looking at correlations in the firing patterns of single neurons in PRC during the presentation of either novel or familiar visual stimuli (Erickson et al., 2000). When comparing two neurons that were "near" (recorded with the same electrode at the same cortical location and at different cortical depths), it was found that a positive correlation was observed, but there was a greater correlation when viewing familiar stimuli (0.28) than when viewing novel stimuli (0.13). This analysis can be replicated using data from the current model by looking at the correlation between the activity of the network units at different stages in the lateral interaction calculations: before (a_j)

and after (a_j''). The current model assumes, for simplicity, that these calculations take place within the same unit, although this is not necessarily the case in the neocortex, where different populations of adjacent cells may play different roles in these calculations or be undergoing different stages of the calculation at the same point in time. The "experienced" simulations presented above were therefore exposed to a collection of 14 familiar stimuli (the 14 neutral or home stimuli) randomly interleaved with 12 novel stimuli (each consisting of a pair of parallel bars differing in orientation and location within the input array) for 10 iterations following the above simulations, and the activity values before and after lateral interaction calculations were collected. These activity values were separated into trials of novel and familiar stimuli, and the correlation coefficient between the average activity value before and after lateral interaction calculations for each unit were calculated for each stimulus type. The model data follow the pattern observed in primates (Erickson et al., 2000), with a greater correlation (r value) when viewing familiar stimuli (0.47 ± 0.013 SEM) than when viewing novel stimuli (0.25 ± 0.013 SEM). As well as providing evidence that the current model ties in with the known attributes of recognition memory in the PRC, this analysis demonstrates that this neuro-realistic model can be tested at a neuro-biological level as well as at a cognitive level.

EXPERIMENT 2: SIMULATION OF STIMULUS CATEGORIZATION

A cognitive function typically associated with the VVS and with temporal lobe cortex, more generally, is the categorization of stimuli based on perceptual dimensions (Keri, 2003), such as the dot pattern classification task (Posner, Goldsmith, & Welton, 1967). In this task, participants learn to sort abstract patterns of dots either using their own subjective criteria or with the help of feedback. A category of dot patterns is created by first creating a random pattern of dots—the prototype—and from this, any number of exemplars can be created by moving each of the dots in the prototype to a greater or lesser extent (Posner et al., 1967). The size of the distortion used can be low or high, making exemplars with varying similarity to the prototype and to each other. After training with several exemplars from three categories, the subject is required to label some of those familiar exemplars, as well as some novel exemplars and the prototype itself, neither of which they have seen before. One major finding from these experiments is that subjects are more accurate at labeling the prototype than the other equally novel exemplars: the prototype effect (Posner & Keele, 1968). The second major finding is that novel exemplars are sorted less accurately than the familiar exemplars regardless of their similarity to the prototype: the exemplar effect (Posner & Keele, 1968).

These major findings inspired the two main theories regarding what information is stored during performance

of the dot pattern classification task (for a review, see Keri, 2003). The prototype account (Posner & Keele, 1968) argues that the prototype effect is evidence that some representation of the prototype was extracted and stored during the initial learning period in addition to a representation of each exemplar. Others have argued that the more accurate labeling of the prototype stimulus does not necessitate the extraction or storage in memory of the prototype itself and can result from generalized labeling using what is known about the stored exemplars. One such exemplar account, the Generalized Context Model (Nosofsky, 1986), proposes that each exemplar is represented and stored as a location in multidimensional space, where the dimensions are based on stimulus attributes such as color, shape, and size. Categorization of a novel stimulus is then based on its summed distance from and therefore its similarity to the representations of previously seen exemplars.

A simulation of the original dot pattern classification task used by Posner and Keele (1968) was run to assess whether the model would demonstrate the same pattern of responses to the novel stimuli: Are the exemplar effect and the prototype effect observed? If the simulations are faithful to the empirical results, examining the mechanism responsible for the emergence of these effects in the model may provide valuable insights into what information is used to perform the task by humans.

The model was run with the default parameters in Table 1. The stimuli and training procedure used in this experiment are detailed in Figure 4. Before testing, the simulations were again exposed to the 14 neutral stimuli used in Experiment 1. The simulations were then trained to label four exemplars from three categories and were finally tested on a range of novel exemplars, familiar ex-

emplars, and the prototype for each category. The stimuli used in this task were random dot patterns consisting of nine dots, with each dot centered on the middle of a 3 × 3 pixel square and with a degree of blurring to the surrounding pixels. Three different prototypes were created and from each, low distortions were created using the 4 bits-per-dot distortion level (Posner et al., 1967) and high distortions were created using the 6 bits-per-dot distortion level (Posner et al., 1967). Ten simulations were run, corresponding to multiple subjects.

Each network had three response units to represent the three categories being trained. Performance of the model in terms of concept identification was assessed by looking at the activity level of the three response units. To transform these activity values into a response, a random "noise" activity value between 0 and 0.5 was added to the activity level of each unit. The unit with the largest total activity value was then taken to represent the model's chosen category response. For each trial, a response was generated and depending on the stimulus this was identified as correct or incorrect, and percent correct over blocks of 10 trials was calculated. Percent correct in the test phase was averaged across all the networks, and responses to the four different stimulus types were compared using paired-sample *t* tests, with Bonferroni correction for multiple comparisons. For novel random dot patterns, there is no correct concept, so performance cannot be meaningfully gauged.

Results and Discussion

The average performance of the 10 networks on the 24 test stimuli is shown in Figure 4. Analysis of the data showed that both the prototype and exemplar effects

Figure 4. Simulation of categorization. Procedure shows the training stages and the stimuli presented to the network. All stimuli are derived from three prototype patterns, each a random arrangement of dots. See Experiment 2 methods for details. **Stimuli that are derived from one prototype are related, although different. They are said to be from the same concept and require the model to generate the same response. Results show performance of ten networks on the categorization task, showing the probability of correctly identifying the concept for each stimulus type. Error bars indicate *SEM*.
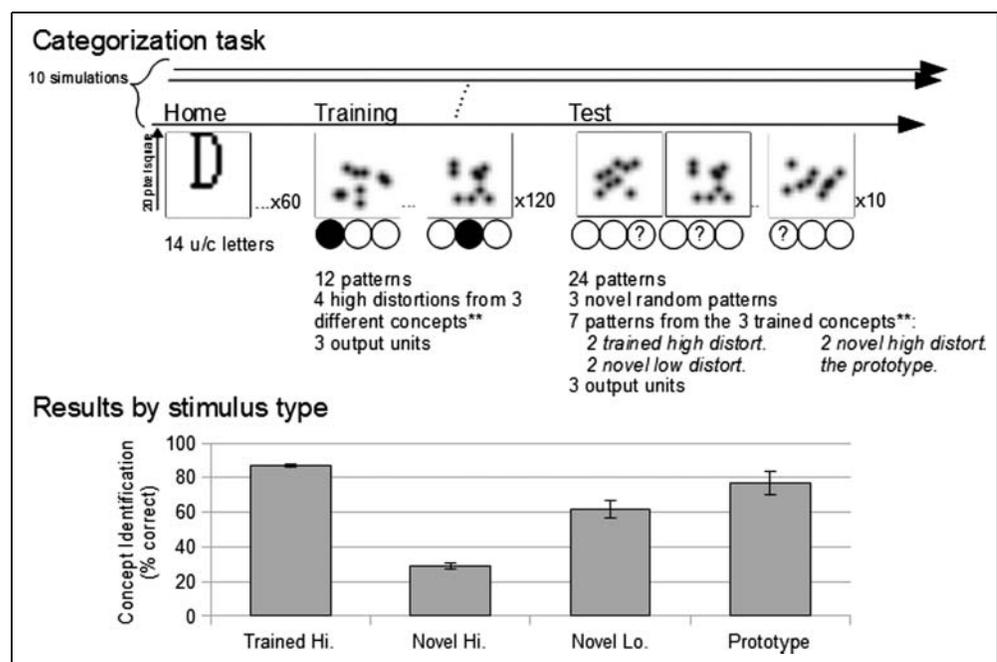
are seen in the model simulations, echoing the pattern of behavior seen in humans (Knowlton & Squire, 1993). Categorization of the prototype was significantly different from that of the novel high distortions ($p < .001$) and from the novel low distortions ($p < .01$), demonstrating a prototype effect. Categorization of the old high distortions was significantly different from that of the novel high distortions ($p < .001$), demonstrating an exemplar effect. In addition, as seen in the human data (Posner & Keele, 1968), the novel low distortions were categorized significantly better than the novel high distortions ($p < .001$).

