

Diluted connectivity in pattern association networks facilitates the recall of information from the hippocampus to the neocortex

Edmund T. Rolls^{*,†,1}

^{*}*Oxford Centre for Computational Neuroscience, Oxford, UK*

[†]*Department of Computer Science, University of Warwick, Coventry, UK*

¹*Corresponding author: e-mail address: edmund.rolls@oxcns.org*

Abstract

The recall of information stored in the hippocampus involves a series of corticocortical back-projections via the entorhinal cortex, parahippocampal gyrus, and one or more neocortical stages. Each stage is considered to be a pattern association network, with the retrieval cue at each stage the firing of neurons in the previous stage. The leading factor that determines the capacity of this multistage pattern association backprojection pathway is the number of connections onto any one neuron, which provides a quantitative basis for why there are as many backprojections between adjacent stages in the hierarchy as forward projections. The issue arises of why this multistage backprojection system uses diluted connectivity. One reason is that a multistage backprojection system with expansion of neuron numbers at each stage enables the hippocampus to address during recall the very large numbers of neocortical neurons, which would otherwise require hippocampal neurons to make very large numbers of synapses if they were directly onto neocortical neurons. The second reason is that as shown here, diluted connectivity in the backprojection pathways reduces the probability of more than one connection onto a receiving neuron in the backprojecting pathways, which otherwise reduces the capacity of the system, that is the number of memories that can be recalled from the hippocampus to the neocortex. For similar reasons, diluted connectivity is advantageous in pattern association networks in other brain systems such as the orbitofrontal cortex and amygdala; for related reasons, in autoassociation networks in, for example, the hippocampal CA3 and the neocortex; and for the different reason that diluted connectivity facilitates the operation of competitive networks in forward-connected cortical systems.

Keywords

Hippocampus, Memory, Recall, Diluted cortical connectivity, Pattern association, Backprojections, Episodic memory, Autoassociation, Cortical backprojections, CA3, CA1, Dentate granule cells

1 INTRODUCTION

The operation of hippocampal circuitry as a memory system, and evidence that supports the theory that has been developed, is considered in detail elsewhere (Kesner and Rolls, 2015; Rolls, 2010, 2013b,c; Rolls and Kesner, 2006), as has the development of the theory (Rolls, 1987, 1989a,b,c, 1996; Rolls and Stringer, 2005; Rolls et al., 2002; Treves and Rolls, 1992, 1994), and its relation to other theories and approaches (Marr, 1971; McClelland et al., 1995; McNaughton and Morris, 1987). This computational theory of hippocampal function remains the only quantitative theory of hippocampal function in memory and its recall to the neocortex.

In this chapter, I focus on how information is recalled from the hippocampus to the neocortex, and introduce new hypotheses and evidence about the advantages of diluted connectivity in this backprojection circuitry, as well as in pattern association networks in the brain in general. I also compare this advantage to the advantages of diluted connectivity in autoassociation networks in the cortex such as those in the CA3 region of the hippocampus and local areas of the neocortex, and to the advantages of diluted connectivity in competitive networks in the brain, including those implemented in the forward connectivity between cortical areas. By diluted connectivity, I mean that there are fewer synaptic connections onto each neuron than there are neurons in the population. If there are $C = 10,000$ connections onto every neuron, and $N = 100,000$ neurons that receive the inputs, then the dilution of the connectivity is 0.1. Full connectivity refers to the situation in which there is one synaptic connection onto every neuron from all of the inputs.

I start with a brief overview of the operation of hippocampal circuitry (which is described in detail elsewhere, Kesner and Rolls, 2015; Rolls, 2008a, 2010), in which I treat the issue of the advantages of diluted connectivity within CA3. Then, I consider the recall of information from the hippocampus to the neocortex, including here the new hypotheses and evidence about the advantages of diluted connectivity in pattern association networks in the cortex.

2 OVERVIEW OF THE OPERATION OF HIPPOCAMPAL CIRCUITRY

2.1 SYSTEMS-LEVEL FUNCTIONS OF THE HIPPOCAMPUS

The hypothesis is that the hippocampus is involved in storing information in an unstructured way that can be used for episodic memory of single events or sequences of

events, and that recall from the hippocampus back to the neocortex is used to help build semantic including autobiographical memories. To implement this, the hippocampus is involved in spatial, object–place, reward–place, and object–temporal sequence memory (Kesner and Rolls, 2015). The systems-level neurophysiology of the primate hippocampus shows that it contains representations of space “out there,” that is, spatial view cells (Georges-François et al., 1999; Robertson et al., 1998; Rolls and O’Mara, 1993, 1995; Rolls and Xiang, 2005, 2006; Rolls et al., 1989, 1997, 1998, 2005), which are appropriate for a human episodic memory system, for which associations between objects or rewards and the places where they are seen are prototypical.¹

2.2 HIPPOCAMPAL CIRCUITRY

Hippocampal circuitry is illustrated in Figs. 1 and 2.

2.3 HIPPOCAMPAL COMPUTATION

2.3.1 Dentate Granule Cells

The theory is that the dentate granule cell stage of hippocampal processing which precedes the CA3 stage acts in a number of ways including competitive learning and the randomizing effect of the small numbers of mossy fiber connections onto each CA3 neuron (see Fig. 2) to produce during learning the sparse, yet efficient (i.e., nonredundant) representation in CA3 neurons that is required for the autoassociation performed in CA3 to perform well (Kesner and Rolls, 2015; Rolls, 1989a,b,d, 1996, 2008a, 2013b,c; Treves and Rolls, 1992). An important property for episodic memory is that the dentate by acting in this way would perform pattern separation (or orthogonalization; Rolls, 1989b, 2008a, 2013b,c; Rolls and Kesner, 2006; Rolls et al., 2006; Treves and Rolls, 1992), enabling the hippocampus to store different memories of even similar events, and this prediction has been confirmed by investigations in rodents (Kesner and Rolls, 2015).

2.3.2 CA3 as an Autoassociation Memory

The hypothesis is that the CA3 operates effectively as a single autoassociation network (see Figs. 2 and 3) to allow arbitrary associations between inputs originating from very different parts of the cerebral cortex to be formed, and later for the whole

¹John O’Keefe was one of the recipients of the Nobel Prize for Physiology or Medicine in 2014, after this paper was written. His work on the discovery of hippocampal place cells in rats (“O’Keefe and Dostrovsky, 1971”) was cited, and he is congratulated. Indeed, the announcement for the award of the Nobel Prize described this system as a “component of a positioning system, an ‘inner GPS’ in the brain.” John O’Keefe has continued to emphasize the role of the hippocampus and rodent place cells in navigation (Hartley et al., 2014; “O’Keefe, 1990”). Rolls’ discoveries and theory are thus somewhat different, in that Rolls has shown that spatial view cells may be especially relevant to the operation of the hippocampus in primates including humans; and in that the roles of the hippocampal system in memory are emphasized (Kesner and Rolls, 2015; Rolls, 2008a; Rolls and Xiang, 2006).

