articles

Inter-trial neuronal activity in inferior temporal cortex: a putative vehicle to generate long-term visual associations

Volodya Yakovlev[1], Stefano Fusi[2,5], Elisha Berman[3] and Ehud Zohary[1,4]

[1] Department of Neurobiology, Institute of Life Science, Hebrew University, Jerusalem, 91904, Israel

[2] Racah Institute of Physics, Hebrew University, Jerusalem, 91904, Israel

[3] Human Biology Research Center, Hadassah University Hospital, Jerusalem 91904, Israel

[4] Center for Neural Computation, Hebrew University, Jerusalem, 91904, Israel

[5] INFN, Universit`a di Roma, La Sapienza, Italy

Correspondence should be addressed to E.Z. (udiz@lobster.huji.ac.il)

When monkeys perform a delayed match-to-sample task, some neurons in the anterior inferotemporal cortex show sustained activity following the presentation of specific visual stimuli, typically only those that are shown repeatedly. When sample stimuli are shown in a fixed temporal order, the few images that evoke delay activity in a given neuron are often neighboring stimuli in the sequence, suggesting that this delay activity may be the neural correlate of associative long-term memory. Here we report that stimulus-selective sustained activity is also evident following the presentation of the test stimulus in the same task. We use a neural network model to demonstrate that persistent stimulus-selective activity across the intertrial interval can lead to similar mnemonic representations (distributions of delay activity across the neural population) for neighboring visual stimuli. Thus, inferotemporal cortex may contain neural machinery for generating long-term stimulus–stimulus associations.
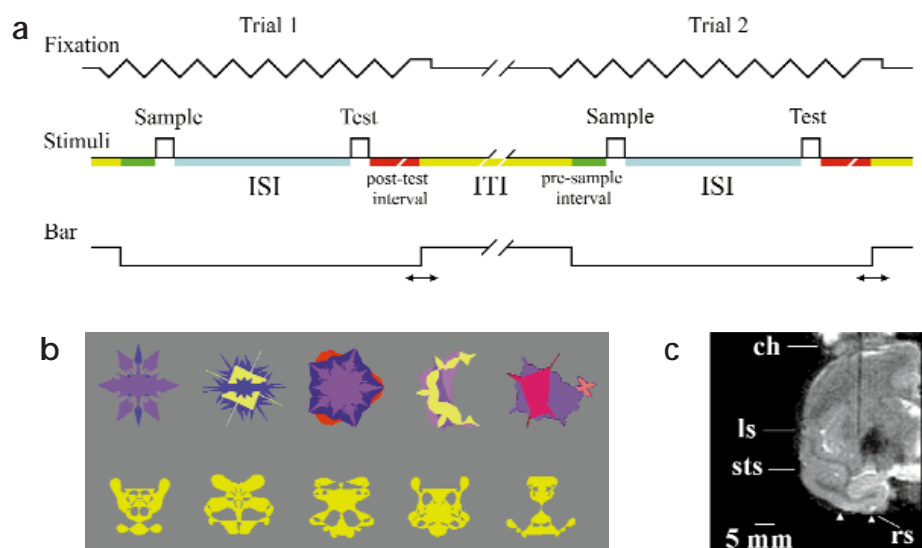
Most of us can remember the next melody on an album once the current tune is over and recall the alphabet in its correct sequential order. These are classic examples of generating an association between stimuli that have been presented in a fixed temporal order. A visual example of such a phenomenon may be the way we navigate an unfamiliar environment. In such circumstances, we typically use remembered snapshots of visual scenes to verify that we are on the correct track, with one serving as a cue to lead us to the next expected landmark. Our current understanding of the neuronal basis for the formation of such long-term associative memories is only marginal.

Miyashita and colleagues[1] addressed this question in a recent study. Monkeys were trained on a delayed match-to-sample task, in which they remember the identity of a sample stimulus during a delay interval and indicate whether the following test stimulus is ('match') or is not ('non-match') identical to the sample stimulus. The novel feature in the experimental design was that the sample stimuli were presented in a fixed temporal order. Single neurons in inferior temporal (IT) cortex were recorded during the delay period between the presentation of the sample and test stimuli. Some IT neurons had increased firing rates throughout the delay interval, long after sample-stimulus presentation, as previously reported in IT and prefrontal cortex[2–6]. Although the monkey could perform the task with novel stimuli, only highly familiar stimuli evoked this delay activity. The few visual stimuli that generated delay activity in the same IT neuron were usually nearest neighbors in the fixed temporal sequence during the training period, even though the order of the sample stimuli was totally irrelevant to task performance. This aspect of the neuronal response led the authors to suggest[1] that "the selectivity acquired by these cells represents a neuronal correlate of associative long-term memory of pictures."

Based on this observation, a comprehensive theoretical framework for understanding the development of associative long-term memory was proposed[7,8]. According to this approach, the sustained delay activity is a feature of the pattern of connectivity between neurons, rather than of a single neuron. The persistent delay activity is maintained by recurrent synaptic feedback between interconnected neurons within a local module, built up as stimuli become familiar. The memory process is initiated by presentation of the visual stimulus, which generates a pattern of response across the neuronal population. Following removal of the visual stimulus, because of the feedback connections within the neuronal population, the dynamics of the network is such that it settles into a stable state (the attractor), in which most neurons are firing at their spontaneous level, but some distinct neurons continue firing at elevated levels even though the visual stimulus is no longer present. The stable state implies that this pattern of firing continues until a new afferent input (from a new, effective visual stimulus) changes the state of the network components. Because each visual stimulus evokes a characteristic pattern of delay activity, the delay-activity distribution is the neuronal engram of the last familiar stimulus seen. The distributed nature of the representation allows storage of a large number of patterns (stable delay-activity distributions) in the same neural module, by the same synaptic structure.

