

The perirhinal cortex and long-term familiarity memory

E. T. Rolls, L. Franco, and S. M. Stringer

University of Oxford, UK

To analyse the functions of the perirhinal cortex, the activity of single neurons in the perirhinal cortex was recorded while macaques performed a delayed matching-to-sample task with up to three intervening stimuli. Some neurons had activity related to working memory, in that they responded more to the sample than to the match image within a trial, as shown previously. However, when a novel set of stimuli was introduced, the neuronal responses were on average only 47% of the magnitude of the responses to the set of very familiar stimuli. Moreover, it was shown in three monkeys that the responses of the perirhinal cortex neurons gradually increased over hundreds of presentations (mean = 400 over 7–13 days) of the new set of (initially novel) stimuli to become as large as those to the already familiar stimuli. Thus perirhinal cortex neurons represent the very long-term familiarity of visual stimuli. Part of the impairment in temporal lobe amnesia may be related to the difficulty of building representations of the degree of familiarity of stimuli. A neural network model of how the perirhinal cortex could implement long-term familiarity memory is proposed using Hebbian associative learning.

In this paper, we produce a new model of how the perirhinal cortex implements a long-term form of familiarity memory. In the Introduction, we first describe the neurophysiological data that implicate the perirhinal cortex in long-term familiarity memory and set out what needs to be modelled. Then we compare this type of memory to other types of memory in which the perirhinal cortex is implicated. These other types of memory include, as described below, recognition memory as measured in delayed match-to-sample tasks with short time delays, in which perirhinal cortex neurons typically respond more to the sample than the match stimulus, and delayed match-to-sample tasks with intervening stimuli in which perirhinal cortex neurons may respond more to the match stimulus than the sample; and paired associate learning.

Evidence that the perirhinal cortex is involved in long-term familiarity memory comes from a neuronal recording study in which it was shown that perirhinal cortex neuronal responses in the rhesus macaque gradually increase in magnitude to a set of stimuli as that set is repeated for 400 presentations each 1.3 s long (Hölscher, Rolls, & Xiang, 2003). The

Correspondence should be addressed to Professor Edmund T. Rolls, University of Oxford, Department of Experimental Psychology, South Parks Road, Oxford OX1 3UD, UK. Email: Edmund.Rolls@psy.ox.ac.uk

This research was supported by MRC Programme Grant PG9826105 and by the Medical Research Council Interdisciplinary Research Centre for Cognitive Neuroscience.

single neurons were recorded in the perirhinal cortex in monkeys performing a delayed matching-to-sample task with up to three intervening stimuli, using a set of very familiar visual stimuli used for several weeks. When a novel set of stimuli was introduced, the neuronal responses were on average only 47% of the magnitude of the responses to the familiar set of stimuli. It was shown in eight different replications in three monkeys that the responses of the perirhinal cortex neurons gradually increased over hundreds of presentations of the new set of (initially novel) stimuli to become as large as those to the already familiar stimuli. The mean number of 1.3 s presentations to induce this effect was 400 occurring over 7–13 days. These results show that perirhinal cortex neurons represent the very long-term familiarity of visual stimuli. A representation of the long-term familiarity of visual stimuli may be important for many aspects of social and other behaviour, and part of the impairment in temporal lobe amnesia may be related to the difficulty of building representations of the degree of familiarity of stimuli. It is this type of memory that is modelled in this paper.

The perirhinal cortex is also involved in recognition memory in that damage to the perirhinal cortex produces impairments in recognition memory tasks in which several items intervene between the sample presentation of a stimulus and its presentation again as a match stimulus (Malkova, Bachevalier, Mishkin, & Saunders, 2001; Zola-Morgan, Squire, Amaral, & Suzuki, 1989; Zola-Morgan, Squire, & Ramus, 1994). Indeed, damage to the perirhinal cortex rather than to the hippocampus is believed to underlie the impairment in recognition memory found in amnesia in humans associated with medial temporal lobe damage (Buckley, Booth, Rolls, & Gaffan, 2001; Buckley & Gaffan, 2000). Neurophysiologically, it has been shown that many inferior temporal cortex (a term we use to refer to area TE) neurons (Rolls, 2000; Rolls & Deco, 2002), which provide visual inputs to the perirhinal cortex (Suzuki & Amaral, 1994a, 1994b), respond more to the first than to the second presentation of a stimulus in a running recognition task with trial-unique stimuli (Baylis & Rolls, 1987). In this task, there is typically a presentation of a novel stimulus, and after a delay, which may be in the order of minutes or more and in which other stimuli may be shown, the stimulus is presented again as “familiar”, and the monkey can respond to obtain food reward. Most neurons responded more to the “novel” than to the “familiar” presentation of a stimulus, where “familiar” in this task reflects a change produced by seeing the stimulus typically once (or a few times) before. (A small proportion of neurons respond more to the familiar (second) than to the novel (first) presentation of each visual stimulus.) In the inferior temporal cortex this memory spanned up to 1–2 intervening stimuli between the first (novel) and second (familiar) presentations of a given stimulus (Baylis & Rolls, 1987), and as recordings are made more ventrally, towards and within the perirhinal cortex, the memory span increases to several or more intervening stimuli (Brown & Xiang, 1998; Wilson, Riches, & Brown, 1990; Xiang & Brown, 1998).