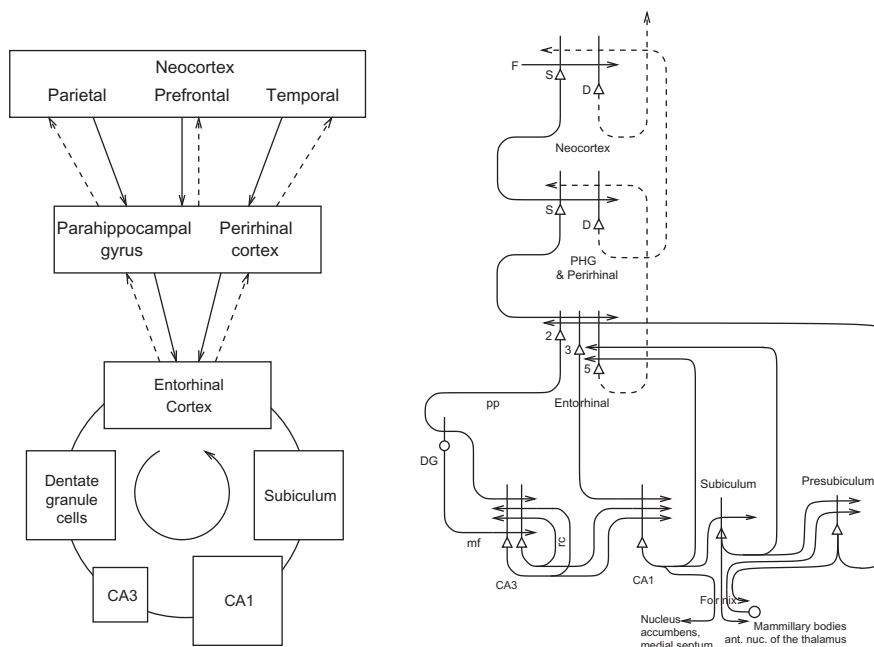
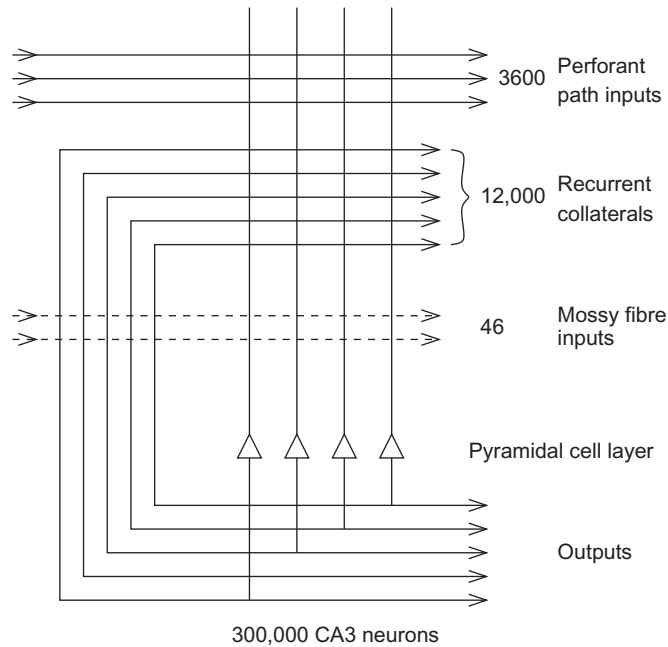


FIGURE 1

Forward connections (solid lines) from areas of cerebral association neocortex via the parahippocampal gyrus and perirhinal cortex, and entorhinal cortex to the hippocampus; and backprojections (dashed lines) via the hippocampal CA1 pyramidal cells, subiculum, and parahippocampal gyrus to the neocortex. There is a great convergence in the forward connections down to the single network implemented in the CA3 pyramidal cells and a great divergence again in the backprojections. Left: block diagram. Right: more detailed representation of some of the principal excitatory neurons in the pathways. D, deep pyramidal cells; DG, dentate granule cells; F, forward inputs to areas of the association cortex from preceding cortical areas in the hierarchy; mf, mossy fibers; PHG, parahippocampal gyrus and perirhinal cortex; pp, perforant path; rc, recurrent collateral of the CA3 hippocampal pyramidal cells; S, superficial pyramidal cells; 2, pyramidal cells in layer 2 of the entorhinal cortex; 3, pyramidal cells in layer 3 of the entorhinal cortex. The thick lines above the cell bodies represent the dendrites.

memory to be recalled from any part in the process termed completion. We have extended previous formal models of autoassociative memory (see Amit, 1989; Hopfield, 1982) by analyzing a network with graded response units, so as to represent more realistically the continuously variable rates at which neurons fire, and with incomplete connectivity and with sparse representations (Treves, 1990; Treves and Rolls, 1991). We have found that, in general, the maximum number p_{\max} of firing patterns that can be (individually) retrieved is proportional to the number C^{RC} of (associatively) modifiable recurrent collateral synapses per cell, by a factor that increases roughly with the inverse of the sparseness a of the neuronal representation.

**FIGURE 2**

The numbers of connections from three different sources onto each CA3 cell from three different sources in the rat.

After Rolls and Treves (1998) and Treves and Rolls (1992).

(The sparseness for a binary representation is the proportion of neurons firing for any one pattern.) Approximately,

$$p_{\max} \cong \frac{C^{\text{RC}}}{a \ln\left(\frac{1}{a}\right)} k \quad (1)$$

where k is a factor that depends weakly on the detailed structure of the rate distribution, on the connectivity pattern, etc., but is roughly in the order of 0.2–0.3 (Treves and Rolls, 1991).

Figure 2 shows that in the rat, there are approximately 300,000 CA3 neurons, but only 12,000 recurrent collateral synapses per neuron. The dilution of the connectivity is thus $12,000/300,000 = 0.04$. We have shown how analysis of the capacity of attractor networks (Hopfield, 1982) can be extended to the case with diluted connectivity, and also with sparse representations with graded firing rates (Rolls and Treves, 1990; Rolls and Webb, 2012; Treves, 1990, 1991; Treves and Rolls, 1991; Webb et al., 2011).

However, the question has recently been asked about whether there are any advantages to autoassociation or attractor networks with diluted connectivity compared to fully connected attractor networks (Rolls, 2012a). One biological property that

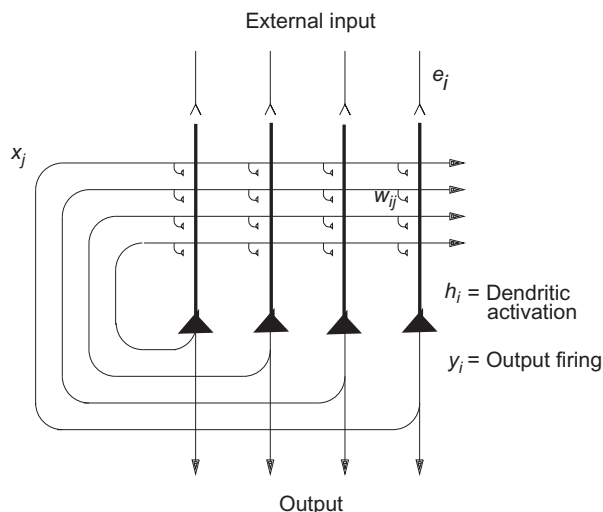


FIGURE 3

Autoassociation memory. The architecture of an autoassociation memory. The external input e_i is applied to each neuron i by unmodifiable synapses. This produces firing y_i of each neuron. Each output neuron i is connected by a recurrent collateral connection to the other neurons in the network, via modifiable connection weights w_{ij} . This architecture effectively enables the output firing vector \mathbf{y} to be associated during learning with itself. Later on, during recall, presentation of part of the external input will force some of the output neurons to fire, but through the recurrent collateral axons and the modified synapses, other neurons in \mathbf{y} can be brought into activity. This process can be repeated a number of times, and recall of a complete pattern may be perfect. Effectively, a pattern can be recalled or recognized because of associations formed between its parts. This of course requires distributed representations.

may be a limiting factor is the number of synaptic connections per neuron, which is 12,000 in the CA3–CA3 network just for the recurrent collaterals (see Fig. 2). The number may be higher in humans, allowing more memories to be stored in the hippocampus than order 12,000. I note that the storage of large number of memories may be facilitated in humans because the left and right hippocampus appear to be much less connected between the two hemispheres than in the rat, which effectively has a single hippocampus (Rolls, 2008a). In humans, with effectively two separate CA3 networks, one on each side of the brain, the memory storage capacity may be doubled, as the capacity is set by the number of recurrent collaterals per neuron in each attractor network (Eq. 1). In humans, the right hippocampus may be devoted to episodic memories with spatial and visual components, whereas the left hippocampus may be devoted to memories with verbal/linguistic components, i.e., in which words may be part of the episode (e.g., who said what to whom and when; Barkas et al., 2010; Bonelli et al., 2010; Sidhu et al., 2013).