Fig. 1. Experimental task, visual stimuli and cortical area explored in this study. **(a)** Schematic sequence of events in the delayed match-to-sample task. A trial began with the presentation of a flickering dot (at 1 Hz) at the center of a computer screen. The monkey was required to press a lever in response to the flicker onset. Bar press led to the presentation of the sample stimulus one second later. The test stimulus was presented after a fixed ISI (usually five seconds). The test stimulus matched the sample stimulus in half the trials. Both stimuli were presented at the center of the screen. After a variable post-test stimulus interval (500–1500 ms), the central dot stopped flickering and turned bright. This served as a 'go' signal for the monkey to shift the bar (two-sided arrow) left if the test stimulus matched the sample stimulus, or right if the two were different, and then to release the lever to get a fruit-juice reward. Both monkeys performed over 90% of trials correctly. A set of 30 color stimuli were presented in a fixed temporal order during the training session. **(b)** Examples of stimuli used. Top row, fractals; bottom row, Fourier descriptors. **(c)** A coronal MRI image of the right hemisphere of the brain of one of the studied monkeys. ch, recording chamber; ls; lateral sulcus; sts, superior temporal sulcus; rs, rhinal sulcus. During the imaging, a tungsten electrode was placed at the center of the recording chamber (dark vertical shadow). Its depth corresponds to the approximate location of the tip of the guide tube during recording. The area between the rhinal sulcus and anterior-medial-temporal sulcus recorded during the experiments is marked by triangles.

One key property of such a network is its pattern-completion abilities. The distributed representation across a large neuronal population makes it relatively insensitive to noise. If the pattern of activity during the presentation of a modified or degraded visual stimulus resembles the pattern evoked by the original stimulus, the network will reach the same neuronal delay-activity pattern (the dynamics will flow toward the same attractor). This type of neuronal behavior, IT delay activity that is insensitive to moderate levels of noise in a visual stimulus, has been reported[9].

It is important to note that the stable attractors are formed during a slow learning process, which shapes the synaptic structure between the network members. Therefore, delay activity should be evident only for stimuli that have been repeatedly presented to the animal[1]. The memories are embedded in the synaptic structure through an unsupervised Hebbian learning rule. Thus, no special assumptions or requirements are needed to generate the appropriate synaptic structure.

Last, and most important, this framework can lead to an association between stimuli repeatedly presented in temporal proximity because the delay activity can link events separated in time. Neurons that are part of an attractor of one stimulus will remain active during the delay period, until the presentation of the next stimulus. This joint activity (within a time window of tens of milliseconds) allows for Hebbian strengthing of the synapses between neurons belonging to the two populations. If the stimuli are systematically presented in a fixed temporal order, this Hebbian learning will eventually lead to similar mnemonic representations (patterns of firing rates) for the two stimuli, forming an associative memory.

According to this view, sustained activity for sets of neighboring stimuli presented in a fixed temporal sequence is a single-neuron manifestation of this association, which can only form if the memory trace following one stimulus is maintained across the intertrial interval (ITI). Theoretical considerations predict that the sustained activity following a specific stimulus will be evident during the ITI as in the interstimulus interval (ISI) because the activity evolves automatically, in a mechanical fashion, irrespective of the behavioral relevance of the stimulus. Because the sample and test stimuli are identical in half the trials in the delayed match-to-sample task, this propagation of stimulus-selective activity during the ITI could transmit information about the temporal order of the sample stimuli. Here we report that IT neurons indeed have a stimulus-selective sustained activity during the ITI. We use a simulation of a large network of (integrate and fire) neurons to illustrate the development of this sequence of events. In this simulation, the sustained activity during the ITI generates temporal correlations in the delay activity, as reported by Miyashita[1].

Results

Figure 1 illustrates the sequence of events and stimuli used in this study. We recorded the activity of 314 visually responsive cells in IT cortex. Twenty-three neurons (7.3%) showed stimulus-selective delay activity. Figure 2a and b shows the responses of two such neurons to three different stimuli presented as the sample stimulus or as the test stimulus. Note that the rasters and histograms on the left and right of the figure are not temporally contiguous because we were interested in the neuronal activity elicited by a specific stimulus, when presented as a test stimulus, irrespective of whether it matched or did not match the sample stimulus.

Stimulus #14 elicited the most vigorous firing from the IT neuron shown in Fig. 2a, which had highly selective activity, during and after its presentation. Note that the delay activity is evident after both the test stimulus and the sample stimulus. Furthermore, the delay activity evoked by stimulus #14 as the test

articles

stimulus continued throughout the ITI, which lasted six to seven seconds, and was evident until the presentation of the next sample stimulus in the following trial (**Fig. 2a**). The few trials in which activity could be seen in the pre-sample intervals are all cases when the test stimulus in the previous trial was stimulus #14. (Individual rasters are marked by different levels of the shaded area.) Thus, the conventionally defined 'spontaneous' activity is affected by the identity of the last stimulus seen. We therefore