In a similar task, though, typically performed with non-trial-unique stimuli—a delayed matching-to-sample task with up to several intervening stimuli—some neurons respond more to the match stimulus than to the sample stimulus (Miller, Li, & Desimone, 1998). Many neurons in this task respond more to the sample (“novel”) than to the match (“familiar”) presentations of the stimuli, and this short-term memory is reset at the start of the next trial (Hölscher & Rolls, 2002). The resetting at the start of each trial shows that the perirhinal cortex is actively involved in the task demands.

A fourth type of memory in which the perirhinal cortex is implicated is paired associate learning (a model of semantic long-term memory), which is represented by a population of neurons in a restricted part of area 36 where the neuronal responses may occur to both members of a pair of pictures used in the paired association task (Miyashita, Kameyama, Hasegawa, & Fukushima, 1998; Miyashita, Okuno, Tokuyama, Ihara, & Nakajima, 1996).

We emphasize that the type of familiarity memory modelled here is a very long-term type of familiarity or recognition memory, which reflects the gradual build-up of neuronal responses over several hundred presentations of a stimulus, and which may thus represent the degree of long-term familiarity of stimuli. We call this long-term familiarity memory in this paper. Figure 1 shows some of the neurophysiological data on long-term familiarity memory from Hölischer et al. (2003) that we wish to model. Figure 1 shows the results of three different replications of the whole investigation in macaque BL. Each replication consisted of starting with a completely new set of “novel” images and using this set for 10 days of testing, in which experiments were performed on many different neurons. The ordinate shows the mean response of a neuron to the set of stimuli in the novel set expressed as a percentage of the response to the set of stimuli in the familiar set. In each replication of the overall investigation, for many neurons early on after the novel set of stimuli was introduced, there were highly significant differences between the mean responses of the neurons to the set of familiar and novel stimuli, as shown by nonparametric (Mann–Whitney U) tests. Indeed, for many cells the difference between the responses to novel and familiar stimuli on the days soon after the novel stimuli were introduced were significant at $p < 10^{-5}$. For the first replication, the degree of variation is indicated by the standard errors of the mean responses of each cell. The slope of the (linear) regression line for each replication (BL1–BL3) was calculated and was highly significant. The intercept of the regression line indicates the average percentage of the neuronal response to novel stimuli compared to very familiar stimuli at the start of testing with novel stimuli. The regression lines show how long it takes neurons to take to respond to the novel set of images as well as to the highly familiar set, shown for hundreds of previous trials so that their maximal response level had been reached. Similar results were found in five further replications in two further monkeys.

Overall, it was found that stimulus selectivity was less than that in the inferior temporal cortex (Rolls, 2000; Rolls & Deco, 2002). This was even confirmed by direct comparison in one of the macaques used in the study, in which when the macaque performed the same task, 15/22 (68%) of inferior temporal cortex (IT) neurons had selective responses to the same set of stimuli. Moreover, the selectivity of the inferior temporal cortex neurons was greater than that of the neurons in the perirhinal cortex (that is, the sparseness of the representation for inferior temporal cortex neurons is lower, see Rolls, 2000; Rolls & Deco 2002). Related perhaps partly to the specificity of the responses of many inferior temporal cortex neurons, in the experiments described here there was no difference apparent in the responses of IT neurons to the novel and the long-term familiar stimuli used. This means at least that long-term familiarity is not made explicit in the responses of inferior temporal cortex neurons, whereas it is made explicit in the responses of neurons in the perirhinal cortex. By “made explicit”, we mean that it can be easily read off from the firing rates of one or a small number of neurons, by, for example, dot product decoding (Rolls & Deco, 2002; Rolls & Treves, 1998; Rolls, Treves, & Tovee, 1997a; Rolls, Treves, Tovee, & Panzeri, 1997b).