Thus, the model can account for several key features of human categorization performance when presented with stimuli that closely approximate those used in behavioral studies of categorization. How the model achieves this is of interest. On test, the model generates a pattern of activity across its units in response to a given stimulus and uses this to generate a category response. Both novel and familiar stimuli are treated equally—the primary difference being that the pattern of activity generated in response to a novel stimulus will be weak and distributed across many units, whereas that for a familiar stimulus will be stronger and involve fewer units as a direct consequence of learning and some more active units outcompeting their neighbors (Figure 2). The extent to which a clear category is given in response to this pattern will depend on which units are active and whether they are strongly associated with only one of the three trained categories. Thus, novel exemplars are less able to generate the correct concept responses than familiar exemplars due to a weak pattern of activity but are able to generate above chance performance (33.3%) due to a similar pattern of activity to that evoked by the trained exemplars. In the isolated case of the prototype, this weak distributed pattern of activity is unusual in that it overlaps to a very large extent with the units active for all of the trained exemplars and is therefore better able to generate a correct concept response than the familiar exemplars in spite of its novelty and weak distributed activity. Critically, this performance is achieved using the same computational algorithm that was used to simulate recognition memory in Experiment 1 (and to simulate further tasks in Experiments 3 and 4).

At the moment, there is not a clear consensus in the literature regarding the role of teaching signals in category learning. The most well-known computational models of human categorization (e.g., ALCOVE, Kruschke, 1992; the generalized context model, Nosofsky & Palmeri, 1997; the Rational model, Anderson, 1991; and connectionist approaches, Rogers & McClelland, 2004) have mainly been explored in the context of supervised learning. However, there also exist many unsupervised models of category learning, in which no explicit teaching signals are provided, but instead items are grouped into categories based on their observed properties, and then these categories are used to make inferences about a new item's class membership. Indeed, unsupervised competitive learning or Kohonen networks are quite good at solving categorization problems as long as the data clusters are relatively easily separable (Rumelhart & Zipser, 1986; Kohonen, 1982; Grossberg, 1976a, 1976b). Our model falls into this latter camp and, in principle, should be able to perform the same types of classification problems as other competitive or Kohonen learning models. The sort of categorization that our model performs may well be somewhat different from other, semantically richer forms of categorization: As mentioned previously, unsupervised learning may be sufficient to solve categorization problems in which the data are easily separable, but a teaching signal may become necessary as the classification becomes more difficult (e.g., see Kohonen's LVQ2.1 algorithm; Kohonen, 1990). The present work therefore does not represent a repudiation of extant work on categorization that incorporates a teaching signal but is consistent with extant unsupervised models that indicate that certain types of categorization problem are solvable with an unsupervised network.

Interestingly, recent work exploring the role of teaching signals on category learning in humans suggests that teaching signals are not as essential as has been traditionally assumed (e.g., Kalish, Rogers, Lang, & Zhu, 2011). In these studies, it was found that unlabeled experiences, where no category information is present, can alter beliefs about category structure, but only if these unlabeled trials are drawn from a shifted distribution of categories to the original trained trials. Such a finding highlights the extent to which category learning may not require any teaching signals to shape internal representations and is consistent with our model, in which learning takes place in all trials, regardless of category information, to alter the landscape of stimulus representation within the self-organizing network.

EXPERIMENT 3: SIMULATION OF PERCEPTUAL LEARNING

Perceptual learning is thought to be a form of nondeclarative implicit learning (Schacter, Chiu, & Ochsner, 1993). It was first shown with rats learning to discriminate two similar geometric figures (Gibson & Walk, 1956), and has subsequently been demonstrated in humans and other species with a range of stimulus types (for a review, see Gilbert et al., 2001). The basic phenomenon is that pre-exposure to the stimuli enables faster subsequent learning of different responses to those stimuli, but this occurs only for difficult discriminations between very similar stimuli (Oswalt, 1972).

An initial explanation for this phenomenon attributed it to an increased ability to discriminate more properties of the stimuli (Gibson & Gibson, 1955), making an individual more sensitive to the differences that existed between the stimuli. Although alternative accounts exist

Figure 5. Simulation of perceptual learning. Procedure shows the training stages and the stimuli presented to the network. Results show the average performance of all networks on the stimulus discrimination over 10 blocks of 10 iterations (a total of 100 iterations) both with (diamonds) and without (crosses) preexposure. Chance performance is at 50%. Error bars indicate *SEM*.
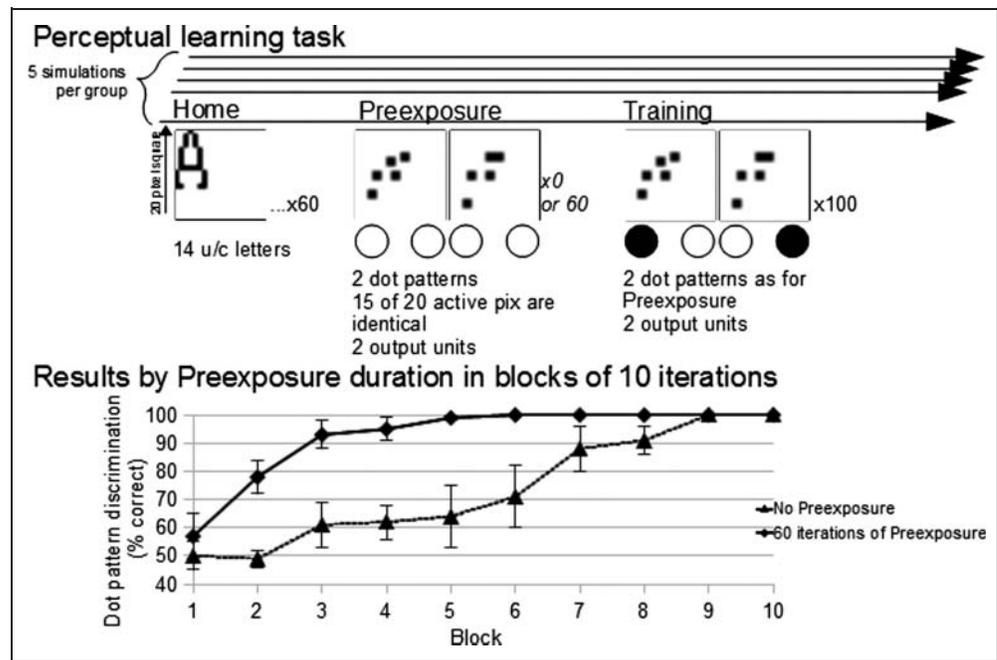
(McLaren, Kaye, & Mackintosh, 1989), some recent theoretical models are sympathetic with Gibson's idea of changes taking place in the stimulus representation (Saksida, 1999; Gaffan, 1996): A SOFM exposed to two similar stimuli learns to devote a larger number of units to representing those stimuli and so reduces the overlap in the stimulus representation and enables faster discrimination learning to take place when compared with a non-preexposed group (Saksida, 1999).

The model was run with the default parameters in Table 1. The stimuli and training procedure used in this experiment are detailed in Figure 5. Ten simulations were run to replicate multiple subjects, with five being run on a perceptual learning task in which networks received preexposure to the stimuli before acquisition of the discrimination problem and five on a control task in which networks were trained on the discrimination problem with no prior stimulus exposure. Once again training began with exposure for all simulations to the 14 neutral stimuli. The pre-exposure simulations were then presented with the two test stimuli for a fixed number of trials before all simulations were trained on a discrimination between the two test stimuli.

Each network had two response units. The performance of the model in terms of expectation of reward was assessed by looking at the activity level of the two response units, representing expectation of the presence or absence of reward. To transform these activity values into a probability of responding, a random "noise" activity value between 0 and 0.5 was added to the activity level of each unit. The unit with the largest total activity value was then taken to represent the model's action, either expecting or not expecting reward. For each trial, the response was scored as correct or incorrect, depending on the stimulus. This, in turn, was used to calculate a per-

cent correct over blocks of 10 trials, which was averaged across all preexposure networks and all non-pre-exposure networks and analyzed by ANOVA.