analyzed the activity prior to the sample stimulus according to the identity of the previous test stimulus (**Fig. 2a** and **b**). The cell shown in Fig. 2b has a more widely distributed selective delay activity, which is maintained throughout the ITI following a specific test stimulus until the presentation of the next sample stimulus. The scatterplots in Fig. 2c and d show the delay activity in the last period of the ITI (stimulus activity before next sample) as a function of the delay activity in the first interval of the ITI (post-
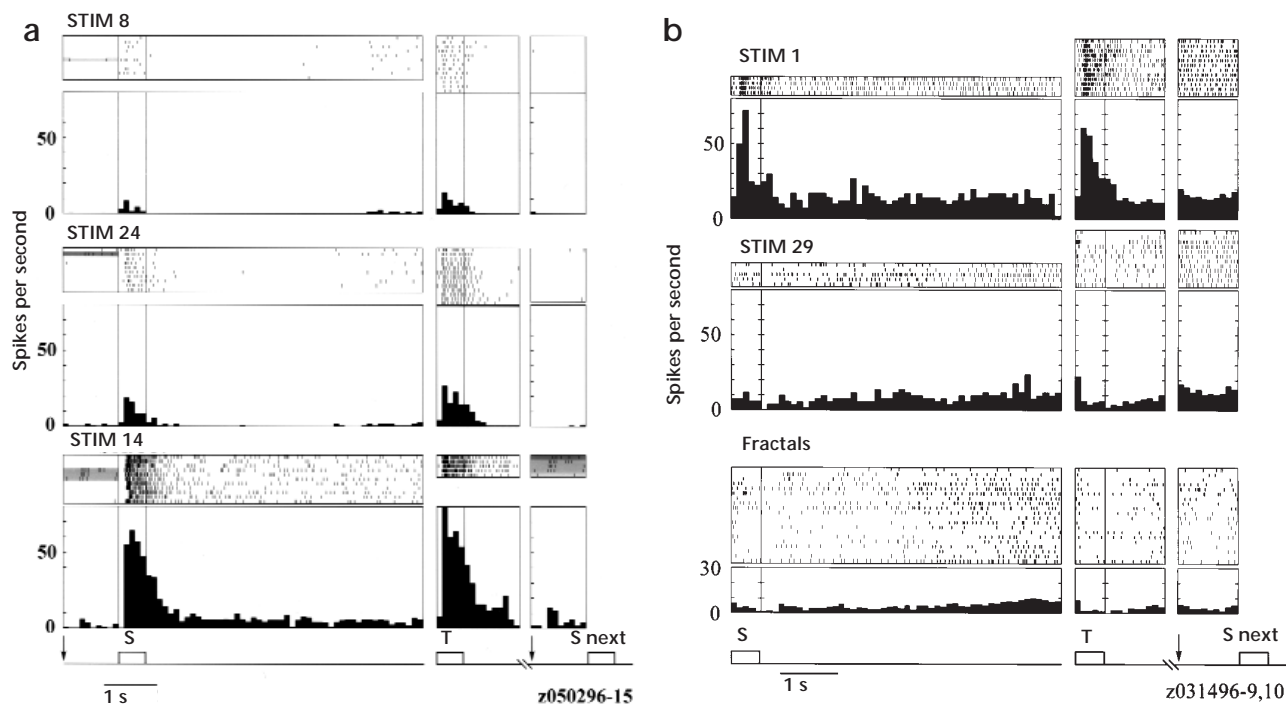
Fig. 2. Stimulus-selective sustained activity in the ITI. (a) and (b) show two example neurons with sustained activity throughout the ITI. Note that the different intervals within the trial are sorted and organized according to the identity of the corresponding stimulus. Consequently, the number and order of the rasters for the sample and test stimuli are not in register. The data shown during and following the test stimulus are combined for the same test stimulus across match and non-match conditions. S, sample stimulus; T, test stimulus; S next, sample stimulus in the following trial. **(a)** A neuron with highly selective delay activity. The right down arrow corresponds to a bar press in the beginning of the next trial, six to seven seconds after the termination of the trial (indicated by broken lines). Note that stimulus-selective delay activity is as clear following the test as following sample stimulus #14. Sustained activity following the test stimulus was evident throughout the ITI, until the onset of the sample stimulus of the next trial. Almost all the spikes recorded in the pre-sample period were a result of sustained activity following the presentation of stimulus #14 as a test stimulus in the previous trial. (Corresponding individual rasters are marked by shaded area.) **(b)** An example of a neuron with more widely distributed selective delay activity. The sustained activity following a specific test stimulus is maintained throughout the ITI until the presentation of the next sample stimulus. (See stimuli #1 and #29 compared to all the fractal stimuli, which do not elicit delay activity.) **(c, d)** Scatterplots of the average delay activity in the last period of the ITI (one second before the sample stimulus of the next trial), as a function of the average sustained activity in the beginning of the ITI (following the test stimulus). Scatter plots (c) and (d) correspond to the data from the neurons shown in (a) and (b), respectively. Each data point is an average across different presentations of a given test stimulus. (Numerous stimuli did not elicit any activity in the cell depicted in (a), and therefore the data point at the origin of (c) represents multiple stimuli.) The + sign denotes the response to the best stimulus (#14 in c, #1 in d), and the X symbols depict the response to the ineffective stimuli (#24 and #8 in c) and less effective stimulus (#29) in (d). The diagonal lines indicate points of equal response in the two time epochs.

Fig. 3. Sustained activity measures across a population of IT neurons. **(a)** A histogram depicting the distribution of the delay selectivity index across the population of selective sustained activity neurons (*n* = 23). The index is the difference in activity in the end of the ITI, following the 'best' and 'worst' stimuli, divided by the sum of the two responses. The 'best' and 'worst' stimuli were chosen according to the activity they evoked during the beginning of the ITI (the post-test period). Maintained selectivity throughout the ITI corresponds to positive values. The average sustained selectivity index was 0.44 (indicated by the arrow). This corresponds to a response more than twofold stronger for the 'best' stimulus compared to the 'worst' stimulus in the end of the ITI. **(b)** Average neuronal activity for the 'best' and 'worst' stimuli during the ISI and the beginning (post-test) and end (before next sample) periods of the ITI. Error bars indicate standard error. Note that the neuronal activity during the ISI was based on the trials in which the 'best' and 'worst' stimuli appeared as sample stimuli, whereas the activity in the ITI was according to the identity of the test stimulus. **(c)** The difference between the activity elicited by the 'best' and 'worst' stimuli during the corresponding initial and last periods of the ISI and ITI, shown individually for each neuron. The vast majority of the neurons maintain their differential response during the ITI, as in the ISI.
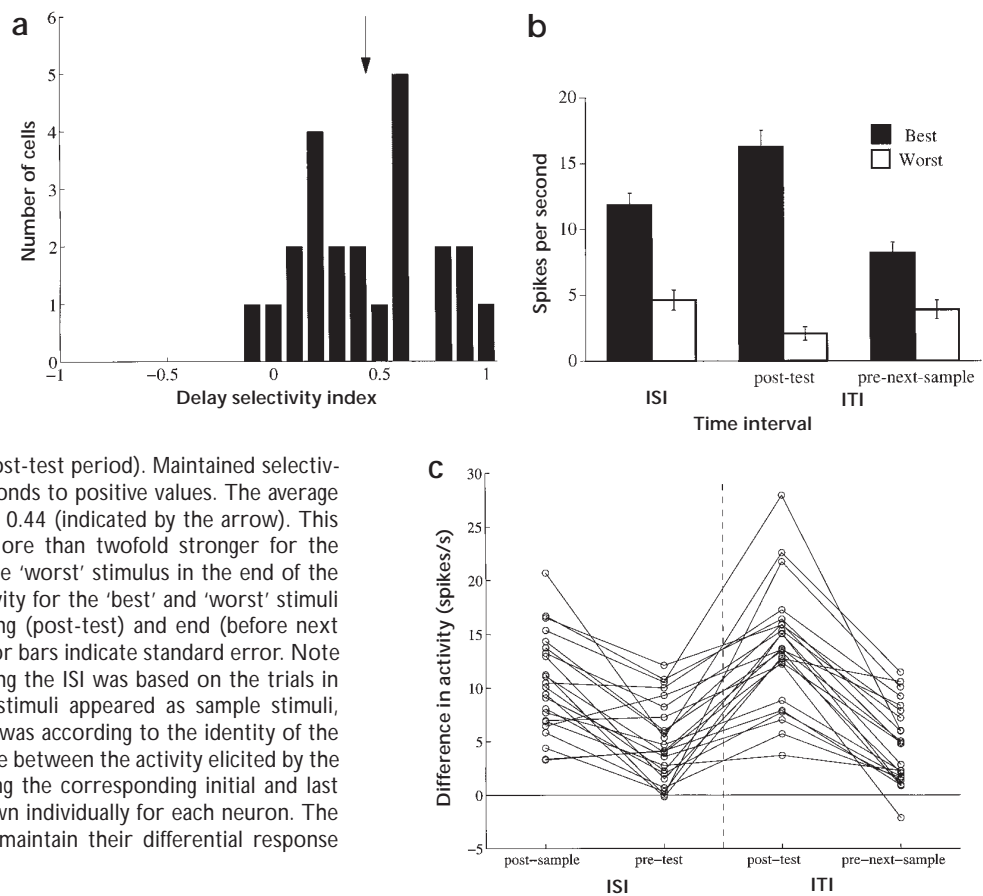
test stimulus activity) for the neurons shown in Fig. 2a and b, respectively. Each point is an average across all trials with the same test stimulus.