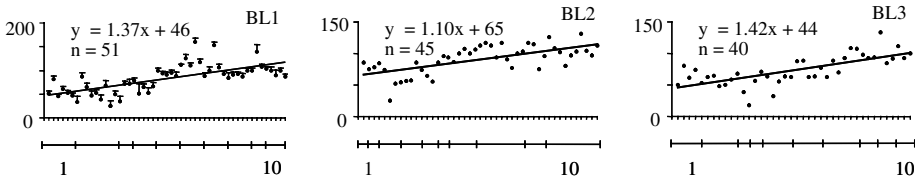


Figure 1. Regressions showing the relative response of each neuron to novel vs. familiar stimuli (expressed as a percentage) as a function of the number of experiments since the novel set of stimuli were investigated (abscissa), for three complete replications (BL1, BL2, and BL3) in macaque BL. Each point on the graph shows the results of one experiment involving 80 trials of the delayed matching-to-sample task on one neuron. The 80 trials included 40 with the novel stimulus set and 40 with the familiar stimulus set, with 5 stimuli on each trial. On some days more than one neuron was analysed in a separate experiment, and the number of days since introduction of the novel set of stimuli is also shown on the abscissa. Results for three separate replications of the whole investigation in one monkey (BL) are shown. Each replication involved starting with a completely novel set of images, and using that novel set on 10 days of testing in which on any day as many experiments as possible were performed, each experiment with a different neuron, and each experiment involving 40 trials with the novel set and 40 trials with the familiar set of images. The first replication (left) involved recordings in 51 experiments from 51 separate neurons over 10 testing days. The slope and intercept of the regression line are shown. The intercept indicates the magnitude of the response to novel stimuli expressed as a percentage of that to familiar stimuli at the start of the replication. During an experiment on each neuron, the set of novel stimuli was shown for approximately 12.5 of the 1.3 s presentations of each novel stimulus during the delayed matching-to-sample task. The results for replication BL1 show the standard error of the mean response of the neuron to the novel relative to the familiar stimuli to give an indication of the degree of accuracy with which this could be estimated. The error bars are omitted from the other replications for clarity.

The model

To analyse how neurons in a receiving brain area such as the perirhinal cortex might build long-term familiarity representations, let us consider the model shown in Figure. 2. This consists of a set of perirhinal cortex neurons y_i receiving synaptic connections w_{ij} from a set of neurons x_j (such as those in the inferior temporal visual cortex), which are tuned to respond to different stimuli and which do not show long-term familiarity-related responses.

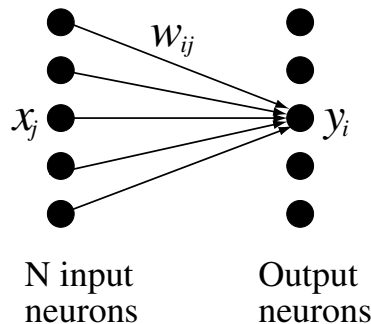


Figure 2. The network model for the formation of neurons with responses related to the long-term familiarity of visual stimuli. This consists of a set of perirhinal cortex neurons y_j receiving synaptic connections w_{ij} from a set of neurons x_j (such as those in the inferior temporal visual cortex) which are tuned to respond to different stimuli and which do not show long-term familiarity-related responses.

We first consider analytically cases under which it would be possible to observe a linear increase in the response of the cells y_i to the familiar set of patterns as they are presented very many times (of order 400) to the network. If we consider a case of neurons with a linear transfer function and a Hebb (associative) learning rule, there will be an exponential growth of the synaptic weights and correspondingly of the firing rates of the output neurons to the familiar stimuli. The activation h_i (sometimes termed the net input to a neuron, and corresponding to the depolarization of a neuron produced by its inputs) of neuron i is

$$h_i = \sum_{j=1}^N w_{ij} x_j \quad 1$$

This can be read as indicating that the depolarization h_i of the neuron is the sum of the activations produced through each synapse, which depend on the presynaptic firing rate x_j and the synaptic strength w_{ij} .

For a linear activation function $y_i = h_i$ relating the firing rate y_i of neuron i to its depolarization or activation h_i , the output, y_i^p , of a neuron i after p presentations of the familiar set of stimuli can be calculated to be

$$y_i^p = \sum_{j=1}^N w_{ij}^{p-1} x_j \quad 2$$

where w_{ij}^{p-1} are the values of the synaptic weights after the $(p-1)$ th presentation, and x_j is the firing rate of input neuron j .

For binary (i.e., 0,1) input patterns with sparseness a (where the sparseness can be thought of as the proportion of the input neurons that respond to any one pattern and thus takes the value 0.2 if 20% of the neurons are active for a stimulus, and 80% are not firing, see Rolls & Deco, 2002; Rolls & Treves, 1998), and for the case of N input neurons we obtain

$$y_i^p = aNw_{ij}^{p-1} \quad 3$$

If the weights w_{ij} are updated according to a Hebbian associative learning rule in which the change of weight depends on the product of the presynaptic firing rate x_j and the postsynaptic firing rate y_i shown below in Equation 4,

$$\delta w_{ij} = ky_i x_j \quad 4$$

where k is the learning rate, the change in weights at the p th presentation can be calculated to be

$$\delta w_{ij}^p = k(aNw_{ij}^{p-1}) \quad 5$$

for active input neuron x_j (using substitution of Equation 3 into Equation 4).

