COGNITIVE NEUROPSYCHOLOGY, 2005, 22 (5), 559–587

Psychology Press
Taylor & Francis Group

TRACELINK: A MODEL OF CONSOLIDATION AND AMNESIA

Martijn Meeter
Vrije Universiteit Amsterdam and University of Amsterdam, The Netherlands

Jaap M. J. Murre
University of Amsterdam and University of Maastricht, The Netherlands

A connectionist model is presented, the TraceLink model, that implements an autonomous "off-line" consolidation process. The model consists of three subsystems: (1) a trace system (neocortex), (2) a link system (hippocampus and adjacent regions), and (3) a modulatory system (basal forebrain and other areas). The model is able to account for many of the characteristics of anterograde and retrograde amnesia, including Ribot gradients, transient global amnesia, patterns of shrinkage of retrograde amnesia, and correlations between anterograde and retrograde amnesia or the absence thereof (e.g., in isolated retrograde amnesia). In addition, it produces normal forgetting curves and can exhibit permastore. It also offers an explanation for the advantages of learning under high arousal for long-term retention.

INTRODUCTION

Where are memories located? Although new imaging techniques allow visualisation of brain activity, memory loss after brain damage is still a vital source of evidence as to where memory resides in the brain. There are two kinds of amnesia, with the division between the two being along the time axis. The onset of the amnesia is taken as the point of reference. Any memory loss before that moment is defined as retrograde amnesia, whereas any memories not acquired or not retained thereafter fall under anterograde amnesia (at least, if they would have been acquired and retained under normal circumstances). Retrograde amnesia is thus the loss of memories that were accessible before the lesion, while anterograde amnesia reflects an inability to form or retain new memories.

A major characteristic of memory loss after brain damage is Ribot's Law of retrograde amnesia (Ribot, 1881). It states that there is a time-gradient in retrograde amnesia, such that recent memories are more likely to be lost. In the past century, this time gradient, usually referred to as the Ribot gradient, has been reported time and again. Not all studies of retrograde amnesia report the Ribot gradient, and the extent and length of the gradient can vary considerably from case to case. However, the sheer number of studies in which a Ribot has been found suggest that the Ribot gradient is not a pattern restricted to one category of patients or to one

Correspondence should be addressed to M. Meeter, Sectie Cognitieve Psychology, Vrije Universiteit, Vd Boechorststraat 1, 1081 BT Amsterdam, The Netherlands (Email: m@meeter.nl).

This research was supported by a PIONIER grant from the Netherlands Society for Scientific Research (NWO) to the second author.

http://www.tandf.co.uk/journals/pp/02643294.html
DOI:10.1080/02643290442000194

type of memory test (A. S. Brown, 2002; Meeter & Murre, in press; Squire, 1992).

The Ribot gradient is all the more striking in the light of normal retention. Since Ebbinghaus, it has been known that memories typically show a monotonically decreasing retention curve, which holds under a large number of conditions (Baddeley, 1990; Bahrick, Bahrick, & Wittlinger, 1975; Ebbinghaus, 1885; Rubin & Wenzel, 1996; Slamecka & McElree, 1983). A theory of amnesia thus has to explain—or at least allow for—the fact that we find opposite temporal gradients in normal and disturbed memory. In normal subjects recent memories are retrieved better than old ones (the forgetting curve), whereas with severe brain damage remote memories tend to be preserved better then recent ones (Ribot's Law).

Many authors (e.g., Alvarez & Squire, 1994; McClelland, McNaughton, & O'Reilly, 1995; Milner, 1989; Squire, 1992; Squire, Cohen, & Nadel, 1984) have explained the paradox by assuming two processes that influence trace strength. New experiences are first stored in the hippocampus, a relatively fast process that leads to near-immediate prominence of novel episodic memories. Thereafter, a second, relatively slow process makes some memories more persistent with time: During "off-line" periods such as sleep, these memories become reactivated and are transferred to the neocortex. This process of memory transfer has been referred to as "memory consolidation." The above explanation has fallen on such fertile grounds that Nadel and Moscovitch (1997) refer to it as the "Standard View" of amnesia.

The standard view has been implemented in several connectionist models, one of which is the TraceLink model (Murre, 1994, 1996, 1997). Although retrograde amnesia and its gradient were successfully simulated in several models, none has been applied to amnesia in its diverse forms, and few of the implications for remote memory have been thoroughly studied. Here, we will do just that. The TraceLink model will be applied to a number of amnesic conditions, to forgetting, and to variance in memory strength. This will allow a thorough evaluation of consolidation and its ramifications. Before we discuss amnesia, however, we will present the architecture of TraceLink, starting with two neurobiological concepts that inform our model.