The answer that has been suggested to why the connectivity of the CA3 autoassociation network is diluted (and why neocortical recurrent networks are also diluted) is that this may help to reduce the probability of having two or more synapses between any pair of randomly connected neurons within the network, which it has been shown greatly impairs the number of memories that can be stored in an attractor network, because of the distortion that this produces in the energy landscape (Rolls, 2012a). In more detail, the hypothesis proposed is that the diluted connectivity allows biological processes that set up synaptic connections between neurons to arrange for there to be only very rarely more than one synaptic connection between any pair of neurons, assuming that synapses are made at random between neurons. If the average connectivity between neurons was 1, then some neurons would receive more than one synaptic input from a given afferent neuron, with the proportions of multiple synapses set by the Poisson distribution with $\lambda = 1$. If probabilistically there were more than one connection between any two neurons, it was shown by simulation of an autoassociation attractor network that such connections would dominate the attractor states into which the network could enter and be stable, and thus strongly reduce the memory capacity of the network (the number of memories that can be stored and correctly retrieved), below the normal large capacity for diluted connectivity (Rolls, 2012a). Diluted connectivity between neurons in the cortex thus has an important role in allowing high capacity of memory networks in the cortex, and helping to ensure that the critical capacity is not reached at which overloading occurs leading to an impairment in the ability to retrieve any memories from the network (Rolls, 2012a). The diluted connectivity is thus seen as an adaptation that simplifies the genetic specification of the wiring of the brain, by enabling just two attributes of the connectivity to be specified (e.g., from a CA3 to another CA3 neuron chosen at random to specify the CA3–CA3 recurrent collateral connectivity), rather than which particular neuron should connect to which other particular neuron (Rolls, 2012a; Rolls and Stringer, 2000). Consistent with this hypothesis, there are NMDA receptors with the genetic specification that they are NMDA receptors on neurons of a particular type, CA3 neurons (as shown by the evidence from CA3-specific vs. CA1-specific NMDA receptor knockouts; Nakazawa et al., 2002, 2003, 2004; Rondi-Reig et al., 2001). A consequence is that the vector of output neuronal firing in the CA3 region, i.e., the number of CA3 neurons, is quite large (300,000 neurons in the rat). The large number of elements in this vector may have consequences for the noise in the system (Rolls and Webb, 2012).

Part of the answer to why there is a large number of neurons in CA3 (300,000 in the rat) compared to the number of synapses onto each neuron (12,000 in the rat, see Fig. 2) is that this provides for diluted connectivity in the CA3–CA3 system to minimize the number of multiple connections between any pair of CA3 neurons, which would degrade the memory storage capacity of CA3 very considerably (Rolls, 2012a).

The theory is also that the perforant path inputs to CA3 cells with the necessary associative synaptic modifiability initiate recall in CA3 and contribute to generalization (Treves and Rolls, 1992).

2.3.3 CA1 Cells

The CA3 cells connect to the CA1 cells by the Schaeffer collateral synapses. Associativity in these connections increases the number of patterns that can be correctly transferred to CA1; the information from the parts of a memory such as about object and place, necessarily separate in CA3 so that they can be associated, can be combined in CA1 to make a more efficient retrieval cue for the whole memory; and expansion of cell numbers occurs, in preparation for the massive divergence needed to address large areas of the neocortex during recall (Kesner and Rolls, 2015; Rolls, 1995; Schultz and Rolls, 1999; Treves, 1995; Treves and Rolls, 1994).

3 BACKPROJECTIONS TO THE NEOCORTEX, EPISODIC MEMORY RECALL, AND CONSOLIDATION

3.1 HOW BACKPROJECTIONS FROM THE NEOCORTEX TO THE HIPPOCAMPUS COULD IMPLEMENT RECALL

The need for information to be retrieved from the hippocampus to affect other brain areas was noted in Section 1. The way in which this could be implemented via backprojections to the neocortex (Rolls, 1995, 1996, 2008a, 2010; Treves and Rolls, 1994) is considered here in the context of recalling a complete memory representation in the complete set of cortical areas that provide inputs to the hippocampus (see Fig. 1).

It is suggested that the modifiable connections from the CA3 neurons to the CA1 neurons allow the whole episode in CA3 to be produced in CA1. This may be assisted by the direct perforant path input to CA1 (Treves and Rolls, 1994). This might allow details of the input key for the recall process, as well as the possibly less information-rich memory of the whole episode recalled from the CA3 network, to contribute to the firing of CA1 neurons. The CA1 neurons would then activate, via their termination in the deep layers of the entorhinal cortex, at least the pyramidal cells in the deep layers of the entorhinal cortex (see Fig. 1). These entorhinal cortex layer 5 neurons would then, by virtue of their backprojections (Lavenex and Amaral, 2000; Witter et al., 2000) to the parts of cerebral cortex that originally provided the inputs to the hippocampus, terminate in the superficial layers (including layer 1) of those neocortical areas, where synapses would be made onto the distal parts of the dendrites of the (superficial and deep) cortical pyramidal cells (Markov et al., 2014; Rolls, 1989a,b,d). The areas of cerebral neocortex in which this recall would be produced could include multimodal cortical areas (e.g., the cortex in the superior temporal sulcus which receives inputs from temporal, parietal, and occipital cortical areas, and from which it is thought that cortical areas such as 39 and 40 related to language developed; and the orbitofrontal and anterior cingulate cortex to retrieve the reward/affective aspects of an episodic memory; Rolls, 2014a,b) and also areas of unimodal association cortex (e.g., inferior temporal visual cortex; Lavenex and Amaral, 2000). The backprojections, by recalling previous episodic events, could provide information useful to the neocortex in the building of new representations

in the multimodal and unimodal association cortical areas, which by building new long-term representations (sometimes called schemas; [Preston and Eichenbaum, 2013](#)) can be considered as a form of memory consolidation ([Rolls, 1989a,b,d, 1990a,b, 2008a](#)), or in organizing actions.

The hypothesis of the architecture with which this would be achieved is shown in [Fig. 1](#). The feedforward connections from association areas of the cerebral neocortex (solid lines in [Fig. 1](#)) show major convergence as information is passed to CA3, with the CA3 autoassociation network having the smallest number of neurons at any stage of the processing. The backprojections allow for divergence back to neocortical areas. The way in which we suggest that the backprojection synapses are set up to have the appropriate strengths for recall is as follows ([Rolls, 1989a,b,d, 2008a](#)). During the setting up of a new episodic memory, there would be strong feedforward activity progressing toward the hippocampus. During the episode, the CA3 synapses would be modified, and via the CA1 neurons and the subiculum, a pattern of activity would be produced on the backprojecting synapses to the entorhinal cortex. Here, the backprojecting synapses from active backprojection axons onto pyramidal cells being activated by the forward inputs to entorhinal cortex would be associatively modified. A similar process would be implemented at preceding stages of neocortex, that is, in the parahippocampal gyrus/perirhinal cortex stage and in association cortical areas, as shown in [Fig. 1](#).

The concept is that during the learning of an episodic memory, cortical pyramidal cells in at least one of the stages would be driven by forward inputs but would simultaneously be receiving backprojected activity (indirectly) from the hippocampus, which would, by pattern association from the backprojecting synapses to the cortical pyramidal cells, become associated with whichever cortical cells were being made to fire by the forward inputs. Then later on, during recall, a recall cue from perhaps another part of cortex might reach CA3, where the firing during the original episode would be completed. The resulting backprojecting activity would then, as a result of the pattern association learned previously, bring back the firing in any cortical area that was present during the original episode. Thus, retrieval involves reinstating the activity that was present in different cortical areas that was present during the learning of an episode. (The pattern association is also called heteroassociation, to contrast it with autoassociation. The pattern association operates at multiple stages in the backprojection pathway, as made evident in [Fig. 1](#)). If the recall cue was an object, this might result in recall of the neocortical firing that represented the place in which that object had been seen previously. As noted elsewhere in this chapter and by [McClelland et al. \(1995\)](#), that recall might be useful to the neocortex to help it build new semantic memories, which might inherently be a slow process and is not part of the theory of recall. It is an interesting possibility that recall might involve several cycles through the recall process. After the information fed back from the first pass with a recall cue from perhaps only one cortical area, information might gradually be retrieved to other cortical areas involved in the original memory, and this would then act as a better retrieval cue for the next pass.

The timing of the backprojecting activity would be sufficiently rapid, in that, for example, inferior temporal cortex neurons become activated by visual stimuli with latencies of 90–110 ms and may continue firing for several hundred milliseconds (Rolls, 1992), and hippocampal pyramidal cells are activated in visual object-and-place and conditional spatial response tasks with latencies of 120–180 ms (Miyashita et al., 1989; Rolls and Xiang, 2006; Rolls et al., 1989). Thus, backprojected activity from the hippocampus might be expected to reach association cortical areas such as the inferior temporal visual cortex within 60~100 ms of the onset of their firing, and there would be a several hundred milliseconds period in which there would be conjunctive feedforward activation present with simultaneous backprojected signals in the association cortex.

During recall, the backprojection connections onto the distal synapses of cortical pyramidal cells would be helped in their efficiency in activating the pyramidal cells by virtue of two factors. The first is that with no forward input to the neocortical pyramidal cells, there would be little shunting of the effects received at the distal dendrites by the more proximal effects on the dendrite normally produced by the forward synapses. Further, without strong forward activation of the pyramidal cells, there would not be very strong feedback and feedforward inhibition via GABA cells, so that there would not be a further major loss of signal due to (shunting) inhibition on the cell body and (subtractive) inhibition on the dendrite. (The converse of this is that when forward inputs are present, as during normal processing of the environment rather than during recall, the forward inputs would, appropriately, dominate the activity of the pyramidal cells, which would be only influenced, not determined, by the backprojecting inputs; see Deco and Rolls, 2005b; Rolls, 1989b,d, 2008a).