It can then be shown that

$$w_{ij}^p = (1 + kaN)^p w_{ij}^0 \quad 6$$

and since $(1 + kaN) > 1$, the weights grow exponentially in this case of using a linear transfer function and an associative learning rule; w_{ij}^0 is the initial small synaptic weight, w_{ij}^p is the

weight after the p th presentation, and $(1 + kaN)$ to the power p arises because on every one of the p learning trials, as the synaptic weight increases, it contributes to its own further increase on every trial by influencing the output firing rate on that trial.

Simulation results

We ran simulations using a network comprising 1,000 input neurons projecting to a set of output neurons as shown in Figure 2. The input to the network consisted of random binary patterns that were divided into two sets: the novel and the familiar stimulus sets. In different simulations we tested sparsenesses of the (novel and familiar) input patterns of 0.1 and 0.01.

The growth of the weights summarized in Equation 6 when the network was trained with a Hebb rule and a linear activation function was found in the simulations, and it produced an exponential increase in the firing rate to the familiar patterns, as illustrated in Figure 3. This scenario, therefore, does not account for the approximately linear increase in the firing rate found as a function of the number of training trials, which is illustrated for the real neuronal firing in Figure 1.

Two possible solutions are to use a nonlinear saturating activation function, such as a sigmoid function (which would limit the maximum firing of the output neuron), or to use an associative synaptic modification rule with a decrementing term that captures effects of long-term synaptic depression (LTD). We investigated both.

The results for the case of a sigmoid activation function relating the firing rate y_i of neuron i to its depolarization or activation h_i (shown in Equation 7 where alpha is the threshold, and beta is the slope) are shown in Figure 4. The cell firing rate grows approximately linearly up to its maximum value after approximately 400 repeated presentations.

$$y = \sigma(h) = \frac{1}{1 + e^{-(\beta h + \alpha)}} \quad 7$$

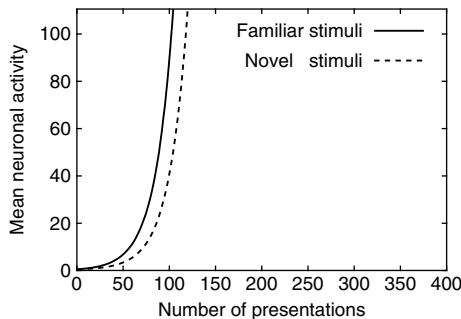


Figure 3. An exponential increase in the firing rate to the familiar patterns as a function of the number of training trials was produced when the network was trained with a Hebb rule and a linear activation function. The familiar set of stimuli was initially novel, and the solid curve shows how the response to this set of stimuli gradually increases as the set becomes familiar over approximately 400 training (i.e., weight modification) presentations. For comparison, the responses to novel stimuli, which were not trained and remained novel, are also shown (dotted curve).

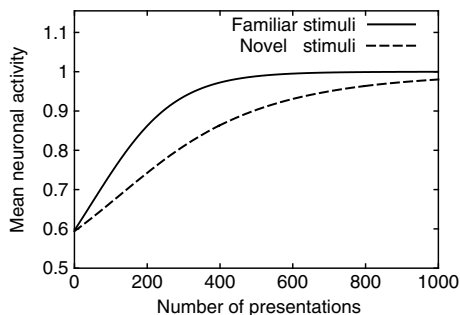


Figure 4. An approximately linear increase in the firing rate to the familiar patterns as a function of the number of training trials was produced when the network was trained with a Hebb rule and a sigmoid activation function (in which $\beta = 15$ and $\alpha = 0.7$). The firing rate saturated at a high value after approximately 500 training trials, and this is a function of the learning rate parameter k in Equation 4. The familiar set of stimuli was initially novel, and the solid curve shows how the response to this set of stimuli gradually increases as the set becomes familiar over approximately 400 training (i.e., weight modification) presentations. For comparison, the responses to novel stimuli, which were not trained and remained novel, are also shown (dotted curve).

We also explored the case of using a modified Hebb rule of the type useful for competitive networks (see Rolls & Deco, 2002; and Rolls & Treves, 1998) in which the synaptic modification, describing heterosynaptic long-term depression in which the synaptic weight can decrease (in proportion to the postsynaptic neuron firing y_i) if the presynaptic firing x_j is lower than the value of the synaptic weight w_{ij} , is:

$$\delta w_{ij} = k(x_j - w_{ij})y_i \quad 8$$

This allows active output neurons to increase their synaptic weight if the input x_j is above the existing weight and to decrease the synaptic weight if the input is below the existing weight. (The fact that the weight is subtracted from the presynaptic firing rate in Equation 8 captures the fact that LTD is easier to obtain after long-term potentiation [LTP] has been produced—see Rolls & Deco, 2002.) The weights are clipped to be non-negative. The simulation results using this learning rule and the linear activation function are presented in Figure 5a. After an initial period of learning, there is an almost linear increase in the firing as a function of the number of training trials in the range 0–400, with saturation of the mean neuronal activity to familiar stimuli occurring after that. The response to the novel stimuli remains low. The sparseness was 0.01.