Hierarchies in the brain

Episodes that are stored in memory consist of a relatively random collection of facts, sensations, verbal descriptions, thoughts, and other experiences, and have visual, auditory, tactile aspects, and so on. Such episodes are, according to virtually all theories of memory, stored as a collection of features (corresponding to aspects of the episode) that are associated during the learning process. Given what is known about the localisation of functions in the brain, it seems plausible that these neural activation patterns will be distributed over many brain areas that are at a sizeable distance from one another, e.g., areas that code for different modalities (Cabeza & Nyberg, 2000). To store an episode, one would thus have to connect aspects of the episode coded for at different places in the brain, which requires long-distance cortico-cortical connections. Those connections are known to exist only sparsely, however (Abeles, 1991; Braitenberg & Schüz, 1991). An individual cortical neuron is typically connected to 5000 others and most of these will be located in the direct neighbourhood (Murre & Sturdy, 1995).

A hierarchical connectivity structure with dense bidirectional connections may solve this problem, analogous to the way that telephone networks are designed to deal with the same issue. Local telephone connections (i.e., cables) are dense. They hook up to local area networks, which in turn are connected to national and international networks. There is ample evidence that the cortex is structured in such a hierarchical fashion, and that the hippocampus is a top-level structure. In the rat, the hippocampus receives inputs from nearly the entire neocortex via the entorhinal, perirhinal, and postrhinal cortices (Burwell, 2000; Witter, Wouterlood, Naber, & Van Haeften, 2000). These connections are largely reciprocal. Further evidence for this hierarchical connection scheme derives from a study by Felleman and Van Essen (1991) in the macaque monkey. They investigated

connectivity patterns between 32 visual areas and estimated that only 40% of the possible area-to-area connections exist. Felleman and Van Essen (1991) arrived at a visual hierarchy including 14 levels of cortical processing with the entorhinal cortex and hippocampus at the top. For the mouse, a similar hierarchical connectivity was found (Braitenberg & Schüz, 1991; Greilich, 1984). TraceLink assumes that there is such a hierarchy. It assumes that there are areas in the brain with high connectivity that have as a principal function to link other areas of the brain.

Modulation

A further consideration in the development of the TraceLink model is the fact that learning is not equally fast at all times. Indeed, it should not be, as some things are worth learning and others are not. The problem of real-life or real-time learning is well formulated by Carpenter and Grossberg (1988) in what they call the "stability-plasticity dilemma": "How can a learning system be designed to remain plastic, or adaptive, in response to significant events and yet remain stable in response to irrelevant events?" (p.77). They have argued that any system that wants to perform significant real-time learning must solve this problem. If a system does not, and instead learns all events with equal intensity, its memory will soon be cluttered with a large collection of irrelevant facts.

The stability-plasticity dilemma has been addressed by relatively few researchers (Carpenter & Grossberg, 1988; Grossberg, 1976, 1987; Hasselmo, 1994; Hasselmo, Bodelón, & Wyble, 2002; Meeter, Talamini, & Murre, in press; Murre, 1992). Their solutions are based in part on monitoring nonspecific aspects of stimuli, such as resonance, competition, and novelty, using these measures to control learning processes (usually, such models tackle categorisation, which is not the aim of TraceLink). Other more psychological factors also seem to influence how strongly a memory is encoded in memory. A number of concepts have become associated with this question, among which are attention, arousal, vigilance, wakefulness, emotional relevance, novelty, motivation, and drive. These terms, though overlapping, serve as a broad outline of general modulatory factors and plasticity-inducing circumstances. In the brain, several central mechanisms have been proposed that modulate learning, for example an amygdaloid system based on norepinephrine (McGaugh, 1990) and a cholinergic system arising from the basal forebrain (Hasselmo, 1995).

Although the details of the modulation of learning are outside the scope of TraceLink, we propose that there are systems with such functions. In the model, regulating plasticity is the function of a separate module called the modulatory system.

The TraceLink model

The hierarchical structure of the brain and the need for modulation of learning inform a model containing three main components: (1) a neocortical "trace" system, (2) a hippocampal "link" system, and (3) a modulatory system.

1. Normally, the greater part of a memory trace will be stored in the connections of the trace system. The trace system represents roughly the neocortical basis of memories. The input to the trace system originates in sensory areas. We assume considerable preprocessing in these areas, which themselves do not form part of the model.