The synapses receiving the backprojections would have to be Hebb-modifiable, as suggested by Rolls (1989b,d). This would solve the deaddressing problem, which is the problem of how the hippocampus is able to bring into activity during recall just those cortical pyramidal cells that were active when the memory was originally being stored. The solution hypothesized (Rolls, 1989b,d) arises because modification occurs during learning of the synapses from active backprojecting neurons from the hippocampal system onto the dendrites of only those neocortical pyramidal cells active at the time of learning. Without this modifiability of cortical backprojections during learning at some cortical stages at least, it is difficult to see how exactly the correct cortical pyramidal cells active during the original learning experience would be activated during recall. Consistent with this hypothesis (Rolls, 1989b,d), there are NMDA receptors present especially in superficial layers of the cerebral cortex (Monaghan and Cotman, 1985), implying Hebb-like learning just where the backprojecting axons make synapses with the apical dendrites of cortical pyramidal cells. The quantitative argument is that the backprojecting synapses in at least one stage have to be associatively modifiable parallels that applied to the pattern retrieval performed at the entorhinal to CA3 synapses (Treves and Rolls, 1992) and at the CA3–CA1 synapses (Schultz and Rolls, 1999) and is that the information retrieved would otherwise be very low. The performance of pattern association networks is considered in detail by Rolls and Treves (Rolls, 2008a; Rolls and Treves, 1990,

1998) and other authors (Hertz et al., 1991). It is also noted that the somewhat greater anatomical spread of the backprojection than the forward connections between two different stages in the hierarchy shown in Fig. 1 would not be a problem, for it would provide every chance for the backprojecting axons to find co-active neurons in an earlier cortical stage that are part of the representation that is relevant to the current memory being formed.

If the backprojection synapses are associatively modifiable, we may consider the duration of the period for which their synaptic modification should persist. What follows from the operation of the system described above is that there would be no point, indeed it would be disadvantageous, if the synaptic modifications lasted for longer than the memory remained in the hippocampal buffer store. What would be optimal would be to arrange for the associative modification of the backprojecting synapses to remain for as long as the memory persists in the hippocampus. This suggests that a similar mechanism for the associative modification within the hippocampus and for that of at least one stage of the backprojecting synapses would be appropriate. It is suggested that the presence of high concentrations of NMDA synapses in the distal parts of the dendrites of neocortical pyramidal cells and within the hippocampus may reflect the similarity of the synaptic modification processes in these two regions (cf. Kirkwood et al., 1993). It is noted that it would be appropriate to have this similarity of time course (i.e., rapid learning within 1–2 s, and slow decay over perhaps weeks) for at least one stage in the series of backprojecting stages from the CA3 region to the neocortex. Such stages might include the CA1 region, subiculum, entorhinal cortex, and perhaps the parahippocampal gyrus/perirhinal cortex. However, from multimodal cortex (e.g., the parahippocampal gyrus) back to earlier cortical stages, it might be desirable for the backprojecting synapses to persist for a long period, so that some types of recall and top-down processing (Rolls, 1989b,d, 2008a; Rolls and Deco, 2002) mediated by the operation of neocortico-neocortical backprojecting synapses could be stable and might not require modification during the learning of a new episodic memory.

An alternative hypothesis to that above is that rapid modifiability of backprojection synapses would be required only at the beginning of the backprojecting stream. Relatively fixed associations from higher to earlier neocortical stages would serve to activate the correct neurons at earlier cortical stages during recall. For example, there might be rapid modifiability from CA3 to CA1 neurons, but relatively fixed connections from there back (McClelland et al., 1995). For such a scheme to work, one would need to produce a theory not only of the formation of semantic memories in the neocortex but also of how the operations performed according to that theory would lead to recall by setting up appropriately the backprojecting synapses.

We have noted elsewhere that backprojections, which included corticocortical backprojections, and backprojections originating from structures such as the hippocampus and amygdala, may have a number of different functions (Rolls, 1989a,b,d, 1990a,b, 1992, 2005, 2008a; Rolls and Deco, 2002) including implementing top-down attention by biased competition (Deco and Rolls, 2003, 2004, 2005a; Deco et al., 2005; Grabenhorst and Rolls, 2010; Rolls, 2008a,b, 2013a; Rolls and Deco,

2002, 2006). The particular function with which we have been concerned here is how memories stored in the hippocampus might be recalled in regions of the cerebral neocortex, and this is not at all incompatible with such theories of top-down attentional control.

3.2 BACKPROJECTIONS TO THE NEOCORTEX—THE LARGE NUMBER OF CONNECTIONS ONTO EACH NEURON

How many backprojecting fibers does one need to synapse on any given neocortical pyramidal cell in order to implement the mechanism outlined above? Consider a polysynaptic sequence of backprojecting stages, from hippocampus to neocortex, as a series of simple (hetero-)associative (i.e., pattern association) memories in which, at each stage, the input lines are those coming from the previous stage (closer to the hippocampus; Rolls, 2008a; Treves and Rolls, 1994; Fig. 1). (The interesting concept here is that one can treat for a capacity analysis the series of backprojection stages to the cerebral cortex, which each involves a pattern association, as an “unrolled” version of an autoassociator. Each backprojection pattern association stage would correspond to one iteration round the autoassociation system.) Implicit in this framework is the assumption that the synapses at each stage are modifiable and have been indeed modified at the time of first experiencing each episode, according to some Hebbian associative plasticity rule. A plausible requirement for a successful hippocampodirected recall operation is that the signal generated from the hippocampally retrieved pattern of activity, and carried backward toward neocortex, remains undegraded when compared to the noise due, at each stage, to the interference effects caused by the concurrent storage of other patterns of activity on the same backprojecting synaptic systems. That requirement is equivalent to that used in deriving the storage capacity of such a series of heteroassociative memories, and it was shown in Treves and Rolls (1991) that the maximum number of independently generated activity patterns that can be retrieved is given, essentially, by the same formula as (1) above where, however, a is now the sparseness of the representation at any given stage, and C is the average number of (back-)projections each cell of that stage receives from cells of the previous one (Treves and Rolls, 1991). In more detail, the number of memory patterns p that can be retrieved in a multistage pattern association network is

$$p \cong \frac{C}{a \ln(1/a)} k' \quad (2)$$

where k is a factor that depends weakly on the detailed structure of the rate distribution, on the connectivity pattern, etc., but is roughly in the order of 0.2-0.3 (Treves and Rolls, 1991). (This result for the storage capacity is derived using threshold linear neurons as model M2 in the appendix of Treves and Rolls (1991). The storage capacity of a one-stage pattern association network is similar, as derived by Rolls and Treves (1990), where a there refers to the sparseness of the output representation Rolls, 2008a, 2014a.)

If p is equal to the number of memories held in the hippocampal memory, it is limited by the retrieval capacity of the CA3 network, p_{\max} . Putting together the formula for the latter with that shown here (2), one concludes that, roughly, the requirement implies that the number of afferents of (indirect) hippocampal origin to a given neocortical stage (C^{HBP}), must be $C^{\text{HBP}} = C^{\text{RC}} a_{\text{nc}}/a_{\text{CA3}}$, where C^{RC} is the number of recurrent collaterals to any given cell in CA3, a_{nc} is the average sparseness of a neocortical representation, and a_{CA3} is the sparseness of memory representations in CA3 (Treves and Rolls, 1994).

The above requirement is very strong: even if representations were to remain as sparse as they are in CA3, which is unlikely, to avoid degrading the signal, C^{HBP} should be as large as C^{RC} , i.e., 12,000 in the rat. Moreover, other sources of noise not considered in the present calculation would add to the severity of the constraint and partially compensate for the relaxation in the constraint that would result from requiring that only a fraction of the p episodes would involve any given cortical area. If then C^{HBP} has to be of the same order as C^{RC} , one is led to a very definite conclusion: a mechanism of the type envisaged here could not possibly rely on a set of monosynaptic CA3-to-neocortex backprojections. This would imply that, to make a sufficient number of synapses on each of the vast number of neocortical cells, each cell in CA3 has to generate a disproportionate number of synapses onto neocortical neurons (i.e., C^{HBP} times the ratio between the number of neocortical and that of CA3 cells). The required divergence can be kept within reasonable limits only by assuming that the backprojecting system is polysynaptic, provided that the number of cells involved grows gradually at each stage, from CA3 back to neocortical association areas (Treves and Rolls, 1994; cf. Fig. 1).