In Figure 5b we show that the model still works well at a higher loading of 0.25. In Figure 5c we show that the model works well with a sigmoid activation function, in which $\beta = 15$ and $\alpha = 0.7$, so that the success of the model is especially related to the modified learning rule, not to the particular activation function chosen.

We then analysed the capacity of the network as the number of patterns stored in the network grows. The capacity in which we are interested is the capacity to respond differently to familiar stimuli after many training trials than to novel stimuli. The results are presented in Figure 6. The training rule shown in Equation 8 was used with a linear activation function. With a loading of 1.0, N patterns are trained when there are N synapses per neuron. (A loading of 1.0 corresponds to the case of training the network with 1,000 random familiar patterns

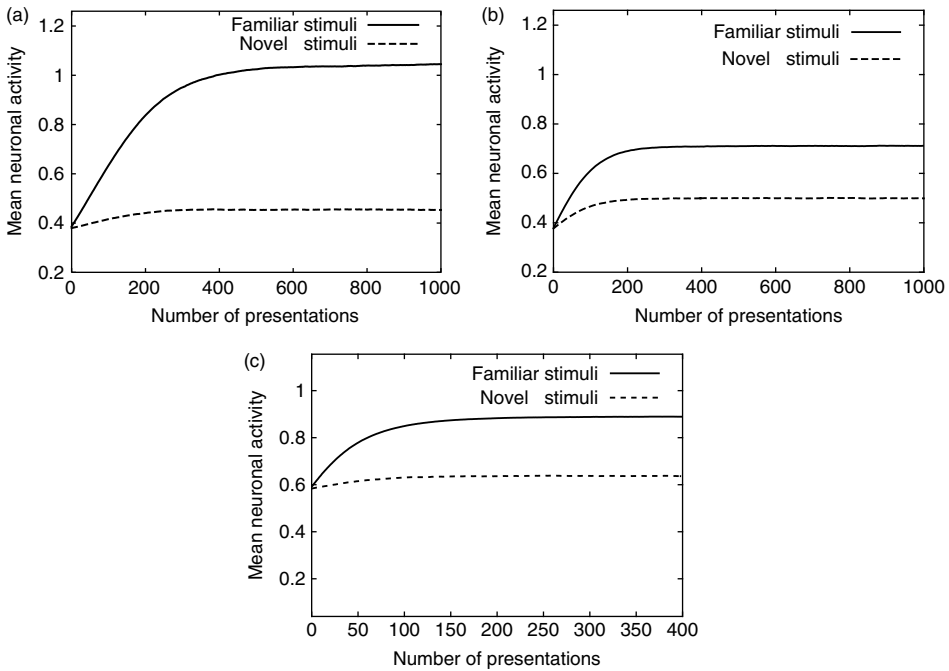


Figure 5. a. (Top left panel) An approximately linear increase in the firing rate to the familiar patterns as a function of the number of training trials was produced when the network was trained with the learning rule shown in Equation 8 and a linear activation function. The firing rate saturated at a high value after approximately 500 training trials, and when this occurs is a function of the learning rate parameter k in Equation 8. The loading was 0.1, and the sparseness a was 0.01. The familiar set of stimuli was initially novel, and the solid curve shows how the response to this set of stimuli gradually increases as the set become familiar over approximately 400 training (i.e., weight modification) presentations. For comparison, the responses to novel stimuli, which were not trained and remained novel, are also shown (dotted curve). b. (Top right panel) The same simulation as that in Figure 5a, but with a higher loading of 0.25. c. (Bottom panel) The same simulation as that in Figure 5a, but with a sigmoid activation function, in which $\beta = 15$ and $\alpha = 0.7$.

and testing with 1,000 novel random patterns, when each neuron in the network has 1,000 inputs.) The results are shown in Figure 6 for sparseness values of 0.01 and 0.1. (The random initial synaptic weights were initialized to the mean asymptotic value.) Given that the actual neurons in the perirhinal cortex respond to novel stimuli with approximately 50% of the firing rate response to familiar stimuli (see Figure 1 and Hölscher et al., 2003), the model operates within approximately the correct area for a sparseness of the patterns of 0.01 up to a loading of approximately 1. With a sparseness of 0.1 of the binary patterns, the network can be loaded less, up to a loading value of approximately 0.1.