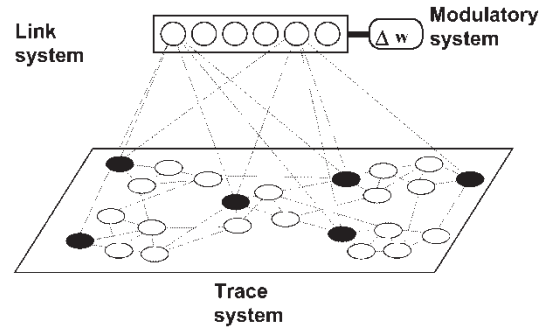

Figure 1. *Overview of the TraceLink model, showing the link system, the trace system, and the modulatory system (indicated by a ΔW sign, symbolising control of learning rate on the connection weights in the system). Only a few nodes and a few connections have been drawn in order to prevent clutter.*

Similarly, output or motor areas are not included. We identify the trace system with association areas in the neocortex, such as the temporal lobe neocortex (Miyashita, 1993) and posterior parietal cortex (Izquierdo et al., 1997).

2. It is the link system's function to connect remote trace elements (i.e., those without direct cortico-cortical connections). The link system has a much smaller number of elements than the trace system. Among other things, this implies that link elements are more likely to be reassigned from an old representation to a new, thus causing interference by new learning. Link elements are connected to each other and to a random subset of the trace elements. Connections involving link elements are much more plastic than trace-to-trace connections. This is illustrated in Figure 1 through the close attachment of the link system to the modulatory system. The link system can be identified with what others have called the hippocampal complex (Nadel & Moscovitch, 1997) or the medial temporal lobe (Alvarez & Squire, 1994): the hippocampal region (hippocampus proper, dentate gyrus, subiculum) and adjacent parahippocampal structures (principally the entorhinal, parahippocampal, and perirhinal cortices).

3. Activation of the modulatory system causes increased plasticity of the link system. The modulatory system may be activated through central states such as arousal and attention, and it can be identified with the basal forebrain, the amygdala, and other subcortical centres that control plasticity in the brain. A structure that certainly also plays a role in the modulatory system is the hippocampus, which many studies have implicated in novelty processing (Johnson & Moberg, 1980; Knight & Nakada, 1998; Montag-Sallaz, Welzl, Kuhl, Montag, & Schachner, 1999; Mumby, Gaskin, Glenn, Schramek, & Lehmann, 2002). The hippocampus thus plays a double role in the TraceLink model: It is part of the link system, but it is also involved in regulating its own plasticity.

Although we do not here explore modulation in great detail, several models have suggested ways in which modulatory systems may work in the brain (Doya, Dayan, & Hasselmo, 2002; Hasselmo, Schnell, & Barkai, 1995; Hasselmo & Wyble, 1997; Meeter et al., in press). For example, a few models have explored how the hippocampus may regulate its own plasticity via a feedback loop with the medial septum. In these models, a novelty signal derived from hippocampal processing triggers the release of acetylcholine in the hippocampus, upregulating plasticity there (Hasselmo et al., 1995; Hasselmo & Wyble, 1997; Meeter et al., in press). Such mechanisms are not implemented in this version of the model, but can be mimicked by simply designating certain stimuli as "interesting" and manually increasing the activation of the modulatory system whenever they are present.

How the model works

Figure 2 illustrates at a conceptual level how, in the model, episodic memory traces are formed under normal circumstances. For the sake of exposition, four stages can be distinguished (these do not have theoretical value).

Stage 1. Through the sensory and motor channels a set of trace nodes is activated (filled circles). This represents an episode to be remembered. The activated trace nodes have no direct trace-to-trace connections to each other but they are connected to a number of link nodes (only two are drawn). We assume that the mechanisms for local inhibition, outlined previously, keep the number of active nodes small.

Stage 2. The trace nodes activate a set of link nodes. In the current model, these are random sets of nodes; in a more detailed model, patterns would consist of the most activated link nodes, with nodes with less activation being suppressed by inhibitory processes. This comes down to pattern formation through self-organisation, as has been implemented in numerous models of the hippocampal region (Doya, 2000; Hasselmo et al., 1995; Jensen, Idiart, & Lisman, 1996; Meeter et al., in press; Norman & O'Reilly, 2003). The modulatory system becomes activated (darkness of shading indicates activation level) and the learning rate increases. As a result of the increased plasticity, connections between link and trace nodes are strengthened (shown by a