Although backprojections between any two adjacent areas in the cerebral cortex are approximately as numerous as forward projections, and much of the distal parts of the dendrites of cortical pyramidal cells are devoted to backprojections, the actual number of such connections onto each pyramidal cell may be on average only in the order of thousands. Further, not all might reflect backprojection signals originating from the hippocampus, for there are backprojections which might be considered to originate in the amygdala (see Amaral et al., 1992) or in multimodal cortical areas (allowing, for example, for recall of a visual image by an auditory stimulus with which it has been regularly associated). In this situation, one may consider whether the backprojections from any one of these systems would be sufficiently numerous to produce recall. One factor which may help here is that when recall is being produced by the backprojections, it may be assisted by the local recurrent collaterals between nearby (~ 1 mm) pyramidal cells, which are a feature of neocortical connectivity. These would tend to complete a partial neocortical representation being recalled by the backprojections into a complete recalled pattern. (Note that this completion would be only over the local information present within a cortical area about, e.g., visual input *or* spatial input; it provides a local “clean-up” mechanism and could not replace the global autoassociation performed effectively over the activity of very many cortical areas which the CA3 could perform by virtue of its widespread recurrent collateral connectivity.) There are two alternative possibilities about how this

would operate. First, if the recurrent collaterals showed slow and long-lasting synaptic modification, then they would be useful in completing the whole of long-term (e.g., semantic) memories. Second, if the neocortical recurrent collaterals showed rapid changes in synaptic modifiability with the same time course as that of hippocampal synaptic modification, then they would be useful in filling in parts of the information forming episodic memories, which could be made available locally within an area of the cerebral neocortex.

This theory of recall by the backprojections thus provides a quantitative account of why the cerebral cortex has as many backprojection connections as forward projection connections (Rolls, 2008a).

These concepts show how the backprojection system to neocortex can be conceptualized in terms of pattern completion, as follows. The information that is present when a memory is formed may be present in different areas of the cerebral cortex, for example, of a face in a temporal cortex face area (Rolls, 2012b), of a spatial location in a neocortical location area, and of a reward received in the orbitofrontal cortex (Rolls, 2014a). To achieve detailed retrieval of the memory, reinstatement of the activity during recall of the neuronal activity during the original memory formation may be needed. This is what the backprojection system described could achieve and is a form of completion of the information that was represented in the different cortical areas when the memory was formed. In particular, the concept of completion here is that if a recall cue from a visual object area is provided, then the emotional parts of the episodic memory can be recalled in the orbitofrontal cortex, and the spatial parts in parietal cortical areas, with the result that a complete memory is retrieved, with activity recalled into several higher-order cortical areas. Because such a wide set of different neocortical areas must be content-addressed, a multistage feedback system is required to keep the number of synapses per neuron in the backprojection pathways down to reasonable numbers. (Having CA1 directly address neocortical areas would require each CA1 neuron to have tens of millions of synapses with cortical neurons. That is part of the computational problem solved by the multistage backprojection system shown in Fig. 1.) Thus, the backprojection system with its series of pattern associators can each be thought of as retrieving the complete pattern of cortical activity in many different higher-order cortical areas that was present during the original formation of the episodic memory.

3.3 DILUTION IN PATTERN ASSOCIATION NETWORKS AND ITS ROLE IN HIPPOCAMPO-NEOCORTICAL RECALL OF MEMORIES STORED IN THE HIPPOCAMPUS

The new hypotheses described and tested in this chapter are that having multiple connections between the inputs and the output neurons can decrease the capacity of pattern association networks; that dilution of connectivity in pattern association networks can minimize this loss of capacity by reducing the probability of multiple synapses if they are made at random between input and output neurons; and that this dilution helps in this way to ensure that the recall of information from the

hippocampus to the neocortex by several stages of pattern association network is efficient and has high capacity. It is further suggested that the accuracy and capacity of the recall process is helped by autoassociation implemented by the recurrent collateral connections between nearby neocortical pyramidal cells, which has previously been shown to be beneficial in autoassociation networks (Rolls, 2012a). The new analyses described here show that a similar process applies in pattern association networks.

The hypotheses were tested by simulations of a pattern association network, the architecture of which is shown in Fig. 4, and the properties of which have been described elsewhere (Rolls, 2008a; with the relevant Appendices available online at www.oxcns.org/papers/RollsMemoryAttentionAndDecisionMakingContents+Appendices1+2.pdf). Random binary pattern vectors of firing rates were used for the inputs to the network (default sparseness 0.05) and for the outputs of the network (default sparseness 0.05). The network was trained with an associative (Hebbian) rule:

$$\delta w_{ij} = k \cdot x_j \cdot y_i \quad (3)$$

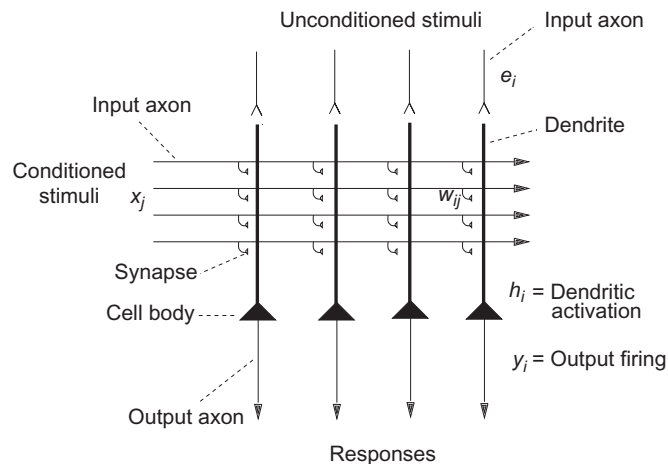


FIGURE 4

A pattern association memory. An unconditioned stimulus has activity or firing rate e_i for the i th neuron and produces firing y_i of the i th neuron. The conditioned stimuli have activity or firing rate x_j for the j th axon. In the context of a pattern association network in the hippocampo-cortical backprojection system, the following correspondences apply. The unconditioned stimulus input is the firing of the postsynaptic neuron during learning of the episodic memory. The conditioned stimulus input is the backprojection input originating in the hippocampus. The output neurons, or neurons connected to them within a cortical area, send back their outputs to act as the recall cue for the preceding cortical area.

where k is a constant, y_i is the activation or firing rate of the postsynaptic neuron, x_j is the presynaptic firing rate, and δw_{ij} is the change of synaptic weight (Rolls, 2008a). During recall, the activations of the neurons h_i were calculated using a dot product between the presynaptic firing rate vector, the conditioned stimulus or recall cue, and the synaptic weights on the dendrite of each neuron as:

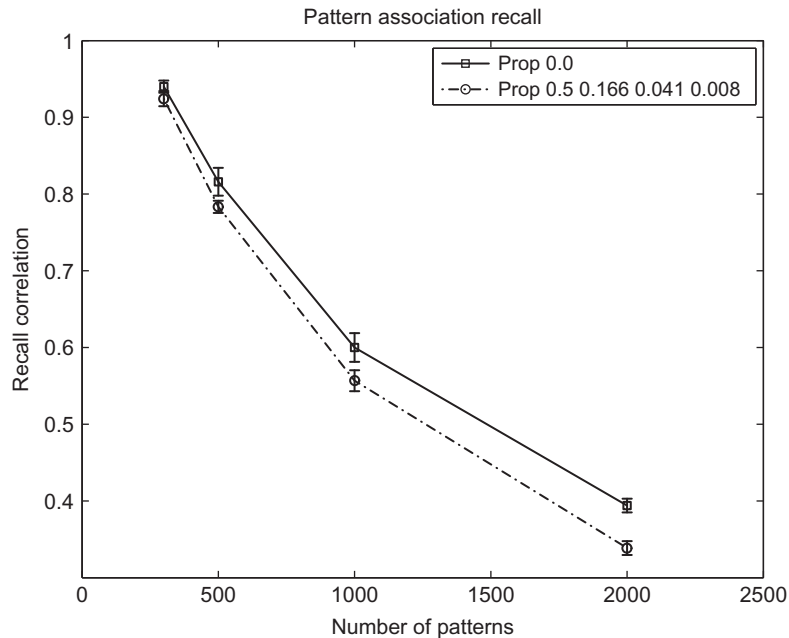
$$h_i = \sum_j x_j \cdot w_{ij} \quad (4)$$

and converted to binary 1,0 firing rates r_i with a sparseness a_o (default 0.05) by setting the threshold appropriately. The performance of the network was measured by the (Pearson) correlation between the output vector of firing rates \mathbf{r} and the trained vector of output firing rates for each conditioned stimulus input pattern.