To evaluate the reliability of transmission of information across the ITI in the population of delay activity neurons, we compute for each neuron a delay selectivity index: We define the 'best' and 'worst' stimuli as the ones that elicit the strongest and weakest activity, respectively, during the beginning of the ITI (the post-test stimulus period). The delay selectivity index is defined as $(R_{best} - R_{worst})/(R_{best} + R_{worst})$ where R is the activity during the last period of the ITI (before the next sample). The delay selectivity index is bounded between the values of [-1,1]. A value of zero indicates that at the end of the ITI there is no difference between the responses to stimuli that elicited a very different response during the beginning of the ITI (i.e., no propagation of information). More positive values indicate the maintenance of the differential response across the ITI. Note that by definition, the best and worst stimuli will elicit a different response in the first period of the ITI, even if the neuron's delay activity is not truly stimulus selective (if the difference between the best and worst stimuli is due only to random fluctuations in the response). The crucial point is whether this differential response is maintained throughout the ITI and evident in the last period of the ITI.

Histograms of the delay selectivity index for the 23 neurons with selective sustained activity (**Fig. 3a**) show that their activity tended to continue through the ITI. The average value of the delay selectivity index for this group of neurons was 0.44 (similar to the level of selectivity, 0.52, when the same measure was

applied to the ISI). This index was significantly different from zero (one-group *t*-test, *p* < 0.0001). The average activity following the best and worst stimulus for the different time intervals (**Fig. 3b**) shows that the difference in response is evident in the ISI, in the classical delay period. This difference was also highly significant in the last period of the ITI, before the presentation of the next sample stimulus (paired *t*-test, *p* < 0.0001). Finally, a cell-by-cell analysis of the activity in the initial and final periods of the ISI and the ITI (**Fig. 3c**) demonstrates that in the vast majority of neurons, the difference in response between the best and worst stimulus was maintained across the ITI. In fact the magnitude of the differential response in the end of the ITI (5.0 spikes per s) was almost identical to that at the end of the ISI (4.6 spikes per s). The sustained activity following the 'best' sample stimulus was disrupted if the test stimulus was different from the sample, in accordance with previous findings[10]. Thus, the sustained activity depended on the identity of the last stimulus seen, whether it was a sample or test stimulus.

There was a strong positive correlation between the visual response to the sample stimulus and the delay activity in the following ISI, when the average activity for each stimulus was considered (average Pearson *r* = 0.69, *n* = 23). However, the correlation between the activity during the presentation of the best sample stimulus and the ISI delay activity on a trial-by-trial basis was much weaker. In fact, the visual response and the activity in the last second of the ISI were generally uncorrelated (average Pearson *r* = 0.10). The difference in sustained activity (between the best and worst stimulus) was greater during the post-test period (13.66 spikes per s) than during the corre-

articles

a

b
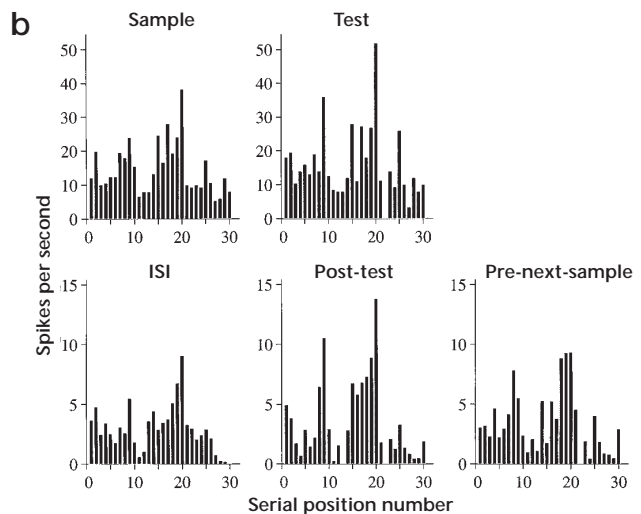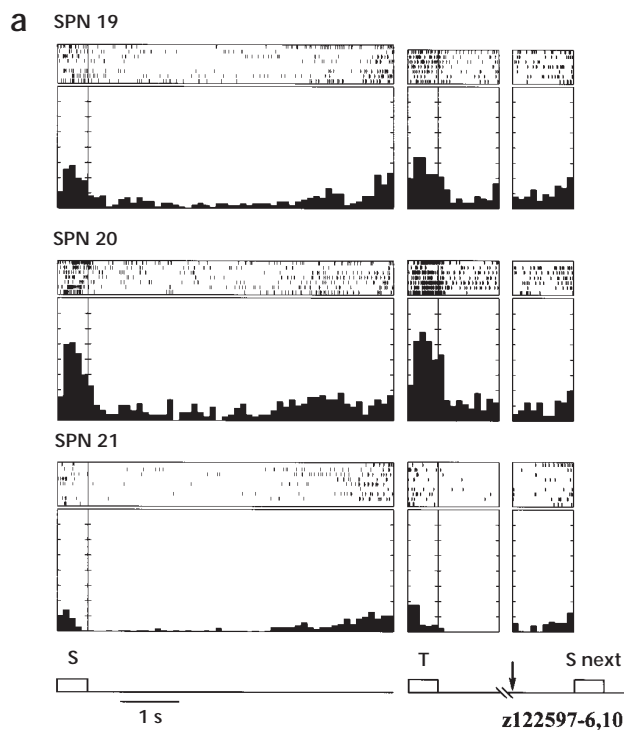
Fig. 4. Clustering of delay activity to neighboring stimuli in a fixed sequence. **(a)** Raster displays and spike density histograms of one neuron for three consecutive stimuli of the thirty stimuli that were presented in a fixed temporal order during the training stage. SPN denotes the serial position number of the stimulus in the training sequence. All other aspects of the data are presented in the same way as in Fig. 2b. **(b)** A histogram depicting the response of the same neuron during the presentation of the sample and test stimuli (top) as well as the sustained activity (bottom) during the ISI and the two ends of the ITI (post-test-stimulus period, and before next sample period) as a function of the SPN. Similar clustering of the sustained activity according to the SPN is observed in all time epochs.