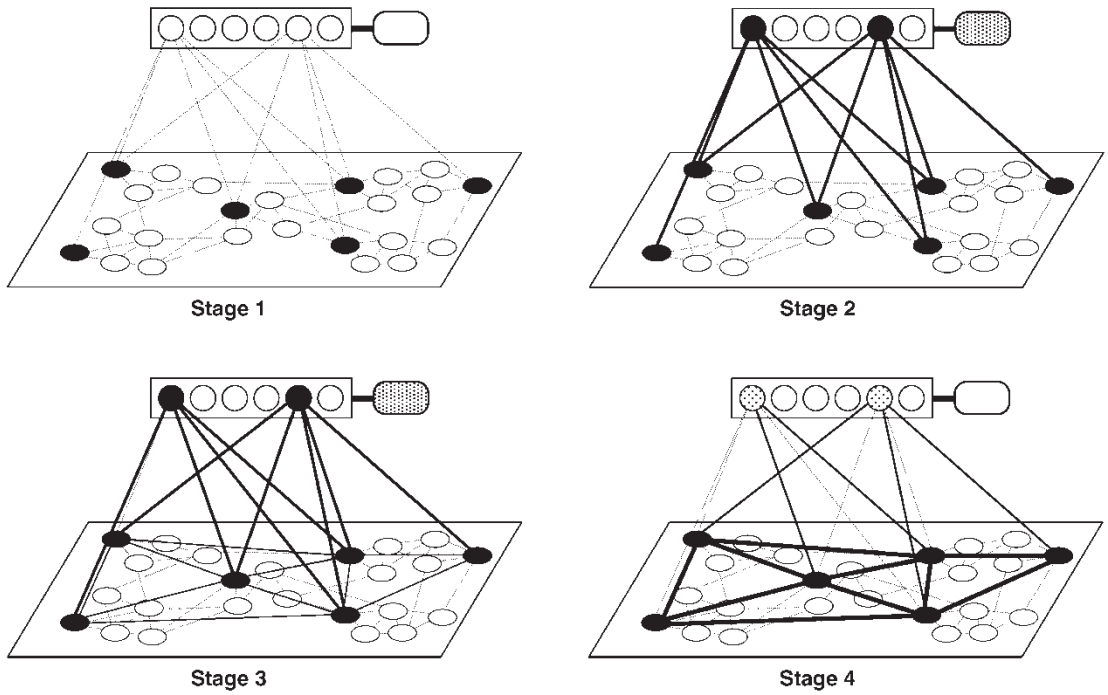

Figure 2. *Four stages in the normal formation of episodic memories in the TraceLink model.* Stage 1: *A new memory representation activates a number of trace elements (shown as filled black circles).* Stage 2: *Several link elements are activated and the relevant trace–link connections are strengthened (shown as thicker connections). Also, the modulatory system has been activated.* Stage 3: *Weak trace–trace connections are developing. The modulatory system is weakly activated.* Stage 4: *Strong trace–trace connections have been formed. Trace–link connections have decayed and the modulatory system does not necessarily respond to the stimulus.*

thickening of the connections). This can take place in minutes or seconds.

Stage 3. Prolonged or repeated activation of the memory trace through the link system will lead to the gradual formation of trace-to-trace connections. With subsequent reactivations, the modulatory system responds with less and less activation—it gradually habituates to the pattern.

Stage 4. Direct trace-to-trace connections have become very strong. Link–trace connections have either decayed or have been reassigned to other memory traces. The modulatory system shows little reaction to the pattern. In short, the memory trace has become independent of the link system.

Recall is modelled as the retrieval of the whole trace pattern when part of the trace pattern is offered as a cue. In the case of a stage 2 memory, the partial cue will typically activate the link nodes associated with the pattern, which in turn will activate the rest of the trace pattern. In the case of a stage 4 memory, the cue will be able to activate the rest of the trace pattern directly through strong trace-to-trace connections.

Consolidation of memories is the transformation of stage 2 memories into stage 4 memories through repeated reactivation via the link system (this proposal is similar to that by Alvarez & Squire, 1994; McClelland et al., 1995; and others). The speed of this transformation may vary with type of material and with many other factors. For

some memories, consolidation may take a very long time (up to several decades). The formation of long-range trace connections is not by direct synaptic contact. The possibility of long-range axonal sprouting should not be excluded, but we think it more likely that these connections are established via chains of neurons, as outlined, for example, by Abeles (1991). Establishing a reliable connection will take repeated exposures to the desired connection pattern.