The hypothesis about the effects of having a proportion of the synapses randomly selected with double strength due to two synapses from an input neuron to an output neuron was tested as follows. The fully connected network was first trained on a number $nPatts$ of input–output pattern pairs. Then a proportion of the synapses, reflecting what might be expected from a Poisson distribution with $\lambda = 1$, the mean number of connections from an input to an output neuron (Rolls, 2012a) onto each neuron selected by random permutation was doubled, tripled, etc. in strength, to reflect what would happen if there were two or three, etc., synapses present as specified by a Poisson distribution (see Fig. 5). Illustrative results are shown in Fig. 5. Recall was considerably impaired, measured by the reduction in the correlation between the recalled outputs firing rate vectors and the trained firing rate output vectors, if some multiple synapses were present as might be expected from a Poisson distribution. The effect occurs because the distortion of the synaptic weights can affect the recalled pattern, with the effect particularly evident when the loading is high, and there is interference between the different pattern pairs associated together in the synaptic matrix. In control simulations, it was shown that if the number of double synapses was set to the number expected for a dilution of $\lambda = 0.1$, then the much smaller number of double synapses produced by chance had only small effects on performance, confirming the usefulness of diluted connectivity.

In addition to this finding, before each recall pattern was applied, a small proportion of the elements in the input recall vector were flipped from 1 to 0 or from 0 to 1 using random permutation, to introduce some noise into the recall cue, to require generalization of the pattern association network toward the correct output firing (Rolls, 2008a), and to reflect what might be imperfect recall within the hippocampus or the preceding stage in the recall process. This small variation of the recall cue contributes to the effects described at somewhat lower loadings (measured by the number of stored patterns), for it results in the recall cue perhaps not activating a (strong) double synapse which should have been activated, or in activating a (strong) double synapse that should not have been activated, resulting in distortion of the pattern recalled from the pattern association network.

Thus, having a proportion of synapses on an output neuron arising from the same input neuron in a pattern association network can impair the retrieval of information from the pattern association network. This provides a potentially fundamental

**FIGURE 5**

Effects of multiple synapses from an input neuron to an output neuron in a pattern association network. In the simulations with random binary input and output patterns each with sparseness $a=0.05$, there were $N=1000$ output neurons and $C=1000$ synaptic inputs onto each neuron. In the case representing the effects of some multiple synapses between input neurons and output neurons due to random connectivity and modeled with a Poisson distribution, the proportion of double-strength synapses was 0.5, of triple strength was 0.166, of quadruple strength was 0.041, and of quintuple strength was 0.008. The case for single synapses is labeled Prop 0.0. Performance was measured by the correlation between each recalled output vector and the corresponding output vector that had been trained. Each data point shows the mean and standard deviation calculated over 5–10 simulations with different random binary pattern vectors to be associated. The effects are shown for different numbers of trained pattern associations. The effects illustrated were without the introduction of noise into the recall vector.

approach to why connectivity in cortical systems is diluted, to be typically <0.1 , which results in the proportion of synapses that are replicated between any pair of neurons being sufficiently low that retrieval is not impaired.

In the hippocampo-cortical backprojection system used for recall in which there are multiple stages of pattern association networks (Fig. 1), any distortion of what is recalled at any one stage (Fig. 5) will be magnified and exacerbated by the distortion being further produced at each stage of the recall process. This means that the effect described here is likely to be very relevant indeed in the recall of memories from the

hippocampus to the neocortex and to provide an important reason why this connectivity must be diluted. Another advantage of the dilution in this connectivity is that it reflects the divergence from a relatively small number of CA3 neurons (300,000 in the rat) to the enormous numbers of neocortical neurons, which can be achieved by making this a multistage recall pathway with diluted connecting and with approximately 10,000 backprojection synapses onto each receiving neuron from the previous stage.

In the context of any distortion produced at each stage of the hippocampo-cortical recall process, part of the hypothesis is that local recurrent collaterals between nearby neocortical neurons within each stage will operate as an autoassociation network to clean up the retrieved pattern, and it has been shown that this effect operates sufficiently fast (within 20 ms per stage) with integrate-and-fire neurons to contribute to the correction of the retrieval process (Panzeri et al., 1998).

The findings described here on the advantages of dilution in the connectivity of pattern associators in the cerebral cortex provide a fundamental advance in understanding cortical computation, for pattern association networks appear to be used in very many cortical systems, including the corticocortical backprojections used in memory recall and in top-down selective attention (Deco and Rolls, 2005b; Rolls, 2008a,b, 2013a), but also in systems in the orbitofrontal cortex and amygdala involved in stimulus-reinforcement association learning for emotion (Rolls, 2014a,b). The new evidence adds to the advantages of dilution in the connectivity of attractor networks in the cerebral cortex, where multiple connections between any pair of neurons potentially distort the basins of attraction, and thereby impair memory recall, short-term memory, etc. (Rolls, 2012a).

In another class of network believed to be fundamental to cortical computation, competitive networks, which can build new representations (Rolls, 2008a, 2012b), dilution in connectivity may be advantageous, but for very different reasons. Dilution in the connectivity of competitive networks can help to break the symmetry, so that some neurons are likely to be allocated to some patterns, and other neurons to other patterns (because of the particular set of input connections of each neuron), and this can help to stabilize a competitive network, by making it difficult for neurons to drift during further learning if the input patterns drift (Rolls, 2008a). In such a system, multiple connections from an input neuron to an output neuron might even be advantageous, for the reasons just given. This emphasizes the point that multiple connections between neurons are found in some parts of the cortex, including, for example, thalamocortical inputs, where a strong selective drive to some neurons may be important, and the presence of such multiple contacts involving some cortical neurons does not invalidate at all the hypotheses and arguments presented here. The points made here are that in pattern association systems in the brain, as well as in autoassociation systems in the brain, used for processes requiring high capacity such as memory storage and recall, there are strong disadvantages to multiple contacts between pairs of neurons, and dilution of connectivity is one way of minimizing these effects. Further effects of dilution in connectivity are described elsewhere (Rolls and Webb, 2012), as is some of the evidence for dilution in cortical connectivity (Rolls, 2008a).

4 DISCUSSION

A theory closely related to the present theory of how the hippocampus operates has been developed by [McClelland et al. \(1995\)](#). It is very similar to the theory we have developed ([Rolls, 1987, 1989a,b,d, 2008a, 2010, 2013b,c](#); [Treves and Rolls, 1992, 1994](#)) at the systems level, except that it takes a stronger position on the gradient of retrograde amnesia, emphasizes that recall from the hippocampus of episodic information is used to help build semantic representations in the neocortex, and holds that the last set of synapses that are modified rapidly during the learning of each episode are those between the CA3 and the CA1 pyramidal cells (see [Fig. 1](#)). It also emphasizes the important point that the hippocampal and neocortical memory systems may be quite different, with the hippocampus specialized for the rapid learning of single events or episodes, and the neocortex for the slower learning of semantic representations, which may necessarily benefit from the many exemplars needed to shape the semantic representation. In the formulation by [McClelland et al. \(1995\)](#), the entorhinal cortex connections via the perforant path onto the CA1 cells are non-modifiable (in the short term) and allow a representation of neocortical long-term memories to activate the CA1 cells. The new information learned in an episode by the CA3 system is then linked to existing long-term memories by the CA3–CA1 rapidly modifiable synapses. The new hypotheses described here about the role of dilution in pattern association networks would apply to the CA3–CA1 connection in their conceptualization. All the connections from the CA1 back via the subiculum, entorhinal cortex, parahippocampal cortex, etc., to the association neocortex are held to be unmodifiable in the short term, during the formation of an episodic memory, but the same new hypothesis about the advantages of dilution pattern association networks would apply to these backprojections.

In conclusion, our theory of the hippocampus ([Kesner and Rolls, 2015](#); [Rolls, 2008a, 2010](#); [Treves and Rolls, 1994](#)) has been extended here by showing how dilution in the connectivity of pattern association networks in the backprojection system from the hippocampus to the neocortex could facilitate the recall of memories from the hippocampus by reducing the proportion of multiple connections between neurons in the backprojecting pathways with the impairment that this would produce in the pattern association recall.