Results and Discussion

ANOVA of the percent correct performance (Figure 5) showed a significant main effect of Block ($F(9, 72) = 29.20, p < .001$), a significant main effect of Pre-exposure Group ($F(1, 8) = 17.33, p < .01$), and a significant interaction between Test Block and Pre-exposure Group ($F(9, 72) = 6.28, p < .001$). These results demonstrate that all networks were able to improve discrimination performance over the series of 100 trials. However, there was a significant difference in discrimination performance following pre-exposure when compared with performance without pre-exposure. A post hoc analysis of the interaction between the block of testing and the exposure group (comparison of means with Bonferroni-corrected p values) showed that the exposure conditions differed significantly only in Blocks 2, 3, 4, 5, and 6 (all $p < .001$). Thus networks in the preexposure and naive conditions showed the same initial discrimination performance at the start of testing, reached the same asymptotic level of performance at the end of testing but showed significantly different rates of acquisition.

This classic observation of perceptual learning is explained by the model recruiting more units to represent the stimuli during exposure and is in line with other models of the phenomenon (Saksida, 1999). This happens because similar stimuli generate similar patterns of activity in the model and with repeated presentation two major changes occur in these patterns of activity: the peak activity value increases as these active units in the model update their weights to better reflect the stimulus, and the activity

patterns for the two stimuli overlap less. Having well-separated patterns of activity representing the stimuli at the beginning of the discrimination task clearly facilitates learning an association of only one stimulus with reward. Without preexposure, the separation must occur simultaneously with the stimulus–reward learning, slowing the learning of the task. The current model therefore appeals to a nonassociative perceptual account of perceptual learning, in line with an existing account of this task (Saksida, 1999).

EXPERIMENT 4: SIMULATION OF V1 RETINOTOPY AND ORIENTATION SELECTIVITY

Beginning with the classic work of Hubel and Wiesel (1962, 1963), the selective response properties of cells in primary visual cortex, V1, have been extensively researched and the precise topography has been mapped. Cells in V1 are retinotopically mapped (Tootell et al., 1988; Talbot & Marshall, 1941) and are selectively responsive to the orientation of a line stimulus in space, direction of motion, color and which eye is being stimulated (Tootell et al., 1988; Hubel & Wiesel, 1962, 1965, 1972). More recently, the orientation specificity of V1 cells has been shown to occur in a spatial pattern across the surface of the cortex, referred to as "pin wheels" or singularities (Bartfeld & Grinvald, 1992; Blasdel & Salama, 1986).

In addition to the patterns of topographic orientation selectivity that develop during infancy, the plasticity of topographic maps in adulthood has also been demonstrated following prolonged exposure to a limited range of stimuli (Jenkins, Merzenich, Ochs, Allard, & Guic-Robles, 1990). This finding demonstrates that the mechanisms driving changes in topographic mapping are not restricted to early life but are present into adulthood in primary sensory cortex and, therefore, may have mechanistic similarities with learning mechanisms that take place in other cortical structures in adulthood.

Many models have provided excellent and detailed simulations of V1 development (Goodhill & Richards, 1999; Barrow, Bray, & Budd, 1996; Swindale, 1996; Goodhill, 1993; Obermayer, Blasdel, & Schulten, 1992; Durbin & Mitchison, 1990; Willshaw & von der Malsburg, 1976, 1979). We used the current model—which we have already used to simulate high-level processes such as categorization and perceptual learning—to simulate the development of retinotopy and the subsequent plasticity of the retinotopy due to overexposure to a restricted set of stimuli. We also simulated the development of orientation selectivity and assessed the resulting spatial pattern.

The initial weights between the input space and the network layer were random apart from a small bias replicating biases used by other models of V1 (Goodhill, 1993; Willshaw & von der Malsburg, 1979) designed to mimic the resultant effects of the chemical axonal path-finding mechanisms. Specifically, the weight value for each connection, ranging from 0 to 1, was made up of two equally weighted terms—the normalized Cartesian distance between the location within the input layer of the input unit and the location within the network layer of the network unit and a random variable. The model was run with the default parameters from Table 1 with one exception: A larger network layer size of 40 × 40 units was used to better visualize the emerging representations. The stimuli and training procedure used in this experiment are detailed in Figures 6 and 7. Two models were run, one to simulate topography generation and plasticity in adulthood (Figure 6) and the other to simulate orientation selectivity (Figure 7).

We visually examined the nature of the model's pattern development. Topography can be observed by plotting each network unit's center of mass in the input space; that is, the location in the input with the strongest weights on average to that network unit (Goodhill, 1993). In the case of orientation selectivity, a number of features are consistently observed (Swindale, 1996): (1) the periodicity of the pattern, (2) linear zones in the pattern where regions of iso-orientation lie in parallel to each other, (3) saddle points that are both a local peak in orientation in one direction and a local valley in the orthogonal direction, (4) singularities at which a full set of orientation domains meet at a point, and (5) fractures where there is a larger step change in orientation. The presence or absence of these features will be discussed.

Results and Discussion

The main finding of this simulation is that the model develops topographic mapping in a similar manner to that of primary visual cortex. For the first 800 iterations the network was presented with stimuli extending over the entire input space; it can be seen in Figure 6 that the units in the network represent the entirety of input space with a roughly even distribution. For the last 200 iterations of the simulation (800–1000), the network was presented with stimuli that occurred only in the top left-hand corner of the input space. Following this phase, the network units are no longer evenly distributed over input space— adjacent units have been recruited to represent more densely the space where this restricted stimulus set is located. This finding of topography reflects a fundamental property of Kohonen networks (Kohonen, 1984). The additional finding of recruitment of additional units to over-represented stimulus–space also follows from the self-organizing nature of this and other similar models: because small amounts of learning occur in each trial, any stimulus that is seen in a greater number of trials can evoke larger levels of learning and pull in more units to represent it better. This second finding is also consistent with the empirical finding that, in primates, repeated stimulation to a restricted location in input space causes more cortical cells

to come to represent that stimulated area (Jenkins et al., 1990).

The model demonstrates a good approximation to the pattern of orientation selectivity seen in V1 (Figure 7). Of the five key features of orientation selectivity patterns seen empirically (Swindale, 1996), four have been highlighted in Figure 7. These are linear zones, saddle points, singularities and fracture points. The last feature of periodicity cannot readily be highlighted, but it is evident from viewing the figure that the different "stripes" of orientation selectivity are of roughly equal width across the network. Thus, all of the key features of V1 orientation selectivity are seen in the present simulation results.

The findings that topographic mapping and orientation selectivity can be simulated by the present self-organizing model show that it is able to simulate some fundamental features of V1 developmental learning. However, there is nothing in the design of the model that reproduces any aspect of cortical circuitry that is unique to V1. This finding, therefore, suggests that the development of the retinotopic mapping of V1 is less determined by any of the unique cytoarchitecture that distinguishes V1 from other neocortical areas, but rather by the inputs that V1 receives. This idea has already been empirically demonstrated: retinal projections that target V1 can be re-routed into primary auditory cortex by deafferentation of the thalamic medial geniculate nucleus shortly after birth in the ferret. When the adult primary auditory cortex is then studied, it is found to contain many of the features normally observed in V1 and also seen in the simulations