What TraceLink shares with previous work

The assumptions incorporated in TraceLink are not highly original, nor do they have to be: TraceLink aims to encapsulate and implement in some detail the main lines of thought about memory, amnesia, and the brain that can be traced in the neuropsychological literature of the past 50 years. This earlier theoretical and modelling work has inspired the TraceLink model (e.g., Eichenbaum, Cohen, Otto, & Wible, 1992; Marr, 1971; Milner, 1989; Squire, 1992; Squire et al., 1984; Wickelgren, 1974, 1979). Since the first work on the model (Murre, 1994), several other models of amnesia have been published that focus on the role of the hippocampus (e.g., Alvarez & Squire, 1994; McClelland et al., 1995). TraceLink is perhaps most similar to the model by Alvarez and Squire (1994). The latter model also uses a hierarchical structure, with a link system connecting two cortical modules, and whereby a gradual consolidation process causes a movement from "cortico-hippocampal" dependence to a purely "cortico-cortical" basis. This type of consolidation process is similar to one we have adopted.

On a conceptual level, TraceLink also shares many elements with the McClelland et al. (1995) model. However, their implementation is very different, with a backprop network as their "trace" system, and their "link" system is implemented as a probability distribution with which consolidation trials for old patterns are interleaved with acquisition trials for new patterns. This implementation follows quite naturally from their central rationale for two memory systems. They point out that purely sequential learning may not lead to useful internal representations, and that a case can be made for the necessity of a more interleaved mode of learning. In particular, newly learned deviant patterns may disturb already learned representations. Occasional learning trials for old patterns counteract this disturbance. There are thus good reasons for a slow (interleaved) learning process such as consolidation of long-term memory.

Neuropsychological data to be explained

In the next section we shall illustrate with connectionist simulations how the TraceLink model may account for normal episodic learning and recall, and for findings in the literature on amnesia. Before describing these simulations, we will list what, in our view, are the main conclusions in the literature on amnesia that have to be explained by a comprehensive model of amnesia:

1. Although the exact brain areas involved remain unclear, damage to certain medial structures—in particular, the hippocampus, adjacent medial temporal lobe structures, and the medial diencephalons—can cause both retrograde and anterograde amnesia. The retrograde amnesia shown by patients suffering such lesions typically shows a temporal gradient in accordance with Ribot's Law (Reed & Squire, 1998; Rempel-Clower, Zola, Squire, & Amaral, 1996; Squire, Haist, & Shimamura, 1989).

2. Anterograde and retrograde amnesia are partially correlated, with the correlation varying from one patient group to the next (Kopelman, Wilson, & Baddeley, 1989; Russel & Nathan, 1946; Shimamura & Squire, 1986). Indices of anterograde amnesia correlate more with retrograde amnesia for the periods right before the lesion than with retrograde amnesia for more remote periods (Schmidtke & Vollmer, 1997).

3. In patients, substantial anterograde amnesia is nearly always accompanied by at least some retrograde amnesia. It occurs in complete isolation of retrograde amnesia, however, after injections of scopolamine, a cholinergic blocker (Kopelman & Corn, 1988). This suggests that multiple causes may underlie anterograde amnesia.

4. Retrograde amnesia for more remote periods, including isolated retrograde amnesia, appear to be

caused by neocortical lesions. "Isolated" retrograde amnesia is not accompanied by anterograde amnesia (or very little) but always involves at least some initial anterograde amnesia in the period following the onset of the lesion (Kapur, 1993).

5. Transient forms of amnesia show that retrograde amnesia can involve a retrieval deficit; after the amnesia has resolved, the patient is typically able to retrieve most memories that were unavailable during the amnesia episode (Cahill & Frith, 1995; Kritchevsky, 1992). The patient is usually left without memories of the episode itself, which suggests that transient forms of amnesia also involve impaired learning.

6. During recovery from retrograde amnesia, memories from more remote times tend to come back faster than memories from more recent times; this is referred to as shrinkage (Russel & Nathan, 1946; Whitty & Zangwill, 1977).

7. Memory impairments in amnesia are probably limited to explicit memory: Amnesic patients show normal or near-normal implicit memory. Implicit memory seems to involve incremental learning in neocortical processing areas (Gabrieli, 1998; Schacter, 1992).

SIMULATIONS

Details of the connectionist model

The model used to simulate the findings listed above consists of two components: the trace system and the link system. Since none of the simulations depends on more than a relatively coarse notion of modulation, we made no attempt to model the third component of the TraceLink model, the modulatory system. Instead, its functions are assumed here (see Meeter et al., in press, for a recent implementation of a modulatory system by our group). The trace system is modelled as a layer of 200 nodes, the link system as a layer of 42 nodes. Both layers have internal connections, and are connected with one another. Every two nodes can in principle be connected. As is the case with a majority of neurons that are involved in learning in the cortex, nodes only have excitatory synapses. The nodes model groups of neurons that are at some distance from each other.