ACKNOWLEDGMENTS

Different parts of the research described here were supported by Programme Grants from the Medical Research Council, by a Human Frontier Science program grant, by an EEC BRAIN grant, by the MRC Oxford Interdisciplinary Research Centre in Cognitive Neuroscience, by the Oxford McDonnell-Pew Centre in Cognitive Neuroscience, and by the Oxford Centre for Computational Neuroscience. The author has performed the experimental and theoretical work that is incorporated in some of the ideas presented here on the hippocampus with many colleagues, including Alessandro Treves, Simon Stringer, Robert Robertson, Pierre Georges-François, Shane O'Mara, Ray Kesner, Yasushi Miyashita, Janet Feigenbaum, Ivan de Araujo,

Simon Schultz, Stefano Panzeri, David Foster, Peter Cahusac, Hiroaki Niki, and Alain Berthoz, and their contributions are sincerely acknowledged.

REFERENCES

- Amaral, D.G., Price, J.L., Pitkanen, A., Carmichael, S.T., 1992. Anatomical organization of the primate amygdaloid complex. In: Aggleton, J.P. (Ed.), *The Amygdala*. Wiley-Liss, New York, pp. 1–66.
- Amit, D.J., 1989. *Modeling Brain Function*. Cambridge University Press, Cambridge.
- Barkas, L.J., Henderson, J.L., Hamilton, D.A., Redhead, E.S., Gray, W.P., 2010. Selective temporal resections and spatial memory impairment: cue dependent lateralization effects. *Behav. Brain Res.* 208, 535–544.
- Bonelli, S.B., Powell, R.H., Yogarajah, M., Samson, R.S., Symms, M.R., Thompson, P.J., Koepp, M.J., Duncan, J.S., 2010. Imaging memory in temporal lobe epilepsy: predicting the effects of temporal lobe resection. *Brain* 133, 1186–1199.
- Deco, G., Rolls, E.T., 2003. Attention and working memory: a dynamical model of neuronal activity in the prefrontal cortex. *Eur. J. Neurosci.* 18, 2374–2390.
- Deco, G., Rolls, E.T., 2004. A neurodynamical cortical model of visual attention and invariant object recognition. *Vis. Res.* 44, 621–644.
- Deco, G., Rolls, E.T., 2005a. Attention, short-term memory, and action selection: a unifying theory. *Prog. Neurobiol.* 76, 236–256.
- Deco, G., Rolls, E.T., 2005b. Neurodynamics of biased competition and co-operation for attention: a model with spiking neurons. *J. Neurophysiol.* 94, 295–313.
- Deco, G., Rolls, E.T., Zihl, J., 2005. A neurodynamical model of visual attention. In: Itti, L., Rees, G., Tsotsos, J. (Eds.), *Neurobiology of Attention*. Elsevier, San Diego, CA, pp. 593–599 (Chapter 597).
- Georges-François, P., Rolls, E.T., Robertson, R.G., 1999. Spatial view cells in the primate hippocampus: allocentric view not head direction or eye position or place. *Cereb. Cortex* 9, 197–212.
- Grabenhorst, F., Rolls, E.T., 2010. Attentional modulation of affective vs sensory processing: functional connectivity and a top-down biased activation theory of selective attention. *J. Neurophysiol.* 104, 1649–1660.
- Hartley, T., Lever, C., Burgess, N., O’Keefe, J., 2014. Space in the brain: how the hippocampal formation supports spatial cognition. *Phil. Trans. Roy. Soc. Lond.* 369, 20120510.
- Hertz, J., Krogh, A., Palmer, R.G., 1991. *An Introduction to the Theory of Neural Computation*. Addison-Wesley, Wokingham.
- Hopfield, J.J., 1982. Neural networks and physical systems with emergent collective computational abilities. *Proc. Natl. Acad. Sci. U. S. A.* 79, 2554–2558.
- Kesner, R.P., Rolls, E.T., 2015. A computational theory of hippocampal function, and tests of the theory: new developments. *Neurosci. Biobehav. Rev.* 48, 92–147.
- Kirkwood, A., Dudek, S.M., Gold, J.T., Aizenman, C.D., Bear, M.F., 1993. Common forms of synaptic plasticity in the hippocampus and neocortex in vitro. *Science* 260, 1518–1521.
- Lavenex, P., Amaral, D.G., 2000. Hippocampal-neocortical interaction: a hierarchy of associativity. *Hippocampus* 10, 420–430.
- Markov, N.T., Vezoli, J., Chameau, P., Falchier, A., Quilodran, R., Huissoud, C., Lamy, C., Misery, P., Giroud, P., Ullman, S., Barone, P., Dehay, C., Knoblauch, K., Kennedy, H.,

2014. Anatomy of hierarchy: feedforward and feedback pathways in macaque visual cortex. *J. Comp. Neurol.* 522, 225–259.
- Marr, D., 1971. Simple memory: a theory for archicortex. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 262, 23–81.
- McClelland, J.L., McNaughton, B.L., O'Reilly, R.C., 1995. Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. *Psychol. Rev.* 102, 419–457.
- McNaughton, B.L., Morris, R.G.M., 1987. Hippocampal synaptic enhancement and information storage within a distributed memory system. *Trends Neurosci.* 10, 408–415.
- Miyashita, Y., Rolls, E.T., Cahusac, P.M., Niki, H., Feigenbaum, J.D., 1989. Activity of hippocampal formation neurons in the monkey related to a conditional spatial response task. *J. Neurophysiol.* 61, 669–678.
- Monaghan, D.T., Cotman, C.W., 1985. Distribution of N-methyl-D-aspartate-sensitive L-[3H] glutamate-binding sites in the rat brain. *J. Neurosci.* 5, 2909–2919.
- Nakazawa, K., Quirk, M.C., Chitwood, R.A., Watanabe, M., Yeckel, M.F., Sun, L.D., Kato, A., Carr, C.A., Johnston, D., Wilson, M.A., Tonegawa, S., 2002. Requirement for hippocampal CA3 NMDA receptors in associative memory recall. *Science* 297, 211–218.
- Nakazawa, K., Sun, L.D., Quirk, M.C., Rondi-Reig, L., Wilson, M.A., Tonegawa, S., 2003. Hippocampal CA3 NMDA receptors are crucial for memory acquisition of one-time experience. *Neuron* 38, 305–315.
- Nakazawa, K., McHugh, T.J., Wilson, M.A., Tonegawa, S., 2004. NMDA receptors, place cells and hippocampal spatial memory. *Nat. Rev. Neurosci.* 5, 361–372.
- O'Keefe, J., Dostrovsky, J., 1971. The hippocampus as a spatial map: preliminary evidence from unit activity in the freely moving rat. *Brain Res.* 34, 171–175.
- O'Keefe, J., 1990. A computational theory of the hippocampal cognitive map. *Progress in Brain Res.* 83, 301–312.
- Panzeri, S., Battaglia, F.P., Rolls, E.T., Lavis, R., Giles, J., Treves, A., 1998. Speed of information retrieval in multilayer networks with spiking neurons. *Soc. Neurosci. Abstr.* 24, 2095.
- Preston, A.R., Eichenbaum, H., 2013. Interplay of hippocampus and prefrontal cortex in memory. *Curr. Biol.* 23, R764–R773.
- Robertson, R.G., Rolls, E.T., Georges-François, P., 1998. Spatial view cells in the primate hippocampus: effects of removal of view details. *J. Neurophysiol.* 79, 1145–1156.
- Rolls, E.T., 1987. Information representation, processing and storage in the brain: analysis at the single neuron level. In: Changeux, J.-P., Konishi, M. (Eds.), *The Neural and Molecular Bases of Learning*. Wiley, Chichester, pp. 503–540.
- Rolls, E.T., 1989a. Functions of neuronal networks in the hippocampus and cerebral cortex in memory. In: Cotterill, R.M.J. (Ed.), *Models of Brain Function*. Cambridge University Press, Cambridge, pp. 15–33.
- Rolls, E.T., 1989b. Functions of neuronal networks in the hippocampus and neocortex in memory. In: Byrne, J.H., Berry, W.O. (Eds.), *Neural Models of Plasticity: Experimental and Theoretical Approaches*. Academic Press, San Diego, CA, pp. 240–265.
- Rolls, E.T., 1989c. Parallel distributed processing in the brain: implications of the functional architecture of neuronal networks in the hippocampus. In: Morris, R.G.M. (Ed.), *Parallel Distributed Processing: Implications for Psychology and Neurobiology*. Oxford University Press, Oxford, pp. 286–308.
- Rolls, E.T., 1989d. The representation and storage of information in neuronal networks in the primate cerebral cortex and hippocampus. In: Durbin, R., Miall, C., Mitchison, G. (Eds.), *The Computing Neuron*. Addison-Wesley, Wokingham, pp. 125–159.