Figure 6. Simulation of topography in V1. Procedure shows the stimuli presented to the network, 800 iterations of blobs followed by 200 iterations of squares. Results show the center of mass of the weights as training progresses. The network starts with an initial small bias in the weights, demonstrated in the small amount of spread already present after 10 iterations. For the first 800 iterations the weights of the model modify so that the units acquire a center of mass reflecting the location of the stimuli—covering the entire input space. By 800 iterations, the network is topographically mapped. For the last 200 iterations the weights of the model learn to represent a fixed set of 30 stimuli in only the top left quadrant of input space. The units representing this location in space have pulled together, with more units being recruited from adjacent locations in space to better represent the restricted set of stimuli. By 1000 iterations, the grid pattern of the locations where each stimulus was is visible in the center of mass of the weights.
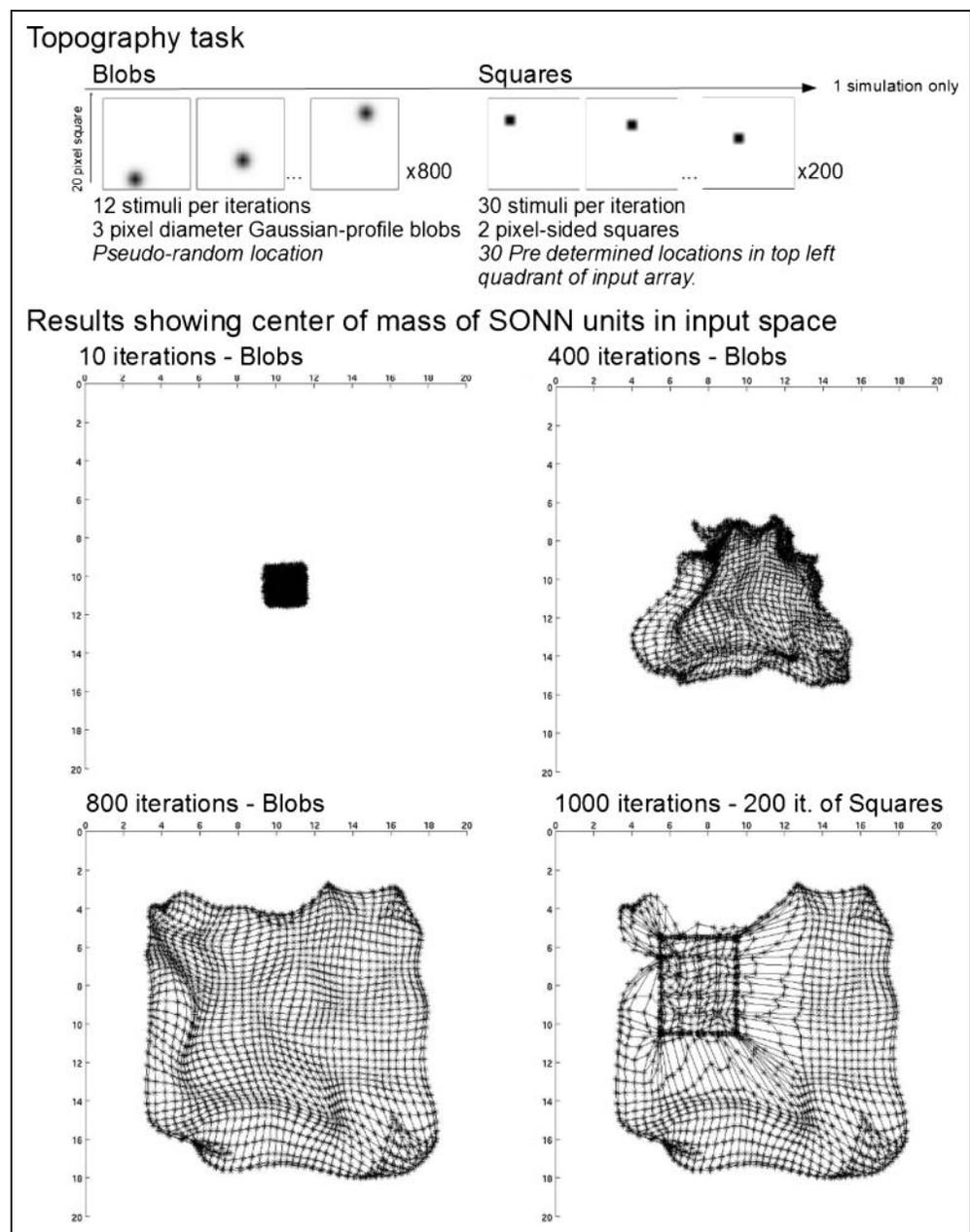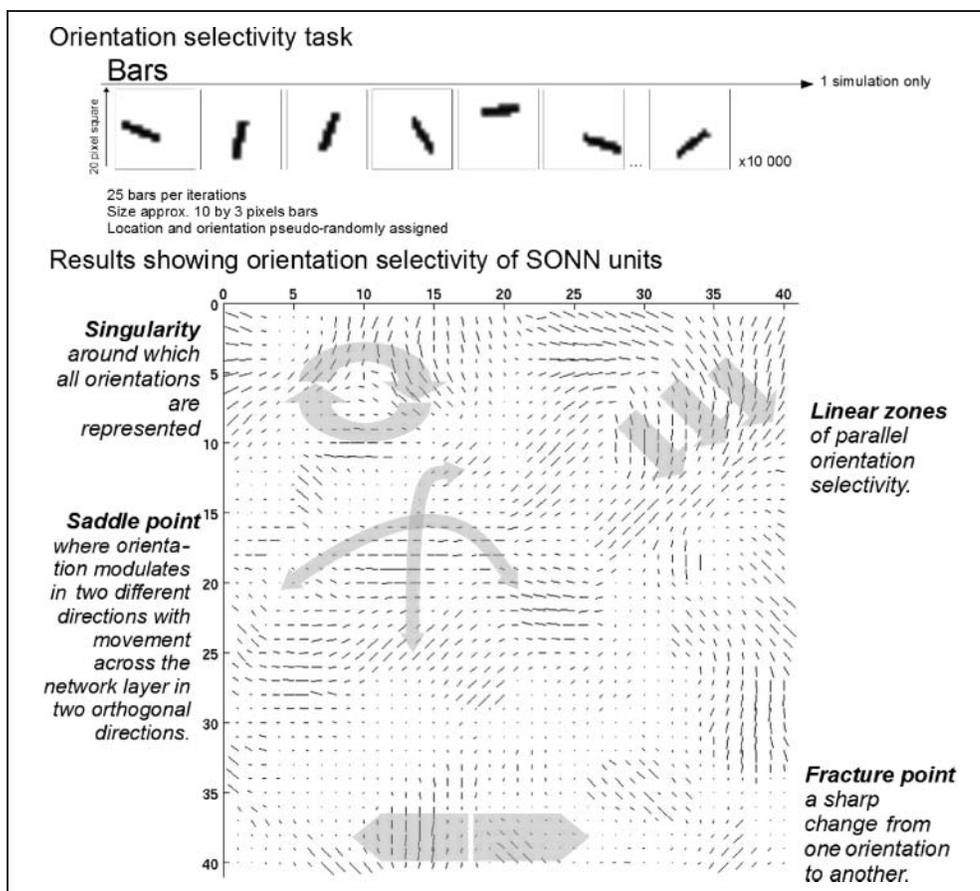
Figure 7. Simulation of orientation selectivity in V1. Procedure shows the stimuli presented to the network for 10,000 iterations. Results shows the orientation selectivity of each unit on the network layer after 10,000 iterations. The four features present in the representation are highlighted and labeled.

presented here, including orientation singularities and saddle points (Sharma, Angelucci, & Sur, 2000). This would support our claim that the pattern of orientation selectivity observed may be produced by a learning algorithm that is generic to many areas of neocortex.

GENERAL DISCUSSION

The brain is capable of many feats of visual processing, including functions as diverse as visual priming; perceptual learning; simultaneous discrimination of visual stimuli; object identification; and recognition memory for objects, faces, scenes, and even simple patterns. Under the traditional view of visual cognition, many of these functions are seen as being underpinned by separate regions of the brain. For example, visual priming is typically assigned to posterior VVS whereas recognition memory is localized to medial-temporal lobe structures (Squire, Stark, & Clark, 2004; Tulving & Schacter, 1990). Furthermore, low-level functions such as the development of visual cortex are typically studied completely separately from higher-level functions such as recognition memory, and it is rare to consider such phenomena together.

In contrast, under the representational–hierarchical account, the VVS is thought of as a functional continuum: Any region can potentially contribute to any of the foregoing functions, whether "perceptual" or "mnemonic,"

to the extent that performing the function requires the kinds of stimulus representation residing in that brain region. Because the complexity of stimulus representations increases continuously from simple features in posterior VVS to complex feature conjunctions in anterior regions, cognitive function under this account varies continuously from posterior to anterior regions. All tasks requiring representations of simple visual features—whether for discrimination on the basis of simultaneously presented features (in a perceptual task) or discrimination on the basis of familiarity (in a recognition memory task)—will depend on regions in posterior VVS, and all tasks requiring discrimination at the level of whole objects will depend on anterior regions. Under the representational–hierarchical account, the contribution of any region to visual cognition is determined by the stimulus inputs received by the region and the consequent nature of the stimulus representations it contains.

If different regions in VVS each contribute to cognition in the same way, with their function modified only by the particular flavor of stimulus representation that they contain, an important implication is that different regions may be using the same cortical mechanism (simply operating on different representations) for any given task. That is, these regions might be neurocomputationally homogenous, but by virtue of the different stimulus inputs they receive, give the appearance of possessing different

specialized cognitive mechanisms. This idea, although advocated by us on numerous occasions, has never previously been formally computationally tested.