Both layers have binary stochastic nodes that are "on" or "off " with a certain likelihood. This likelihood depends on the balance between excitatory input and inhibition. The excitatory input to a node is the weighted sum of the activation of all nodes connected to it. From this excitatory input, inhibition is then subtracted. Inhibition is constantly fine-tuned so as to keep the average number of active cells in a layer as close to a preset number (k) as possible. The inhibition mechanism models the working of inhibitory neurons, which may have an important function in keeping activity in their region within bounds (Braitenberg & Schüz, 1991; Minai & Levy, 1994). Inhibition is increased when too many nodes are active (i.e., more than k), and lowered when too few are (i.e., less than k). The number k is set separately for every layer; consequently inhibition is regulated separately in every layer.

The weights of excitatory connections are the locus of learning. Weights can vary between 0 and 1, and are changed in a simple, linear fashion according to a variant of Hebb's rule (Hebb, 1949). This learning rule allows for learning as well as unlearning as a function of contingent and non-contingent activity of the nodes. The activation, inhibition, and learning rules are explained in detail in Appendix A.

Learning is not equally fast for all connections. The learning rate—the rate at which changes are made to the weights—is much lower for the within-trace connections than for the connections within the link layer, or between the link layer and the trace layer. Connections between the trace and link layers and within the link layer have equal learning rates. Since nodes model groups of neurons, the connection between two nodes is a function of both the number of synapses between the groups of neurons and their average strength. Although the hippocampus is known to be unusually plastic (Lopes da Silva, Witter, Boeijinga, & Lohman, 1990), the higher learning rate in the connections involving the link layer is primarily intended to model the higher connectivity in the regions of the link system (Treves & Rolls, 1994).

A pattern in the learning set consists of a group of trace nodes and a group of link nodes, which are activated when a pattern is presented. Patterns can overlap in both the trace and the link layer. The number of nodes in the trace layer and the link layer that belong to a pattern is equal to the number of nodes active in the layer in equilibrium (k). Because patterns are chosen independently of each other, every two patterns share on average a k/m proportion of their nodes in a layer, where m is the number of nodes in the layer. In both layers k is relatively low compared to m, which has two distinct consequences: first, that at any time in the simulations only a few nodes of both layers are active, and second, that the patterns are sparsely coded with little overlap.

Because the hippocampal system is much smaller than the neocortex, the overlap of patterns is greater in the link system than in the trace system. This is merely intended to model the smaller capacity of the hippocampal region as compared to the neocortex. In particular, it leads to faster forgetting in the link system than in the trace system. However, this choice of parameters would seem to put TraceLink into direct conflict with theories predicated on sparse representations in the hippocampus (McClelland et al., 1995; O'Reilly & McClelland, 1994; O'Reilly & Rudy, 2000). Some fields of the hippocampus, notably the dentate gyrus and to a lesser extent CA1, are indeed known to exhibit sparsely firing neuron populations (Amaral, Ishizuka, & Claiborne, 1990). It is less obvious that the remaining fields of the hippocampus and other regions of the medial temporal lobe have sparser firing than the neocortex. We chose not to alter our overlap parameters, however, as this would have engendered the inclusion of specific forgetting parameters. In this context, it is noteworthy that neocortical LTP decays more slowly than does hippocampal LTP (Trepel & Racine, 1998), and that within the hippocampal region the dentate gyrus combines extremely sparse firing with relatively rapidly decaying LTP (e.g., Ezrokhi, Zosimovskii, Korshunov, & Markevich, 1999).

The values of what one could call "equilibrium parameters" are given in Appendix A. The function of these parameters is to keep the number of active nodes as close as possible to the equilibrium value k, and to prevent wild swings in the number of active nodes from one iteration to the next. These parameters only influence the results in the sense that they make meaningful results possible; they do not have an effect on the pattern of findings. Excluding the "equilibrium parameters," we are left with a few parameters that can perhaps count as free: the learning rates, and the proportion k of active nodes in equilibrium in both layers relative to the size of the layers (see Table 1). The main findings in the simulations were quite robust, with the form of most functions staying the same with many different sets of parameters. One set of parameter values, that given in Table 1, was used in all simulations.

Simulation 1: Normal learning and recall

Method

We first simulated the normal workings of the model: normal learning, consolidation, and recall. This simulation also served as a control for our simulations of different amnesic states. In it, the model went through two distinct phases, a learning phase and a test phase. The learning phase consisted of two alternating subphases, acquisition of a pattern and consolidation (see Figure 3). During acquisition, the model learned one new pattern. This was followed by a period of consolidation, after which another pattern was acquired. In a simulation, the model learned a set of 15 or more patterns with interspersed consolidation periods, after which all the patterns were tested. We replicated each simulation 200 times.