sponding period in the ISI (10.34 spikes per s), but this difference was not evident at the end of the two intervals (5.0 versus 4.6 spikes per s, respectively). The response in the post-test period was also usually somewhat attenuated when the test stimulus matched the sample stimulus (13.92 versus 17.71 spikes per s for the best stimulus in the match versus non-match conditions), but again it was not significantly different at the end of the ITI (7.80 versus 8.05 spikes per s, respectively). In summary, although the delay activity immediately following a specific stimulus depended on the magnitude of the visual response, which could vary from trial to trial for different reasons, the final level of delay activity (a few seconds later) was constant. All these pieces of evidence support the suggestion that the delay activity is a result of the neural network properties, rather than a change in the state of the single neuron alone, triggered by the visual response (see also ref. 9 and below).

Is there a functional role for the propagation of delay activity across the ITI? We suggest that it may allow the generation of sustained activity for neighboring stimuli that are repeatedly shown in a fixed temporal order. Indeed, the few stimuli that evoked sustained activity (example in **Fig. 4**) were often neighboring stimuli, as previously reported[1]. Clustering of delay activity according to the serial position number of the stimulus is obvious both in the ISI and in the two ends of the ITI.

Such a context-dependent associative memory is formed in three stages according to the attractor model. In the first stage, the uncorrelated (context-independent) attractors build up. In the second stage, information about activity patterns is propagated from trial to trial by the sustained activity in the ITI, which leads to the buildup of correlations between stimuli from

consecutive trials. In the third stage, the pattern of connectivity is such that the representation of every stimulus reflects its temporal context: delay-activity distributions corresponding to neighboring stimuli in the training sequence are more correlated than the ones corresponding to distant stimuli in the training sequence. The ITI-selective activity is an essential building block for the detection and memorization of temporal correlations in the statistics of the flow of stimuli, and it is sufficient to correlate not only the nearest neighbors but also stimuli that are further apart in a temporal sequence. In the next section, following refs 7 and 11, we exemplify the mechanism underlying the formation of the temporal correlations by taking three snapshots of the behavior of the modeled network corresponding to these three stages.

MODEL NEURAL NETWORK

We present a model neural network[11] to illustrate how such a context-dependent associative memory can be formed (Methods). To focus on the role of the ITI-selective activity, we expanded the analysis of the dynamics and show the typical behavior of the model neurons during all the learning stages, as they would appear in cortical recording in each interval of the trial (visual response, ISI and ITI).

The network is composed of excitatory and inhibitory neurons, represented (for simplicity) by afferent currents and output rates. Each neuron in the network receives three types of input: recurrent excitatory connections from other neurons in the same module; nonselective, excitatory afferents from other areas of cortex; and local, nonselective inhibitory afferents. The statistics of the input currents determine the firing rates as in ref.
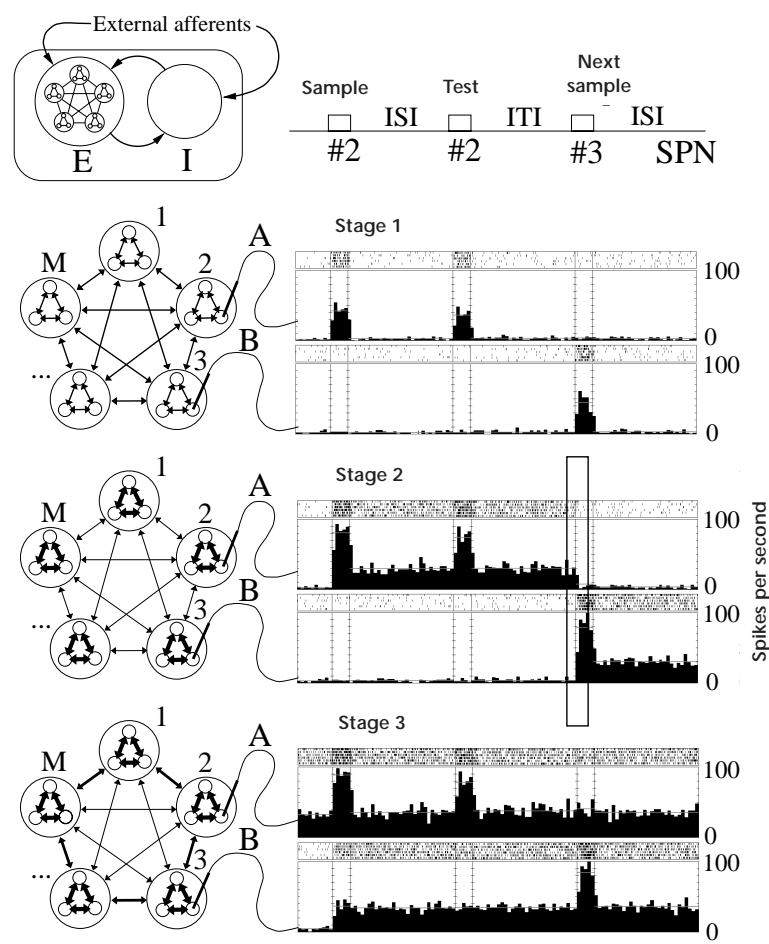
Fig. 5. Three snapshots of the behavior of the modeled network corresponding to the three stages of the development of associative memory. The left column depicts the scheme of the model network in the three stages. A network module is composed of two interconnected populations: stimulus-selective excitatory neurons and inhibitory nonselective neurons. Both populations receive external excitatory afferents from other modules in the cortex (inset in top left panel). In the excitatory population (shown in left panel, below inset), each circle denotes a subpopulation of neurons selective to a specific stimulus. Populations corresponding to stimuli that are nearest neighbors in the training sequence of length M are arranged in the picture so that they are near in the circular chain, for the convenience of presentation. Only five subpopulations are represented in the picture. Arrows denote synaptic connections. There are essentially two classes of connections: intraclass synapses, connecting neurons responding to one specific stimulus, and interclass synapses, connecting neurons responsive to different stimuli. Thicker arrows denote a higher number of potentiated connections. Two neurons (denoted by A and B) are monitored during trials of the delayed match-to-sample task in which stimuli #2 and #3 (numbered according to their serial presentation order, SPN) are repeatedly presented. (For illustration purposes only, we present ten trials in which the sample and test stimulus #2 are the same.) Their neuronal activity in the different stages is shown in the right column. (The dotted gray lines represent the average activity in each interval.) See text for an explanation of the dynamics.