In this study, we developed a neurocomputationally plausible model of visual cortex, consisting of a single layer of stimulus representations that develop through an unsupervised, self-organizing learning algorithm in a manner strongly influenced by the inputs that the network receives. Using the same parameters and learning algorithm throughout, we successfully simulated four tasks with known behavioral or neural outcomes: stimulus recognition memory, categorization of dot patterns, perceptual learning with dot patterns, and the development of orientation-selective topographic features in visual cortex. These four phenomena are associated with distinct regions of the ventral visual-perirhinal stream, namely, PRC (Squire et al., 2007; Murray et al., 2005; Winters et al., 2004), inferior temporal cortex (Keri, 2003), extrastriate cortex (Gilbert et al., 2001), and primary visual cortex (Bartfeld & Grinvald, 1992; Blasdel & Salama, 1986), respectively. For stimulus recognition memory we found, in line with the literature, that the model showed delay dependent performance that was sensitive to interference without showing catastrophic losses (Forwood et al., 2005; Winters et al., 2004), as well as demonstrating repetition-induced response reductions (Zhu & Brown, 1995) in the majority of network units and functional organization (Erickson et al., 2000). In the categorization learning task, the model was able to reproduce the exemplar effect and the prototype effect (Posner & Keele, 1968). In a simulation of perceptual learning, the model demonstrated an advantage in acquiring a visual discrimination problem following simple preexposure to the stimuli subsequently used in the task, in line with many studies of animal behavior (Gilbert et al., 2001; Gibson & Walk, 1956). Finally, the same model, when trained with very simple oriented line stimuli, reproduced all the major features of V1 orientation-selective topography observed from electrophysiological (Jenkins et al., 1990; Hubel & Wiesel, 1962) and cortical imaging studies (Bartfeld & Grinvald, 1992; Blasdel & Salama, 1986), such as plasticity of topography, periodicity of the pattern, linear zones, saddle points, singularities, and fractures (Swindale, 1996).

The principal aim of this study was to test the idea that stimulus inputs and task requirements are sufficient to drive the emergence of distinct regions in VVS that appear to be "specialized" for solving a particular task, if we assume the presence of a common neurocomputational algorithm throughout the VVS. We found that a single, unified cortical algorithm was able to simulate a diverse set of phenomena, traditionally associated with quite distinct areas of VVS. This contributes an important demonstration: the apparent specialization of cognitive function in anatomically distinct regions of visual cortex might simply reflect differences in the stimulus inputs to—and therefore the representational content of—those regions. Accordingly, when drawing inferences about the cognitive specialization of brain regions from either imaging studies or neuropsychological experiments, it is important to consider the representational requirements imposed by the stimuli and the instructions used in the task (Cowell et al., 2010).

Moreover, these simulations indicate that a unified cortical learning mechanism can construct the various layers in the representational hierarchy that we have hitherto simply assumed (Cowell et al., 2006, 2010; Bussey & Saksida, 2002). In previous simulations, we assumed a hierarchy by postulating multiple network layers across which stimulus representations increase in complexity; with the present model, we simulated functions previously associated with each of the separate layers, using the same layer of stimulus representations and changing only the stimulus inputs and the task structure. Interestingly, there are inherent differences between the stimuli used across the different tasks. For example, within-category similarity levels are much higher for simple line stimuli than for complex letter stimuli, because each simple line item contains fewer features; by possessing a greater number of stimulus features, the complex letters effectively reside in a much higher-dimensional space. These properties of the input stimuli influence the stimulus representations that emerge in the network, with the effect that, for each task, the single network layer mimics whichever layer in the hierarchy of our previous models was important for the task. In other words, we allowed the stimulus inputs to the single network layer to drive the emergence of stimulus representations at the appropriate level of complexity, with no assumptions about what that complexity might be or where in the brain it should be found. This provides a very pure test of the ability of stimulus inputs and task requirements to drive the emergence of appropriate representations. The present simulations exploit these diverse representational properties to account for apparent differences in cognitive function across different regions of the VVS, without assuming distinct neural mechanisms (cf. Cowell et al., 2009; Zaki & Nosofsky, 2001; Plaut, 1995).

The view that much of neocortex might function in the same manner is not new; it has a long history going back at least as far as the work of Lashley and his ideas of cortical mass action (Lashley, 1950) and has modern proponents in the work of Fuster (2006, 2009), Foster and Jelicic (1999), and Goldstone and Barsalou (1998). Fuster (2009) has proposed a new paradigm of cortical memory, where memory cognits are composed of distributed patterns of activity spanning multiple cortical areas using both bottom–up and feedback connectivity. This idea is well supported with empirical evidence from imaging and single-unit recording experiments but does not attempt to expand upon what computations may be happening within each cortical area. The current work does this and is broadly in line with Fuster's paradigm. It also contributes to the debate on the localization and

specialization of cognitive function in the cortex by presenting an account of how the cortex might function in a computationally uniform manner, while giving the appearance of cognitive modularity (see also Op de Beeck et al., 2008; Cosmides & Tooby, 1994).

In summary, this study tests the idea that the simple existence of different levels of representational complexity in different parts of the VVS is sufficient to drive the emergence of distinct regions that appear to be specialized for solving a particular task, when a common neurocomputational learning algorithm is assumed across all regions. Of course, the model used here is highly simplified, and we are not claiming that the algorithm and circuitry used here explain everything about the functioning of the cortex. Different neurotransmitter activity, cell types, etc., could and almost certainly do modulate the functions of different cortical regions, endowing them with at least somewhat different properties. However, what we have demonstrated here is that it is not necessary to invoke such differences to understand how different cortical regions can appear to be specialized for what are considered to be very different psychological functions. Potentially much more important than these putative differences are the commonalities across regions, and by focusing on the differences we risk missing the wood for the trees.

APPENDIX 1: COMPUTATIONAL METHODS

The model used in the current article is based on a Kohonen SOFM (Kohonen, 1984); as in the SOFM, the main computations in the model are executed within a single layer of heavily interconnected units. All units in this main layer receive a weighted input from all of the units in an input array and all send projections to the same target outside the layer (Figure 1). As is the case in the cortex, all the units in the main layer engage in lateral excitation and inhibition with their neighbors within the main layer, in other words there are weighted connections between all the units in the main layer. The input–main layer connections change as a function of a learning rule, so that they are updated from one trial to the next as a function of input activity, unit activity, and current weight. The connections within the main layer emulate the lateral connectivity within neocortex: they follow a Gaussian profile, such that close neighbors have stronger connectivity than distance neighbors (Thomson & Deuchars, 1997); the profile of the inhibitory connections is three-fold wider than that of the excitatory connections (Angelucci et al., 2002), and as with extraclassical receptive fields, the width of the Gaussian profile is dynamic, being small when the stimulus is optimal and evokes a high peak activity level and larger when a less optimal stimulus is presented (Angelucci et al., 2002).

Each unit in the model layer has an activation value that represents a firing rate, and spike-timing is not instantiated in this model. Activation values for each computational unit in the current model are limited to within an upper and lower bound: The lower bound, 0, represents the spontaneous noise level of the neuron, and the upper bound, 1, represents the saturation firing rate of the neuron. If the activity value for any unit is calculated to be outside this range, its value is set to 0 or 1 as appropriate. In practice only a small proportion of units are affected by this threshold, so these are not binary units.

The input to the model is a two-dimensional array of units, each one of which can have an activity value of between 0 and 1. In the current model, we use realistic two-dimensional images of visual stimuli—gray-scale representations of lines, shapes, and objects within a 20 × 20 pixel input space—which are, where possible, the exact same stimuli used to collect the original behavioral data.

A single trial proceeds as follows:

1. A stimulus is selected according to the protocol for each task. This is presented to the network layer of units and the resulting activity, a_j, of each unit in the network layer, j, is calculated using

$$a_j = \sum_i \frac{a_i \times w_{ij}}{\sum_i a_i} \tag{1}$$

where a_i is the activity for each input unit i and w_{ij} is the weight between input unit i and network unit j.