- Rolls, E.T., 1990a. Functions of the primate hippocampus in spatial processing and memory. In: Olton, D.S., Kesner, R.P. (Eds.), *Neurobiology of Comparative Cognition*. Lawrence Erlbaum Associates, Hillsdale, NJ, pp. 339–362.
- Rolls, E.T., 1990b. Theoretical and neurophysiological analysis of the functions of the primate hippocampus in memory. *Cold Spring Harb. Symp. Quant. Biol.* 55, 995–1006.
- Rolls, E.T., 1992. Neurophysiological mechanisms underlying face processing within and beyond the temporal cortical visual areas. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 335, 11–21.
- Rolls, E.T., 1995. A model of the operation of the hippocampus and entorhinal cortex in memory. *Int. J. Neural Syst.* 6, 51–70.
- Rolls, E.T., 1996. A theory of hippocampal function in memory. *Hippocampus* 6, 601–620.
- Rolls, E.T., 2005. *Emotion Explained*. Oxford University Press, Oxford.
- Rolls, E.T., 2008a. *Memory, Attention, and Decision-Making: A Unifying Computational Neuroscience Approach*. Oxford University Press, Oxford.
- Rolls, E.T., 2008b. Top-down control of visual perception: attention in natural vision. *Perception* 37, 333–354.
- Rolls, E.T., 2010. A computational theory of episodic memory formation in the hippocampus. *Behav. Brain Res.* 215, 180–196.
- Rolls, E.T., 2012a. Advantages of dilution in the connectivity of attractor networks in the brain. *Biologically Inspired Cognitive Architectures* 1, 44–54.
- Rolls, E.T., 2012b. Invariant visual object and face recognition: neural and computational bases, and a model, *VisNet*. *Front. Comput. Neurosci.* 6 (35), 1–70.
- Rolls, E.T., 2013a. A biased activation theory of the cognitive and attentional modulation of emotion. *Front. Hum. Neurosci.* 7 (74), 1–15.
- Rolls, E.T., 2013b. The mechanisms for pattern completion and pattern separation in the hippocampus. *Front. Syst. Neurosci.* 7, 74.
- Rolls, E.T., 2013c. A quantitative theory of the functions of the hippocampal CA3 network in memory. *Front. Cell. Neurosci.* 7, 98.
- Rolls, E.T., 2014a. *Emotion and Decision-Making Explained*. Oxford University Press, Oxford.
- Rolls, E.T., 2014b. Limbic systems for emotion and for memory, but no single limbic system. *Cortex* 62, 119–157. <http://dx.doi.org/10.1016/j.cortex.2013.12.005>.
- Rolls, E.T., Deco, G., 2002. *Computational Neuroscience of Vision*. Oxford University Press, Oxford.
- Rolls, E.T., Deco, G., 2006. Attention in natural scenes: neurophysiological and computational bases. *Neural Netw.* 19, 1383–1394.
- Rolls, E.T., Kesner, R.P., 2006. A computational theory of hippocampal function, and empirical tests of the theory. *Prog. Neurobiol.* 79, 1–48.
- Rolls, E.T., O'Mara, S., 1993. Neurophysiological and theoretical analysis of how the primate hippocampus functions in memory. In: Ono, T., Squire, L.R., Raichle, M.E., Perrett, D.I., Fukuda, M. (Eds.), *Brain Mechanisms of Perception and Memory: From Neuron to Behavior*. Oxford University Press, New York, pp. 276–300.
- Rolls, E.T., O'Mara, S.M., 1995. View-responsive neurons in the primate hippocampal complex. *Hippocampus* 5, 409–424.
- Rolls, E.T., Stringer, S.M., 2000. On the design of neural networks in the brain by genetic evolution. *Prog. Neurobiol.* 61, 557–579.
- Rolls, E.T., Stringer, S.M., 2005. Spatial view cells in the hippocampus, and their idiothetic update based on place and head direction. *Neural Netw.* 18, 1229–1241.
- Rolls, E.T., Treves, A., 1990. The relative advantages of sparse versus distributed encoding for associative neuronal networks in the brain. *Netw. Comput. Neural Syst.* 1, 407–421.

- Rolls, E.T., Treves, A., 1998. *Neural Networks and Brain Function*. Oxford University Press, Oxford.
- Rolls, E.T., Webb, T.J., 2012. Cortical attractor network dynamics with diluted connectivity. *Brain Res.* 1434, 212–225.
- Rolls, E.T., Xiang, J.-Z., 2005. Reward-spatial view representations and learning in the hippocampus. *J. Neurosci.* 25, 6167–6174.
- Rolls, E.T., Xiang, J.-Z., 2006. Spatial view cells in the primate hippocampus, and memory recall. *Rev. Neurosci.* 17, 175–200.
- Rolls, E.T., Miyashita, Y., Cahusac, P.M.B., Kesner, R.P., Niki, H., Feigenbaum, J., Bach, L., 1989. Hippocampal neurons in the monkey with activity related to the place in which a stimulus is shown. *J. Neurosci.* 9, 1835–1845.
- Rolls, E.T., Robertson, R.G., Georges-François, P., 1997. Spatial view cells in the primate hippocampus. *Eur. J. Neurosci.* 9, 1789–1794.
- Rolls, E.T., Treves, A., Robertson, R.G., Georges-François, P., Panzeri, S., 1998. Information about spatial view in an ensemble of primate hippocampal cells. *J. Neurophysiol.* 79, 1797–1813.
- Rolls, E.T., Stringer, S.M., Trappenberg, T.P., 2002. A unified model of spatial and episodic memory. *Proc. R. Soc. Lond. B Biol. Sci.* 269, 1087–1093.
- Rolls, E.T., Xiang, J.-Z., Franco, L., 2005. Object, space and object-space representations in the primate hippocampus. *J. Neurophysiol.* 94, 833–844.
- Rolls, E.T., Stringer, S.M., Elliot, T., 2006. Entorhinal cortex grid cells can map to hippocampal place cells by competitive learning. *Netw. Comput. Neural Syst.* 17, 447–465.
- Rondi-Reig, L., Libbey, M., Eichenbaum, H., Tonegawa, S., 2001. CA1-specific N-methyl-D-aspartate receptor knockout mice are deficient in solving a nonspatial transverse patterning task. *Proc. Natl. Acad. Sci. U. S. A.* 98, 3543–3548.
- Schultz, S., Rolls, E.T., 1999. Analysis of information transmission in the Schaffer collaterals. *Hippocampus* 9, 582–598.
- Sidhu, M.K., Stretton, J., Winston, G.P., Bonelli, S., Centeno, M., Vollmar, C., Symms, M., Thompson, P.J., Koeppe, M.J., Duncan, J.S., 2013. A functional magnetic resonance imaging study mapping the episodic memory encoding network in temporal lobe epilepsy. *Brain* 136, 1868–1888.
- Treves, A., 1990. Graded-response neurons and information encodings in autoassociative memories. *Phys. Rev. A* 42, 2418–2430.
- Treves, A., 1991. Dilution and sparse coding in threshold-linear nets. *J. Phys. A* 24, 327–335.
- Treves, A., 1995. Quantitative estimate of the information relayed by Schaffer collaterals. *J. Comput. Neurosci.* 2, 259–272.
- Treves, A., Rolls, E.T., 1991. What determines the capacity of autoassociative memories in the brain? *Netw. Comput. Neural Syst.* 2, 371–397.
- Treves, A., Rolls, E.T., 1992. Computational constraints suggest the need for two distinct input systems to the hippocampal CA3 network. *Hippocampus* 2, 189–199.
- Treves, A., Rolls, E.T., 1994. A computational analysis of the role of the hippocampus in memory. *Hippocampus* 4, 374–391.
- Webb, T., Rolls, E.T., Deco, G., Feng, J., 2011. Noise in attractor networks in the brain produced by graded firing rate representations. *PLoS One* 6, e23630.
- Witter, M.P., Naber, P.A., van Haefen, T., Machielsen, W.C., Rombouts, S.A., Barkhof, F., Scheltens, P., Lopes da Silva, F.H., 2000. Cortico-hippocampal communication by way of parallel parahippocampal-subicular pathways. *Hippocampus* 10, 398–410.