Patterns consisted of a random 10 trace nodes and 7 link nodes ($1/20$th of all trace nodes, and $1/6$th of all link nodes). As patterns were random, on average every two patterns shared $1/6$th of their link nodes and $1/20$th of their trace nodes. Each pattern was learned for one iteration with the learning parameter listed in Table 1.

After the acquisition of a pattern, the model entered a period of consolidation. In a consolidation

Table 1. *Parameters in the presented simulations with the TraceLink model*

Learning rate during acquisition	
Within trace	0.06
Within link and between the layers	0.4
Learning rate during consolidation	
Within trace	0.0025
Within link and between the layers	0.0
Total learning during one acquisition	
Within trace	0.06 (0.06 * 1 iteration)
Within link and between the layers	0.4 (0.4 * 1 iteration)
Total learning during one consolidation period	
Within trace	0.06 (0.0025 * 3 trials of 8 iterations)
Within link and between the layers	0.0
Unlearning rate	
In all connections	75% of learning rate
Number of nodes	
In trace	200
In link	42
Number of nodes active in equilibrium k[a]	
In trace	10 (1/20th of the nodes)
In link	7 (1/6th of the nodes)

[a] = number of nodes in one pattern.

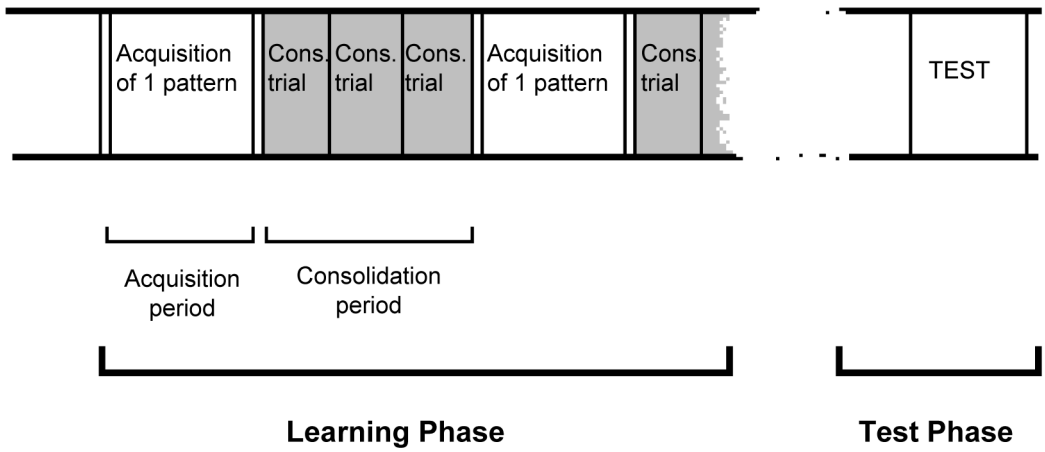

Figure 3. *Diagram showing the order of events in most simulations. A simulation was divided into a learning phase and a test phase. The learning phase was subdivided into alternating acquisition periods, in which one pattern was acquired, and consolidation periods. Consolidation periods consisted of three consolidation trials.*

period, three consolidation trials occurred. A single trial proceeded as follows: The model was set to a random pattern, and then allowed to cycle freely for a fixed number of iterations (150). Whichever pattern was active at the last iteration was consolidated. The dynamics of the model thus selected a pattern to consolidate, with attractors surfacing that were strong in the combined link and trace system. Consolidation was done for eight iterations at a low learning rate (see Table 1;

as the activation would wander a little during the eight iterations, learning for one iteration with a high learning rate would not have been equivalent to learning for eight iterations with a low rate). Because consolidation trials started with random patterns and the activation rule was stochastic, the model usually did not settle on the same pattern for all three consolidation trials. More than one pattern was thus typically consolidated in a consolidation period. Moreover, as the number of active nodes was not always equal to the equilibrium number k, the model could consolidate noisy patterns, mixtures of patterns, or no pattern at all.

Acquisition of the first two patterns was followed by fewer than three consolidation trials so as not give these patterns too much of a head start: As there were no previous patterns for the model to choose, the first patterns were always consolidated in these early consolidation trials. Since the very first pattern was learned in an "empty brain," it is atypical and has been excluded from all analyses and figures showing results.