2. To reduce the levels of activity within the network layer and the number of units able to engage in competition, the activity of most units is reduced to close to 0 and the peak is reduced by the mean activity across the layer, as specified by Equation 2:

$$a_j' = a_j \cdot \left(\frac{a_j}{\max(a_j)}\right)^2 - \bar{a}_j \quad a_j' = \left\{ \begin{array}{l} 1 \text{ if } a_j' > 1 \\ 0 \text{ if } a_j' < 0 \end{array} \right\} \tag{2}$$

3. The neighborhood size N_t for the current trial, t, is then calculated using the new peak activity, $\max(a_j')$. This value lies between the maximal neighborhood value, N_{net}, for very weak peak activity and the minimal neighborhood value, N_{min}, for peak activity levels of 1. This range is a constant parameter of the model. The value of N_t, determines the inhibitory and excitatory neighborhood sizes, Ni_{mn}, the inhibitory neighborhood size, is taken as three times larger than the excitatory neighborhood size, Ne_m, in line with the relative sizes of the effects seen (Angelucci et al., 2002).

$$N_t = N_{min} + (N_{net} - N_{min}) \cdot e^{-\left(\frac{max(a_j')}{0.25}\right)} \quad \begin{array}{l} Ne_t = N_t \\ Ni_t = 3 \cdot N_t \end{array} \tag{3}$$

4. Both the excitatory and inhibitory lateral interaction between each unit and every other unit in the layer are modeled using a matrix of lateral weights. These

weights are defined using a Gaussian profile, as in (4) where x_{jk} is the Cartesian distance between the two network units j and k and the neighborhood sizes are as calculated each trial (3).

$$we_{jk} = e^{-\left(\frac{x_{jk}}{Ne_t}\right)^2} \qquad wi_{jk} = e^{-\left(\frac{x_{jk}}{Ni_t}\right)^2} \qquad (4)$$

Using these weights, the activity of each unit in the layer, a_j'', is then recalculated based on the lateral excitatory and inhibitory weights calculated in (4) and the activity of all the other units in the layer a_k' based on the activity values for each unit generated by (2).

$$a_j'' = \max(a_j' \cdot we_{jk}) \cdot \left(\frac{\max(a_j' \cdot we_{jk})}{\max(a_j' \cdot wi_{jk})}\right)^2$$

$$a_j'' = \begin{Bmatrix} 1 & \text{if } a_j'' > 1 \\ 0 & \text{if } a_j'' < 0 \end{Bmatrix} \qquad (5)$$

The activity value for each unit, a_j'', provides the main output value of the network unit and is used to determine the amount of learning that can take place at each unit in the network.

5. The learning rule used on the input to network layer weights is a form of Hebbian learning. It should be noted that this equation closely resembles Oja's learning rule, a more stable modification of the standard Hebbian Rule and an algorithm for principal components analysis (Oja, 1982).

$$w_{ij}(t+1) = w_{ij}(t) + a_j'' \cdot \lambda \cdot (a_i - w_{ij}(t)) \qquad (5)$$

According to this equation, large changes in a weight can only take place when the network unit is active. The direction of the weight change is determined by the difference term $a_i - w_{ij}(t)$: It is positive if the input unit activity (a_i) is greater than the current weight ($w_{ij}(t)$), and negative if the current weight ($w_{ij}(t)$) is greater than the input activity (a_i). λ represents a learning rate parameter that in the current simulations remains constant.

6. Some of the simulations require the model to learn associations between specific stimuli and a response or outcome. To simulate this, the network layer sends an output to a stimulus–reward associative learning mechanism. The main layer provides a pattern of activity which can engage in error correction learning to associate the stimulus, as represented by the network layer, with a number of response units using a Rescorla–Wagner or delta learning rule (Rescorla & Wagner, 1972; Widrow & Hoff, 1960). In this learning rule (Equation 7), it is the difference between the presence or absence of reward, R, and the expectation of reward signaled by the activity of the response unit, a_r, that

drives weight changes and hence stimulus–response learning. An additional variable in the equation, α, represents a learning rate parameter that determines how quickly learning takes place, and in the following simulations remains constant.

$$a_r = \sum_j \frac{w_{mr} \cdot a_j''}{\sum_j a_j''} \qquad a_r = \begin{Bmatrix} 1 & \text{if } a_r > 1 \\ 0 & \text{if } a_r < 0 \end{Bmatrix} \qquad (6)$$

$$w_{jr}(t+1) = w_{jr}(t) + a_j'' \cdot \alpha \cdot (R - a_r) \qquad (7)$$

The default parameters used in all the simulations presented here are shown in Table 1. Where this is not the case, the parameters used and the justification for the change are described. The overall mechanism of the model and how it deals with repeated stimuli in a dynamic sense is illustrated in Figure 2.

REFERENCES

Anderson, J. R. (1991). The adaptive nature of human categorization. *Psychological Review, 98,* 409–429.

Angelucci, A., Levitt, J. B., Walton, E. J. S., Hupe, J.-M., Bullier, J., & Lund, J. S. (2002). Circuits for local and global signal integration in primary visual cortex. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 22,* 8633–8646.

Barense, M. D., Bussey, T. J., Lee, A. C. H., Rogers, T. T., Davies, R. R., Saksida, L. M., et al. (2005). Functional specialization in the human medial temporal lobe. *Journal of Neuroscience, 25,* 10239–10246.

Barrow, H. G., Bray, A. J., & Budd, J. M. (1996). A self-organizing model of "color blob" formation. *Neural Computation, 8,* 1427–1448.

Bartfeld, E., & Grinvald, A. (1992). Relationships between orientation-preference pinwheels, cytochrome oxidase blobs, and ocular-dominance columns in primate striate cortex. *Proceedings of the National Academy of Sciences, U.S.A., 89,* 11905–11909.

Bartko, S. J., Cowell, R. A., Winters, B. D., Bussey, T. J., & Saksida, L. M. (2010). Heightened susceptibility to interference in an animal model of amnesia: Impairment in encoding, storage, retrieval-or all three? *Neuropsychologia, 48,* 2987–2997.

Blasdel, G. G., & Salama, G. (1986). Voltage-sensitive dyes reveal a modular organization in monkey striate cortex. *Nature, 321,* 579–585.

Bogacz, R., & Brown, M. W. (2003). Comparison of computational models of familiarity discrimination in the perirhinal cortex. *Hippocampus, 13,* 494–524.

Buckley, M. J., Booth, M. C., Rolls, E. T., & Gaffan, D. (2001). Selective perceptual impairments after perirhinal cortex ablation. *Journal of Neuroscience, 21,* 9824–9836.

Buckley, M. J., & Gaffan, D. (1998). Perirhinal cortex ablation impairs visual object identification. *Journal of Neuroscience, 18,* 2268–2275.

Buffalo, E. A., Reber, P. J., & Squire, L. R. (1998). The human perirhinal cortex and recognition memory. *Hippocampus, 8,* 330–339.

Bussey, T. J., Muir, J. L., & Aggleton, J. P. (1999). Functionally dissociating aspects of event memory: The effects of combined perirhinal and postrhinal cortex lesions on object and place memory in the rat. *Journal of Neuroscience, 19,* 495–502.

Bussey, T. J., & Saksida, L. M. (2002). The organization of visual object representations: A connectionist model of effects of lesions in perirhinal cortex. *European Journal of Neuroscience, 15,* 355–364.

Bussey, T. J., & Saksida, L. M. (2005). Object memory and perception in the medial temporal lobe: An alternative approach. *Current Opinion in Neurobiology, 15,* 730–737.

Bussey, T. J., Saksida, L. M., & Murray, E. A. (2002). Perirhinal cortex resolves feature ambiguity in complex visual discriminations. *European Journal of Neuroscience, 15,* 365–374.

Bussey, T. J., Saksida, L. M., & Murray, E. A. (2003). Impairments in visual discrimination after perirhinal cortex lesions: Testing "declarative" vs. "perceptual-mnemonic" views of perirhinal cortex function. *European Journal of Neuroscience, 17,* 649–660.

Cosmides, L., & Tooby, J. (1994). Beyond intuition and instinct blindness: Toward an evolutionarily rigorous cognitive science. *Cognition, 50,* 41–77.

Cowell, R. A., Bussey, T. J., & Saksida, L. M. (2006). Why does brain damage impair memory? A connectionist model of object recognition memory in perirhinal cortex. *Journal of Neuroscience, 26,* 12186–12197.

Cowell, R. A., Bussey, T. J., & Saksida, L. M. (2010). Functional dissociations within the ventral object processing pathway: Cognitive modules or a hierarchical continuum? *Journal of Cognitive Neuroscience, 22,* 2460–2479.

Cowell, R. A., Huber, D. E., & Cottrell, G. W. (2009). *Virtual brain reading: A connectionist approach to understanding fMRI.* Paper presented at the 31st Annual Meeting of the Cognitive Science Society.

Desimone, R., Albright, T. D., Gross, C. G., & Bruce, C. (1984). Stimulus-selective properties of inferior temporal neurons in the macaque. *Journal of Neuroscience, 4,* 2051–2062.

Durbin, R., & Mitchison, G. (1990). A dimension reduction framework for understanding cortical maps. *Nature, 343,* 644–647.

Eacott, M. J., Gaffan, D., & Murray, E. A. (1994). Preserved recognition memory for small sets, and impaired stimulus identification for large sets, following rhinal cortex ablations in monkeys. *European Journal of Neuroscience, 6,* 1466–1478.