Discussion

The analysis and model show that neurons that gradually increase their response approximately linearly over 400 or more presentations of a stimulus cannot be accounted for by a Hebb rule with a linear activation function. If a sigmoid activation function is used with the

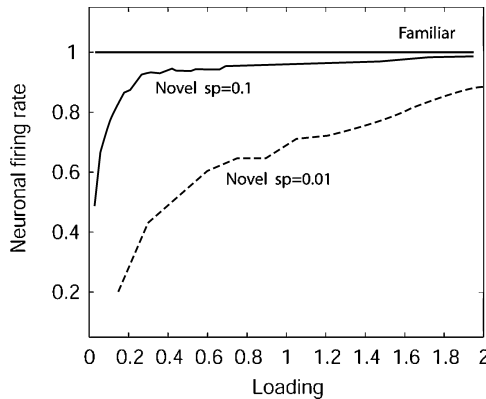


Figure 6. The operation of the model as a function of the number of patterns trained. The training rule shown in Equation 8 was used with a linear activation function. With a loading of 1.0, N patterns are trained when there are N synapses per neuron. (A loading of 1.0 corresponds to the case of training the network with 1,000 familiar patterns and testing with 1,000 novel random patterns, when each neuron in the network has 1,000 inputs.) The results are shown in sparsenesses of the patterns of 0.01 and 0.1.

same associative learning rule, then the limited maximal value of the firing rate results in only a limited maximum weight change on any one trial, due to y_i having a fixed maximum value in Equation 4. With a relatively low learning rate, gradually more and more neurons have synaptic weight increases that result in the neurons reaching their maximal firing rate, as shown after 400 trials in Figure 4a. After that point in the learning, although the firing rates are at a maximal value to the familiar stimuli, the synaptic weights will continue to increase without bound as further learning trials of the familiar stimuli occur. Unbounded synaptic weights are a well-known problem with a simple associative Hebb learning rule. In this situation, to the extent that the random novel patterns overlap with the familiar patterns, the neuronal responses to the novel patterns will continue to increase after Trial 400.

A more attractive model, because the synaptic weights are self-limiting, is therefore a model with heterosynaptic long-term depression of the type shown in Equation 8 in the learning rule. In this case, as illustrated in Figure 4b, the firing rates to the familiar stimuli gradually saturate to a fixed high value at which the synaptic weights to the familiar stimuli are bounded to within a range set by the presynaptic firing rate, as shown in Equation 8 (the term $x_j - w_{ij}$). At the same time, the firing rates to the novel stimuli remain relatively low, because the values of the synapses that are activated by the novel stimuli tend to remain low as a result of the heterosynaptic long-term depression. In this case, the activation function can be linear, threshold linear, or a sigmoid activation function that approximates a threshold linear activation function. Because the modified Hebb rule shown in Equation 8 is self-limiting, and simulations with it can approximate the neuronal data showing long-term familiarity effects in the perirhinal cortex, this simple model is what we propose could implement the neurophysiological results.

With respect to the model, it is also the case that sparse input representations (in each of which few neurons have high firing rates) will allow larger numbers of stimuli to be in the familiar and novel sets, as shown in Figure 6. Indeed, the loading effects shown in Figure 6

show that the number of patterns in the novel and familiar sets that can be discriminated is in the order of the number of synapses per neuron for sparse representations, as is typical for associative networks (Rolls & Treves, 1990, 1998). Further, we note that the capacities for the network scale with the number of synapses per neuron. Further, if there is sparse connectivity from the set of input neurons x_j to the output neurons y_i , it will be possible to load the whole system to higher capacities, because in effect some patterns will tend to be learned more by some neurons, and other patterns more by other neurons. Indeed, this effect could account for why many neurons, rather than just one or very few neurons, are present in the perirhinal cortex. (These neurons are the y neurons in the model.)

We do not know of other models of how long-term familiarity-related neuronal responses could develop in the perirhinal cortex. Bogacz, Brown, and Giraud-Carrier (2001) modelled how, in what is effectively a short-term memory test, the responses to a novel visual stimulus can be large the first time that the stimulus is shown, but small the second time that the same stimulus is shown. That model thus is not of familiarity-related neurons, in that the neurons in that model respond more to novel stimuli. Also, that model is of a short-term memory process that occurs in one trial. (We note incidentally that that model cannot by itself account for the active resetting of perirhinal cortex neurons in a delayed match-to-sample short-term memory task, in that the responses to the sample stimulus are larger than those to the following match stimulus, even if the sample has been seen very recently of the immediately preceding trial, Hölscher & Rolls, 2002.) In contrast, the model described here is of how perirhinal cortex neurons can gradually increase their responses to stimuli as they become familiar over several hundred presentations. McLaren and Mackintosh (2002) discuss how a number of phenomena in associative learning might arise in systems with distributed representations. We note that the modified Hebb rule we use in Equation 8 (used by Willshaw & von der Malsburg, 1976, and developed by Oja, 1982) has self-limiting properties similar to those of an associative learning rule used by Vogel, Brandon, and Wagner (2003).