12. The excitatory neurons in the module belong to subpopulations, each responding (for simplicity) to only one stimulus.

Figure 5 shows the development of delay activity in model neurons during the training process, using stimuli that were repeatedly presented to the network in a protocol identical to the one we used in the physiology experiment above. During stimulation (when the sample or test stimuli are shown), an extra current is injected in the subpopulation of neurons responding to the stimulus presented. The elevated activity of these neurons leads to an increase in the activity of the population of inhibitory neurons, which always reflects the global activation of the excitatory population. As a result, activity is depressed in the other subpopulations that are not activated by the stimulus.

In the first stage, the strength of the interclass (between subpopulations responding to different stimuli) and intraclass (within subpopulations) connections is randomly chosen, and each neuron shows a stimulus-selective visual response but no delay activity. The joint firing of two neurons activated by the same specific stimulus (for instance, neuron A and a similar neuron from the same group of neurons responding to stimulus #2) leads to the potentiation of the connection between the two. Analogously, the interclass connections tend to be depressed. With enough repetitions of the same stimulus, there are enough potentiated synapses that the network can sustain enhanced activity even after the evoking stimulus has been removed: each neuron in the subpopulation excites the others through the potentiated synapses. At the end of this stage, delay activity distributions (attractors) are formed for each specific stimulus. This network property appears suddenly and is observed as a stimulus-specific delay activity (shown in stage 2).

Because of this stimulus specificity, a positive correlation between the visual response and the following delay activity, as we report here, is expected (see also ref. 4). On the other hand, the sustained activity evoked by one specific stimulus is not affected by fluctuations in strength of the visual response from trial to trial, as we reported above. This is because the delay activity is triggered by the visual response, but it is sustained by the pattern of activity of all the neurons in the same module. The visual response determines the initial condition. All the stimuli that evoke patterns of activity in the same basin of attraction lead to the same final steady state (attractor), irrespective of the fluctuations of individual neurons' activity (see also ref. 9).

In stage 2, stimulus-selective delay activity exists, but the patterns of delay activity across the population of neurons are initially not overlapping (i.e., each neuron has sustained activity to only one stimulus). The delay activity of a specific subpopulation is triggered by the presentation of the corresponding stimulus and ends with the presentation of a different stimulus. This is because the global inhibition generated by a different visual stimulus is enough to suppress the activity of this subpopulation to its spontaneous activity level.

When the test stimulus matches the sample, neuron A will have delay activity following test stimulus (#2), until the presentation of the next stimulus, which generates a visual response in

all neurons of subpopulation #3 (including neuron B). This joint activity (within a time window of less than 100 ms, at the end of the ITI for stimulus #2) allows for Hebbian strengthening of the synapse between the two neurons. The result of this unsupervised learning is that in the final stage (stage 3), the neurons show sustained activity also for the neighboring stimuli of their preferred stimulus, as reported above (**Fig. 4**; see also ref. 1). The spreading of the delay activity to the neighboring stimuli is limited to a maximal distance of a few stimuli. This is because the inhibition is faster and stronger than excitation. Thus, the dampening effect of the total inhibition becomes dominant whenever the total excitation tends to grow. Moreover, potentiation depends on the joint level of activity of the two neurons. The delay activity is generally weaker than the visual activity to a given stimulus. Therefore, the delay activity of neuron A elicited by the neighboring stimulus (#3) will usually be weaker than the activity evoked by the original stimulus (#2), and the chain reaction will be limited (see refs 7,11,13). (In our case, the parameters are such that the maximal distance is five. This limitation is not obvious in the figure because we show only the immediate neighboring stimuli in the sequence.)

Discussion

The most prominent and novel finding reported here is that stimulus-selective delay activity in IT cortex persists across the ITI. We suggest that this propagation of activity across the ITI may serve to generate the synaptic structure required to form correlations between the mnemonic representations (delay-activity distributions) of successive stimuli in a sequential training protocol[9]. An analogous type of sustained activity that persisted across the ITI is reported in prefrontal cortex[14]. This activity is not related to eye movements and was seen also in a monkey that was never trained on a memory task, indicating that it evolves automatically. The sub-area within prefrontal cortex where these face-selective neurons are found receives strong input from IT.

Stimulus-selective delay activity was considered to encode the memory trace during the ISI[1,2,15]. Sample-specific delay activity in prefrontal cortex is maintained throughout the trial, even when intervening stimuli are presented, whereas delay activity following the sample stimulus is disrupted by intervening stimuli in IT cortex[10,16]. These authors concluded that prefrontal cortex may subserve 'active' working memory, whereas IT cortex may contribute to an automatic detection of stimulus repetition. Our results are in agreement with the hypothesis of a 'passive,' automatic memory in IT.

One could suggest that the sustained activity following the test stimulus was due to active working memory because the identity (match or non-match) of the test stimulus must be remembered to execute a correct response. However, the stimulus-specific sustained activity following the test stimulus was evident even after the reward, when memory of the stimulus was no longer required. The delay activity also cannot serve as the mnemonic trace of the sample stimulus throughout the trial, as it was disrupted by the presentation of a different test stimulus. Thus, the sustained activity seems to reflect the last familiar stimulus seen, irrespective of its relevance to the behavioral task.