Ennaceur, A., & Delacour, J. (1988). A new one-trial test for neurobiological studies of memory in rats. 1: Behavioral data. *Behavioural Brain Research, 31,* 49–57.

Erickson, C. A., Jagadeesh, B., & Desimone, R. (2000). Clustering of perirhinal neurons with similar properties following visual experience in adult monkeys. *Nature Neuroscience, 3,* 1143–1148.

Fahy, F. L., Riches, I. P., & Brown, M. W. (1993). Neuronal activity related to visual recognition memory: Long-term memory and the encoding of recency and familiarity information in the primate anterior and medial inferior temporal and rhinal cortex. *Experimental Brain Research, 96,* 457–472.

Forwood, S. E., Winters, B. D., & Bussey, T. J. (2005). Hippocampal lesions that abolish spatial maze performance spare object recognition memory at delays of up to 48 hours. *Hippocampus, 15,* 347–355.

Foster, J. K., & Jelicic, M. (1999). *Memory.* Oxford: Oxford University Press.

Fukushima, K. (1980). Neocognitron: A self organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biological Cybernetics, 36,* 193–202.

Fuster, J. M. (2006). The cognit: A network model of cortical representation. *International Journal of Psychophysiology, 60,* 125–132.

Fuster, J. M. (2009). Cortex and memory: Emergence of a new paradigm. *Journal of Cognitive Neuroscience, 21,* 2047–2072.

Gaffan, D. (1996). Associative and perceptual learning and the concept of memory systems. *Cognitive Brain Research, 5,* 69–80.

Gauthier, I., & Tarr, M. J. (1997). Becoming a "Greeble" expert: Exploring mechanisms for face recognition. *Vision Research, 37,* 1673–1682.

Ghirlanda, S. (2005). Retrospective revaluation as simple associative learning. *Journal of Experimental Psychology: Animal Behavior Processes, 31,* 107–111.

Gibson, E. J., & Walk, R. D. (1956). The effect of prolonged exposure to visually presented patterns on learning to discriminate them. *Journal of Comparative and Physiological Psychology, 49,* 239–242.

Gibson, J. J., & Gibson, E. J. (1955). Perceptual learning—Differentiation or enrichment? *Psychological Review, 62,* 32–41.

Gilbert, C. D., Sigman, M., & Crist, R. E. (2001). The neural basis of perceptual learning. *Neuron, 31,* 681–697.

Goldstone, R. L., & Barsalou, L. W. (1998). Reuniting perception and conception. *Cognition, 65,* 231–262.

Goodhill, G. J. (1993). Topography and ocular dominance: A model exploring positive correlations. *Biological Cybernetics, 69,* 109–118.

Goodhill, G. J., & Richards, L. J. (1999). Retinotectal maps: Molecules, models and misplaced data. *Trends in Neurosciences, 22,* 529–534.

Grossberg, S. (1976a). Adaptive pattern-classification and universal recoding: 1. Parallel development and coding of neural feature detectors. *Biological Cybernetics, 23,* 121–134.

Grossberg, S. (1976b). Adaptive pattern-classification and universal recoding: 2. Feedback, expectation, olfaction, illusions. *Biological Cybernetics, 23,* 187–202.

Grossberg, S. (1994). 3-D vision and figure-ground separation by visual cortex. *Perception & Psychophysics, 55,* 48–121.

Hanson, S. J., Matsuka, T., & Haxby, J. V. (2004). Combinatorial codes in ventral temporal lobe for object recognition: Haxby (2001) revisited: Is there a "face" area? *Neuroimage, 23,* 156–166.

Haxby, J. V., Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., & Pietrini, P. (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science, 293,* 2425–2430.

Holdstock, J. S., Gutnikov, S. A., Gaffan, D., & Mayes, A. R. (2000). Perceptual and mnemonic matching-to-sample in humans: Contributions of the hippocampus, perirhinal and other medial temporal lobe cortices. *Cortex, 36,* 301–322.

Hubel, D. H., & Wiesel, T. N. (1959). Receptive fields of single neurones in the cat's striate cortex. *Journal of Physiology, 148,* 574–591.

Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *Journal of Physiology, 160,* 106–154.

Hubel, D. H., & Wiesel, T. N. (1963). Receptive fields of cells in striate cortex of very young, visually inexperienced kittens. *Journal of Neurophysiology, 26,* 994–1002.

Hubel, D. H., & Wiesel, T. N. (1965). Receptive fields and functional architecture in two nonstriate visual areas (18 and 19) of the cat. *Journal of Neurophysiology, 28,* 229–289.

Hubel, D. H., & Wiesel, T. N. (1972). Laminar and columnar distribution of geniculo-cortical fibers in the macaque monkey. *Journal of Comparative Neurology, 146,* 421–450.

Iwai, E., & Mishkin, M. (1968). Two visual foci in the temporal lobe of monkeys. In N. Yoshii & N. Buchwald (Eds.), *Neurophysiological basis of learning and behavior* (pp. 1–11). Japan: Osaka University Press.

Jenkins, W. M., Merzenich, M. M., Ochs, M., Allard, T., & Guic-Robles, E. (1990). Functional reorganization of primary somatosensory cortex in adult owl monkeys after behaviorally controlled tactile stimulation. *Journal of Neurophysiology, 63,* 82–104.

Kalish, C. W., Rogers, T. T., Lang, J., & Zhu, X. J. (2011). Can semi-supervised learning explain incorrect beliefs about categories? *Cognition, 120,* 106–118.

Kanwisher, N. G., McDermott, J., & Chun, M. M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience, 17,* 4302–4311.

Keri, S. (2003). The cognitive neuroscience of category learning. *Brain Research Reviews, 43,* 85–109.

Knowlton, B. J., & Squire, L. R. (1993). The learning of categories: Parallel brain systems for item memory and category knowledge. *Science, 262,* 1747–1749.

Kohonen, T. (1982). *Clustering taxonomy and topological maps of patterns.* Paper presented at the Sixth International Conference on Pattern Recognition, Silver Springs, MD.

Kohonen, T. (1984). *Self-organization and associative memory.* Berlin: Springer-Verlag.

Kohonen, T. (1990). Statistical pattern-recognition revisited. In R. Eckmiller (Ed). *Advanced neural computers* (pp. 137–144). Amsterdam: Elsevier.

Kruschke, J. K. (1992). ALCOVE: An exemplar-based connectionist model of category learning. *Psychological Review, 99,* 22–44.

Lashley, K. S. (1950). In search of the engram. In R. J. Pumphrey (Ed). *Society for Experimental Biology Symposium, No 4: Psychological mechanisms in animal behavior* (pp. 454–482). Cambridge: Cambridge University Press.

Lee, A. C. H., Barense, M. D., & Graham, K. S. (2005). The contribution of the human medial temporal lobe to perception: Bridging the gap between animal and human studies. *The Quarterly Journal of Experimental Psychology: Series B, Comparative and Physiological Psychology, 58,* 300–325.

Lee, A. C. H., Scahill, V. L., & Graham, K. S. (2008). Activating the medial temporal lobe during oddity judgment for faces and scenes. *Cerebral Cortex, 18,* 683–696.

Lopez-Aranda, M. F., Lopez-Tellez, J. F., Navarro-Lobato, I., Masmudi-Martin, M., Gutierrez, A., & Khan, Z. U. (2009). Role of layer 6 of V2 visual cortex in object-recognition memory. *Science, 325,* 87–89.

Manns, J. R., Stark, C. E. L., & Squire, L. R. (2000). The visual paired-comparison task as a measure of declarative memory. *Proceedings of the National Academy of Sciences, U.S.A., 97,* 12375–12379.

McLaren, I. P. L., Kaye, H., & Mackintosh, N. J. (1989). An associative theory of the representation of stimuli: Applications to perceptual learning and latent inhibition.

In R. G. M. Morris (Ed.), *Parallel distributed processing: Implications for psychology and neurobiology* (pp. 102–130). Oxford: Clarendon Press.

Miller, E. K., Li, L., & Desimone, R. (1991). A neural mechanism for working and recognition memory in inferior temporal cortex. *Science, 254,* 1377–1379.

Mishkin, M. (1982). A memory system in the monkey. *Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences, 298,* 83–95.

Murray, E. A., Bussey, T. J., & Saksida, L. M. (2001). *Perirhinal cortex resolves feature ambiguity in complex visual discriminations.* Paper presented at the Cognitive Neuroscience Society Annual Meeting, New York.