With respect to the neurophysiological discovery of perirhinal cortex neurons with activity related to the long-term familiarity of stimuli, with the familiarity building over hundreds of trials (Hölscher et al., 2003), we note that such a time scale has not been investigated previously, since most studies did not record neuronal responses for longer than 24 h and did not allow a slow emergence of increased neuronal responses to images related to their long-term familiarity to be observed (Brown & Xiang, 1998; Erickson & Desimone, 1999; Xiang & Brown, 1998). Although some neurons responding more to familiar than to novel images, and others responding more to novel than familiar images, have been found previously in the perirhinal cortex where novelty and familiarity refer to changes that occur over a few stimulus presentations (Brown & Xiang, 1998; Hölscher & Rolls, 2002; Sobotka & Ringo, 1993), the neurons described by Hölscher et al. (2003) had responses related to a different, long-term type of familiarity in which the increased neuronal responses to familiar images can take days or weeks to develop.

The perirhinal cortex is well placed to form such long-term familiarity representations, because it receives stimulus-selective information about what object is being viewed from the inferior temporal visual cortex. Indeed, it has been shown in a tracer study that the anteroventral part of area TE (TEav) projects diffusely over a wide extent of perirhinal cortex (Saleem & Tanaka, 1996), making the perirhinal cortex anatomically suited for making

associations between features in an object (Bussey, Saksida, & Murray, this issue) or the objects in a scene, or computing a general property of all inputs being received, such as how familiar they are. The perirhinal cortex can thus use the stimulus-selective input from potentially most parts of area TE of the inferior temporal visual cortex to form new associations between such selective inputs and thereby to form a unique representation of complex stimuli to identify familiar objects or scenes.

Xiang and Brown (1998) found that in a recognition memory task with trial-unique visual stimuli, the perirhinal cortex neurons tended to respond more to the first than to the second presentation of the stimuli. With stimuli that were already familiar, even this difference was not clear. This may be because the neurons respond at their highest rates to stimuli that are highly familiar, allowing less room for a smaller response to a match than to a nonmatch stimulus in a recognition memory task.

What advantages might the representation of the long-term familiarity of images, a model for which is reported for the first time in this paper, confer? The potential functions of computing the long-term familiarity of objects or images in the brain are many-fold, and include recognition of complex object-environment configurations such as members of one's own social and family group, recognition of one's own possessions, recognition of one's own territory, and so on. Further, it is notable that the loss of the feeling of familiarity for objects and events introduced after medial temporal lobe damage is one of the important symptoms of medial temporal lobe amnesia (Squire, Stark, & Clark, 2004), and this too may be related to these operations that we suggest are being performed by the perirhinal cortex (Henson, this issue; Holdstock, this issue). What we propose is that the identity of the object or face would be represented by area TE of the inferior temporal visual cortex (in the way reviewed by Rolls & Deco, 2002) and that the long-term familiarity of the object would be represented by how strongly the perirhinal cortex neurons analysed here are firing. Together, the two types of neuronal activity encode both identity and long-term familiarity, but allow each type of information to be read out from the system (by other brain areas) independently of the other.

REFERENCES

- Baylis, G., & Rolls, E. T. (1987). Responses of neurons in the inferior temporal cortex in short term and serial recognition memory tasks. *Experimental Brain Research*, *65*, 614–622.
- Bogacz, R., Brown, M. W., & Giraud-Carrier, C. (2001). Model of familiarity discrimination in the perirhinal cortex. *Journal of Computational Neuroscience*, *10*, 5–23.
- Brown, M., & Xiang, J. (1998). Recognition memory: Neuronal substrates of the judgement of prior occurrence. *Progress in Neurobiology*, *55*, 149–189.
- Buckley, M. J., Booth, M. C. A., Rolls, E. T., & Gaffan, D. (2001). Selective perceptual impairments following perirhinal cortex ablation. *Journal of Neuroscience*, *21*, 9824–9836.
- Buckley, M. J., & Gaffan, D. (2000). The hippocampus, perirhinal cortex and memory in the monkey. In J. J. Bolhuis (Ed.), *Brain, perception, memory* (pp. 279–298). Oxford, UK: Oxford University Press.
- Bussey, T. J., Saksida, L. M., & Murray, E. A., (this issue). The perceptual-mnemonic/feature conjunction model of perirhinal cortex function. *Quarterly Journal of Experimental Psychology*, *58B*, 269–282.
- Erickson, C., & Desimone, R. (1999). Responses of macaque perirhinal neurons during and after visual stimulus association learning. *Journal of Neuroscience*, *19*, 10404–10416.
- Henson, R. (this issue). A mini-review of fMRI studies of human medial temporal lobe activity associated with recognition memory. *Quarterly Journal of Experimental Psychology*, *58B*, 340–360.