Could the sustained activity be a result of unmonitored eye movements? The sustained activity was stimulus specific and reproducible. It occurred both in the ISI, when the animal gazed at the center of screen, as was observed by the video camera, and in the ITI, when the monkey was clearly observed making eye movements. Thus, it is highly unlikely that the delay activity was caused by systematic eye movements following a specific stimu-

lus. Furthermore, eye movements do not influence the delay activity of IT neurons in the ISI[4].

In what circumstances would such a mechanism have behaviorally observed consequences? It was suggested[11] that such activity would be highly effective in a paired association task, in which retrieval from long-term memory of the pair member associated with a given cue is required. Such associations are formed by repeated presentations of the paired associates, and monkeys with lesions of the temporal lobe (rhinal cortex) or the connections between inferotemporal and prefrontal cortex show marked impairment in this task[17,18]. In such a paired association task, neurons in IT cortex selectively respond to both pictures of the paired associates[3]. Furthermore, the neuron shown as an example in that study had a similar level of delay activity following the presentation of either of the two paired stimuli as a cue, suggesting that both stimuli now evoke the same pattern of delay activity (i.e. the same attractor).

A key requirement for the buildup of an attractor network is that neurons are organized in local groups with similar stimulus specificity or that neurons with similar specificity are preferentially connected. Indeed, neighboring neurons in inferotemporal cortex tend to have similar stimulus preferences, determined by single-unit recording and optical imaging techniques[19]. A similar model of attractor dynamics was suggested for the generation of invariant face and object recognition in vision. In essence, it suggests that cells in IT cortex respond similarly to objects seen from different viewing angles because usually faces or objects are seen sequentially from different views in a temporal sequence as one is manipulating an object or moving in space[20,21]. We conclude that the stimulus-selective sustained activity in IT reflects a passive, automatic memory. The persistence of stimulus-selective activity across the ITI may serve as the link to generate associations between neighboring stimuli by modification of the synaptic structure, so that correlations between the neural representations of successive stimuli are formed.

Such a scheme of association may be relevant for navigation in an unfamiliar environment. In such circumstances, we usually remember the specific route we have taken, rather than rely on a cognitive spatial map. Navigation in such circumstances relies heavily on remembered snapshots of visual scenes from specific angles, with one cue leading us to the next expected landmark. Interestingly, lesions in parietal cortex typically lead to a failure to grasp the spatial relationships between places (i.e., a failure to generate a cognitive map) with intact landmark recognition[22]. Temporal lobe lesions in humans, on the other hand, often result in topographical disorientation in novel environments, when landmarks along the route are used[23]. Furthermore, such topographical agnosia often co-occurs with prosopagnosia (inability to recognize familiar faces)[24,25]. This paradoxical finding is more easily understood if attractor dynamics in the temporal lobe is the common neural mechanism underlying the two mnemonic functions.

Methods

BEHAVIORAL TASK AND VISUAL STIMULI. The activity of single neurons was recorded from IT cortex while monkeys performed a visual delayed match-to-sample task. The monkeys were seated in an isolated experimental chamber with a background illumination of 2 cd per m². The only objects in front of the monkey were the computer monitor and a video camera. The background luminance of the screen was 12 cd per m², and the colored images were high-contrast pictures. A set of 30 color stimuli were presented in a fixed temporal order during the training session. Fifteen were fractal stimuli, and the rest were Fourier descriptors.

ANIMALS AND SURGICAL PROCEDURES. Two rhesus monkeys (*Macaca mulatta*) weighting six to seven kg were used. A head post and a record-

ing chamber were implanted above anterior-ventral IT cortex under general anesthesia with nembutal (25–30 mg per kg). The monkeys were given antibiotics and analgesics postoperatively and were allowed sufficient time for recovery after surgery. All experiments, MRI tests and surgical preparations were performed in accordance with NIH and Hebrew University guidelines for use of laboratory animals for experiments.

ANATOMICAL MRI. We applied magnetic resonance imaging (MRI) using a Biospec 47/40 device (Bruker) to verify the position of the recording chamber relative to the area of interest. A series of coronal T2-weighted images (13–15 consecutive 2 mm slices) were recorded covering the whole area of interest in the monkey brain. A tungsten electrode (diameter, 200 μm) was inserted through the chamber center above the area explored during the actual recording sessions (**Fig. 1c**). This area was between the rhinal sulcus and anterior-medial-temporal sulcus. The images were recorded using a spin-echo sequence with the following parameters: field-of view of 13 x 13 cm, 256 x 256 data matrix, RARE factor of 8, TR/TE of 3000/23 ms and 8 scans yielding an effective T2-weighted contrast images corresponding to normal spin-echo taken with TE of 70 ms. The monkeys were anesthetized during the imaging session, which lasted about 15 min.

RECORDING AND DATA ANALYSIS. Single-unit activity was monitored in four hemispheres of two monkeys using standard recording techniques. Because of technical limitations (data transfer between computers, generation of new stimuli, etc.), neuronal activity was registered during the period between the beginning of the trial (presentation of flickering dot) and bar release. Therefore, the activity during the ITI was monitored in two discrete periods: the post-test stimulus activity (the firing rate between the test stimulus offset and the bar release) and the activity before the next sample, defined as the firing rate in the interval prior to the next sample stimulus, from the presentation of the flickering dot to the sample-stimulus onset. The activity during the ISI was defined as the firing rate in the interval between the sample and test stimuli. The first 200 ms following stimulus offset in the ISI and ITI were excluded to avoid the effects of a possible visual response. Neurons were considered to have a stimulus-specific delay activity if the firing rates for the various stimuli during both the ISI and post-test stimulus period were statistically different using one-way analysis of variance (ANOVA, $p < 0.001$).