Murray, E. A., Graham, K. S., & Gaffan, D. (2005). Perirhinal cortex and its neighbours in the medial temporal lobe: Contributions to memory and perception. *Quarterly Journal of Experimental Psychology: Series B, 58,* 378–396.

Norman, K. A., & O'Reilly, R. C. (2003). Modeling hippocampal and neocortical contributions to recognition memory: A complementary-learning-systems approach. *Psychological Review, 110,* 611–646.

Nosofsky, R. M. (1986). Attention, similarity, and the identification-categorization relationship. *Journal of Experimental Psychology: General, 115,* 39–57.

Nosofsky, R. M., & Palmeri, T. J. (1997). An exemplar-based random walk model of speeded classification. *Psychological Review, 104,* 266–300.

Obermayer, K., Blasdel, G. G., & Schulten, K. (1992). Statistical-mechanical analysis of self-organization and pattern-formation during the development of visual maps. *Physical Review A, 45,* 7568–7589.

Oja, E. (1982). A simplified neuron model as a principal component analyzer. *Journal of Mathematical Biology, 15,* 267–273.

Op de Beeck, H. P., Haushofer, J., & Kanwisher, N. G. (2008). Interpreting fMRI data: Maps, modules and dimensions. *Nature Reviews Neuroscience, 9,* 123–135.

Oswalt, R. M. (1972). Relationship between level of visual pattern difficulty during rearing and subsequent discrimination in rats. *Journal of Comparative and Physiological Psychology, 81,* 122–125.

O'Toole, A. J., Jiang, F., Abdi, H., & Haxby, J. V. (2005). Partially distributed representations of objects and faces in ventral temporal cortex. *Journal of Cognitive Neuroscience, 17,* 580–590.

Perrett, D., & Oram, M. (1993). Neurophysiology of shape processing. *Image and Vision Computing, 11,* 317–333.

Plaut, D. C. (1995). Double dissociation without modularity: Evidence from connectionist neuropsychology. *Journal of Clinical Experimental Neuropsychology, 17,* 291–321.

Posner, M. I., Goldsmith, R., & Welton, K. E., Jr. (1967). Perceived distance and the classification of distorted patterns. *Journal of Experimental Psychology, 73,* 28–38.

Posner, M. I., & Keele, S. W. (1968). On the genesis of abstract ideas. *Journal of Experimental Psychology, 77,* 353–363.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. Black & W. F. Prokasy (Eds.), *Classical conditioning II* (pp. 64–99). New York: Appleton-Centure-Crofts.

Riesenhuber, M., & Poggio, T. (1999). Hierarchical models of object recognition in cortex. *Nature Neuroscience, 2,* 1019–1025.

Rogers, T. T., & McClelland, J. L. (2004). *Semantic cognition: A parallel distributed processing approach.* Cambridge, MA: MIT Press.

Rolls, E. T. (1992). Neurophysiological mechanisms underlying face processing within and beyond the temporal cortical visual areas. *Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences, 335,* 11–20; discussion 20.

Rumelhart, D. E., & Zipser, D. (1986). Feature discovery by competitive learning. In D. E. Rumelhart & J. L. McClelland (Eds.), *Parallel distributed processing* (Vol. 1, pp. 151–193). Cambridge, MA: MIT Press.

Sakai, S., & Miyashita, Y. (1993). Memory and imagery in the temporal lobe. *Current Opinion in Neurobiology, 3,* 166–170.

Saksida, L. M. (1999). Effects of similarity and experience on discrimination learning: A nonassociative connectionist model of perceptual learning. *Journal of Experimental Psychology: Animal Behavior Processes, 25,* 308–323.

Schacter, D. L., Chiu, C. Y., & Ochsner, K. N. (1993). Implicit memory: A selective review. *Annual Review of Neuroscience, 16,* 159–182.

Scoville, W. B., & Milner, B. (1957). Loss of recent memory after bilateral hippocampal lesions. *Journal of Neurology, Neurosurgery and Psychiatry, 20,* 11–21.

Sharma, J., Angelucci, A., & Sur, M. (2000). Induction of visual orientation modules in auditory cortex. *Nature, 404,* 841–847.

Spiridon, M., & Kanwisher, N. G. (2002). How distributed is visual category information in human occipito-temporal cortex? An fMRI study. *Neuron, 35,* 1157–1165.

Squire, L. R., Stark, C. E., & Clark, R. E. (2004). The medial temporal lobe. *Annual Review of Neuroscience, 27,* 279–306.

Squire, L. R., & Wixted, J. T. (2011). The cognitive neuroscience of human memory since h.m. *Annual Review of Neuroscience, 34,* 259–288.

Squire, L. R., Wixted, J. T., & Clark, R. E. (2007). Recognition memory and the medial temporal lobe: A new perspective. *Nature Reviews Neuroscience, 8,* 872–883.

Squire, L. R., & Zola-Morgan, S. M. (1991). The medial temporal lobe memory system. *Science, 253,* 1380–1386.

Swindale, N. V. (1996). The development of topography in the visual cortex: A review of models. *Network, 7,* 161–247.

Talbot, S. A., & Marshall, W. H. (1941). Physiological studies on neural mechanisms of visual localization and discrimination. *American Journal of Ophthalmology, 24,* 1255–1264.

Tanaka, K. (2003). Columns for complex visual object features in the inferotemporal cortex: Clustering of cells with similar but slightly different stimulus selectivities. *Cerebral Cortex, 13,* 90–99.

Tanaka, K., Saito, H., Fukada, Y., & Moriya, M. (1991). Coding visual images of objects in the inferotemporal cortex of the macaque monkey. *Journal of Neurophysiology, 66,* 170–189.

Thomson, A. M., & Deuchars, J. (1997). Synaptic interactions in neocortical local circuits: Dual intracellular recordings in vitro. *Cerebral Cortex (New York, N.Y.: 1991), 7,* 510–522.

Tootell, R. B., Dale, A. M., Sereno, M. I., & Malach, R. (1996). New images from human visual cortex. *Trends in Neuroscience, 19,* 481–489.

Tootell, R. B., Switkes, E., Silverman, M. S., & Hamilton, S. L. (1988). Functional anatomy of macaque striate cortex. II. Retinotopic organization. *Journal of Neuroscience, 8,* 1531–1568.

Tsao, D. Y., & Livingstone, M. S. (2008). Mechanisms of face perception. *Annual Review of Neuroscience, 31,* 411–437.

Tulving, E., & Schacter, D. L. (1990). Priming and human memory systems. *Science, 247,* 301–306.

Tyler, L. K., Stamatakis, E. A., Bright, P., Acres, K., Abdallah, S., Rodd, J. M., et al. (2004). Processing objects at different levels of specificity. *Journal of Cognitive Neuroscience, 16,* 351–362.

Wallis, G., & Rolls, E. T. (1997). Invariant face and object recognition in the visual system. *Progress in Neurobiology, 51,* 167–194.

Widrow, B., & Hoff, M. E. (1960). *Adaptive switching circuits.* New York: Institute of Radio Engineers.

Willshaw, D. J., & von der Malsburg, C. (1976). How patterned neural connections can be set up by self-organization. *Proceedings of the Royal Society for London, Series B, Biological Sciences, 194,* 431–445.

Willshaw, D. J., & von der Malsburg, C. (1979). A marker induction mechanism for the establishment of ordered neural mappings: Its application to the retinotectal problem. *Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences, 287,* 203–243.

Winters, B. D., Forwood, S. E., Cowell, R. A., Saksida, L. M., & Bussey, T. J. (2004). Double dissociation between the effects of peri-postrhinal cortex and hippocampal lesions on tests of object recognition and spatial memory: Heterogeneity of function within the temporal lobe. *Journal of Neuroscience, 24,* 5901–5908.

Zaki, S. R., & Nosofsky, R. M. (2001). A single-system interpretation of dissociations between recognition and categorization in a task involving object-like stimuli. *Cognitive, Affective & Behavioral Neuroscience, 1,* 344–359.

Zhu, X. O., & Brown, M. W. (1995). Changes in neuronal activity related to the repetition and relative familiarity of visual stimuli in rhinal and adjacent cortex of the anaesthetised rat. *Brain Research, 689,* 101–110.

Zola-Morgan, S. M., Squire, L. R., Amaral, D. G., & Suzuki, W. A. (1989). Lesions of perirhinal and parahippocampal cortex that spare the amygdala and hippocampal formation produce severe memory impairment. *Journal of Neuroscience, 9,* 4355–4370.