- Holdstock, J. S. (this issue). The role of the human medial temporal lobe in object recognition and object discrimination. *Quarterly Journal of Experimental Psychology*, 58B, 326–339.
- Hölscher, C., & Rolls, E. T. (2002). Perirhinal cortex neuronal activity is actively related to working memory in the macaque. *Neural Plasticity*, 9, 41–51.
- Hölscher, C., Rolls, E. T., & Xiang, J.-Z. (2003). Perirhinal cortex neuronal activity related to long-term familiarity memory in the macaque. *European Journal of Neuroscience*, 18, 2037–2046.
- Malkova, L., Bachevalier, J., Mishkin, M., & Saunders, R. C. (2001). Neurotoxic lesions of perirhinal cortex impair visual recognition memory in rhesus monkeys. *Neuroreport*, 12, 1913–1917.
- McLaren, I. P., & Mackintosh, N. J. (2002). Associative learning and elemental representation: II. Generalization and discrimination. *Animal Learning & Behavior*, 30, 177–200.
- Miller, E.K., Li, L., & Desimone, R. (1998). Activity of neurons in anterior inferior temporal cortex during a short-term memory task. *Journal of Neuroscience*, 13, 1460–1478.
- Miyashita, Y., Kameyama, M., Hasegawa, I., & Fukushima, T. (1998). Consolidation of visual associative long-term memory in the temporal cortex of primates. *Neurobiology of Learning and Memory*, 1, 197–211.
- Miyashita, Y., Okuno, H., Tokuyama, W., Ihara, T., & Nakajima, K. (1996). Feedback signal from medial temporal lobe mediates visual associative mnemonic codes of inferotemporal neurons. *Cognitive Brain Research*, 5, 81–86.
- Oja, E. (1982). A simplified neuron model as a principal component analyzer. *Journal of Mathematical Biology*, 15, 267–73.
- Rolls, E. T. (2000). Functions of the primate temporal lobe cortical visual areas in invariant visual object and face recognition. *Neuron*, 27, 205–218.
- Rolls, E. T., & Deco, G. (2002). *Computational neuroscience of vision*. Oxford, UK: Oxford University Press.
- Rolls, E. T., & Treves, A. (1990). The relative advantages of sparse versus distributed encoding for associative neuronal networks in the brain. *Network*, 1, 407–421.
- Rolls, E. T., & Treves, A. (1998). *Neural networks and brain function*. Oxford, UK: Oxford University Press.
- Rolls, E. T., Treves, A., & Tovee, M. J. (1997a). The representational capacity of the distributed encoding of information provided by populations of neurons in the primate temporal visual cortex. *Experimental Brain Research*, 114, 149–162.
- Rolls, E. T., Treves, A., Tovee, M., & Panzeri, S. (1997b). Information in the neuronal representation of individual stimuli in the primate temporal visual cortex. *Journal Computational Neuroscience*, 4, 309–333.
- Saleem, K. S., & Tanaka, K. (1996). Divergent projections from the anterior inferotemporal area TE to the perirhinal and entorhinal cortices in the macaque monkey. *Journal of Neuroscience*, 16, 4757–4775.
- Sobotka, S., & Ringo, J. (1993). Investigation of long term recognition and association memory in unit responses from inferotemporal cortex. *Experimental Brain Research*, 96, 28–38.
- Squire, L. R., Stark, C. E., & Clark, R. E. (2004). The medial temporal lobe. *Annual Review of Neuroscience*, 27, 279–306.
- Suzuki, W., & Amaral, D. (1994a). Perirhinal and parahippocampal cortices of the macaque monkey: Cortical afferents. *Journal of Comparative Neurology*, 350, 497–533.
- Suzuki, W., & Amaral, D. (1994b). Topographic organization of the reciprocal connections between the monkey entorhinal cortex and the perirhinal and parahippocampal cortices. *Journal of Neuroscience*, 14, 1856–1877.
- Vogel, E. H., Brandon, S. E., & Wagner, A. R. (2003). Stimulus representation in SOP: II. An application to inhibition of delay. *Behavioral Processes*, 62, 27–48.
- Willshaw, D. J., & von der Malsburg C. (1976). How patterned neural connections can be set up by self-organization. *Proceedings of the Royal Society of London. Biological Sciences*, 194, 431–45.
- Wilson, F. A., Riches, I. P., & Brown, M. W. (1990). Hippocampus and medial temporal cortex: Neuronal activity related to behavioural responses during the performance of memory tasks by primates. *Behavioral Brain Research*, 40, 7–28.
- Xiang, J., & Brown, M. (1998). Differential neuronal encoding of novelty, familiarity and recency in regions of the anterior temporal lobe. *Neuropharmacology*, 37, 657–676.
- Zola-Morgan, S., 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.
- Zola-Morgan, S., Squire, L. R., & Ramus, S. J. (1994). Severity of memory impairment in monkeys as a function of locus and extent of damage within the medial temporal lobe memory system. *Hippocampus*, 4, 483–494.