DETAILS OF THE MODEL. The parameters are as described[11]. The statistics of the input currents determine the firing rates as in ref. 12, where the current-to-rate transduction function was calculated for leaky integrate-and-fire neurons. The integration time constant for excitatory (inhibitory) neurons is 10 ms (2 ms), and the spike emission threshold is 20 mV above the resting level. Each neuron receives 10^4 afferents from randomly selected excitatory neurons of the same module, 2×10^3 afferents from the population of inhibitory neurons and an external current from other unspecified areas. The mean synaptic efficacies are chosen in such a way that in the first stage, when the synaptic matrix is still not structured, the average spontaneous activity is 3.0 spikes per s for the excitatory neurons and 4.1 spikes per s for the inhibitory neurons. (The EPSPs are $J_{E \text{ to } E} = 0.035$ mV, $J_{E \text{ to } I} = 0.054$ mV, $J_{I \text{ to } E} = J_{I \text{ to } I} = -0.141$ mV.) The external mean excitatory current is the same as the mean recurrent excitatory current when all the neurons have spontaneous activity. During stimulation, an extra gaussian current is injected in the neurons of the subpopulation (fraction $f = 0.01$ of the excitatory neurons in the network) corresponding to the activated stimulus ($\mu = 8.25$ mV per ms, $\sigma^2 = 0.9$ mV^2 per ms). Only the excitatory synapses in a module are modifiable, and each synapse has two potentiation levels[26]. The high level (potentiated state) corresponds to a synaptic efficacy that is 4.4 times larger than the low level (depressed state). Synaptic transitions between the two levels depend on the mean rates of the pre- and postsynaptic neurons. Long-term potentiation (LTP) corresponds to the transition between the low level and the high level and occurs with probability $p_+ = 0.2$ if the pre- and postsynaptic neurons are simultaneously activated by the stimulus. (In other words, following each repetition, a mean fraction p_+ of the depressed synapses that are connecting active neurons makes a transition to the potentiated state.) If one neuron is activated by a stimulus (high rate) and the other carries selective delay activity elicited by the previous stimulus seen, then potentiation occurs with probability $p'_+ = ap_+$

($a = 0.015$). Long-term depression (LTD) occurs with a probability $p_- = 0.2$ when one neuron is activated by the stimulus while the other is firing at its spontaneous rate.

1. Miyashita, Y. Neuronal correlate of visual associative long-term memory in the primate temporal cortex. *Nature* **335**, 817–820 (1988).
2. Fuster, J. M. & Jervey, J. P. Inferotemporal neurons distinguish and retain behaviorally relevant features of visual stimuli. *Science* **212**, 952–955 (1981).
3. Sakai, K. & Miyashita, Y. Neural organization for the long-term memory of paired associates. *Nature* **354**, 152–155 (1991).
4. Nakamura, K. & Kubota, K. Mnemonic firing of neurons in the monkey temporal pole during a visual recognition memory task. *J. Neurophysiol.* **74**, 162–178 (1995).
5. Wilson, F. A., Scalaidhe, S. P. & Goldman-Rakic, P. S. Dissociation of object and spatial processing domains in primate prefrontal cortex. *Science* **260**, 1955–1958 (1993).
6. Rao, S. C., Rainer, G. & Miller, E. K. Integration of what and where in the primate prefrontal cortex. *Science* **276**, 821–824 (1997).
7. Amit, D. J., Brunel, N. & Tsodyks, M. V. Correlations of cortical hebbian reverberations: Theory versus experiment. *J. Neurosci.* **14**, 6435–6445 (1994).
8. Amit, D. J. The Hebbian paradigm reintegrated: Local reverberations as internal representations. *Behav. Brain Sci.* **18**, 617–657 (1995).
9. Amit, D. J., Fusi, S. & Yakovlev, V. Paradigmatic working memory (attractor) cell in IT cortex. *Neural Comput.* **9**, 1071–1093 (1997).
10. Miller, E. K., Li, L. & Desimone, R. Activity of neurons in anterior inferior temporal cortex during a short-term memory task. *J. Neurosci.* **13**, 1460–1478 (1993).
11. Brunel, N. Hebbian learning of context in recurrent neural networks. *Neural Comput.* **8**, 1677–1710 (1996).
12. Amit, D. J. & Brunel, N. Global spontaneous activity and local structured (learned) delay activity in cortex. *Cereb. Cortex* **7**, 237–252 (1997).
13. Griniasti, M., Tsodyks, M. V. & Amit, D. J. Conversion of temporal correlations between stimuli to spatial correlations between attractors. *Neural Comput.* **5**, 1–9 (1993).
14. Scalaidhe, S. P. O., Wilson, F. A. & Goldman-Rakic, P. S. Areal segregation of face-processing neurons in prefrontal cortex. *Science* **278**, 1135–1138 (1997).
15. Fuster, J. M. Inferotemporal units in selective visual attention and short-term memory. *J. Neurophysiol.* **64**, 681–697 (1990).
16. Miller, E. K., Erickson, C. & Desimone, R. Neural mechanisms of working memory in prefrontal cortex of the macaque. *J. Neurosci.* **16**, 5154–5167 (1996).
17. Murray, E. A., Gaffan, D. & Mishkin, M. Neural substrates of visual stimulus-stimulus association in rhesus monkeys. *J. Neurosci.* **13**, 4549–4561 (1993).
18. Gutnikov, S. A., Yuan-Ye, M. & Gaffan, D. Temporo-frontal disconnection impairs visual-visual paired association learning but not configural learning in macaca monkeys. *Eur. J. Neurosci.* **9**, 1524–1529 (1997).
19. Tanaka, K. Inferotemporal cortex and object vision. *Annu. Rev. Neurosci.* **19**, 109–140 (1996).
20. Wallis, G. & Rolls, E. T. Invariant face and object recognition in the visual system. *Prog. Neurobiol.* **51**, 167–194 (1997).
21. Bartlett, M. & Sejnowski, T. Learning viewpoint invariant representations from visual experience using attractor networks. *Network* (in press).
22. DeRenzi, E., Faglioni, P. & Villa, P. Topographical amnesia. *J. Neurol. Neurosurg. Psychiatry* **40**, 498–505 (1977).
23. Whiteley, A. M. & Warrington, E. K. Selective impairment of topographical memory: a single case study. *J. Neurol. Neurosurg. Psychiatry* **41**, 575–578 (1978).
24. Landis, T. Cummings, J. L. Benson, D. F. & Palmer, E. P. Loss of topographic familiarity. *Arch. Neurol.* **43**, 132–136 (1986).
25. Maguire, E. A., Burke, T. Philips, J. & Staunton, H. Topographical disorientation following unilateral temporal lobe lesions in humans. *Neuropsychologia* **34**, 994–1004 (1996).
26. Petersen, C. C. H., Malenka, R. C., Nicoll, R. A. & Hopfield, J. J. All-or-none potentiation at CA3-CA1 synapses. *Proc. Natl. Acad. Sci. USA* **95**, 4732–4737 (1998).