After learning and consolidation, the model was subjected to a test phase. Each pattern was tested a number of times by activating and clamping part of the pattern in the trace layer (the cue), and letting the model cycle for 70 iterations. No nodes were clamped in the link layer. After the model had gone through the 70 iterations, the active trace nodes of the pattern that were not part of the cue were counted. That number, divided by the number of trace nodes in the pattern not part of the cue, was used as the measure of performance. Less crude measures were tried, but did not yield qualitatively different results. After each test, the model was reset, the cue was activated and clamped anew, and another test was carried out. The scores that will be reported are thus the average proportions of trace nodes in a pattern that were not part of the cue, but were active after 70 iterations in the test.

Results and discussion

In most of the consolidation trials, one relatively intact pattern was found and consolidated. In the simulation of normal learning, this was the case in 86% of the consolidation trials. In 9% of the trials no pattern was found, and in 5% more than one pattern was active at consolidation (usually two). Often the pattern consolidated was the one learned just before the consolidation trial, but it could also be an older pattern. Figure 4 shows the likelihood that a given pattern was consolidated at different times in the experiment. For all patterns, this likelihood was relatively high in the first consolidation periods after the pattern was learned, and dropped off when more and more new patterns were learned.

As is clear from Figure 4, the likelihood that a given pattern was consolidated monotonically decreases with every new pattern that was learned. We fitted several functions on how much the first 10 patterns were consolidated at each time step after learning. A power law decrease fitted best on the consolidation curve of some patterns, logarithmic decrease on others (fits varied from an R^2 of .91 to .97 for power functions and from .92 to .99 for logarithmic functions).

Figure 5 shows the results of the first simulation (the filled circles). Performance was very high for the most recent pattern; the older a pattern was, the lower it scored on the test. The curve was best approximated by a power function, which explained 93% of the variance. This is in accordance with data from human subjects: The power function is often seen as the best approximation of the retention function for human memory (Anderson & Schooler, 1991; Rubin & Wenzel, 1996; Wixted & Ebbesen, 1991).

Figure 5 also shows a chance level in the simulations. Since TraceLink strives for a given number of active nodes in a layer (parameter k), during testing there will always be some active nodes that may or may not belong to the pattern. We determined the chance level by defining an extra pattern that was not learned, testing this pattern in the usual way. The figure shows that although forgetting was substantial, pattern recall remained well above chance.

These results show the forgetting that occurs in the model when new patterns are learned. The mechanism that accounts for forgetting in TraceLink is interference through overlap of patterns. When a pattern is learned, it is immediately stored in the link system in a way that enables

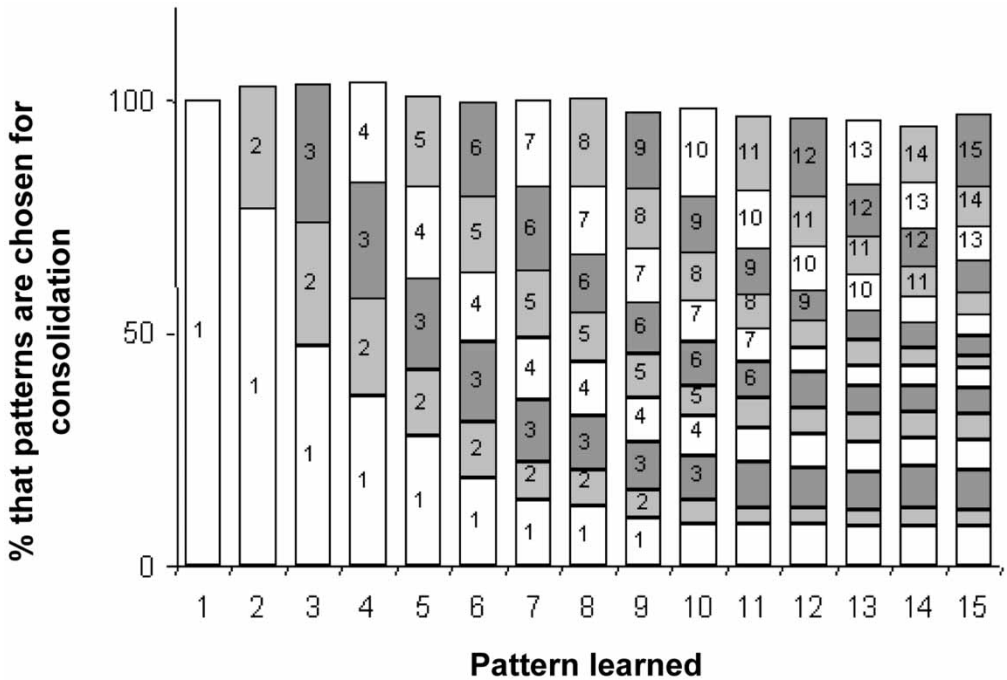

Figure 4. *The likelihood that, in 400 replications of simulation 1, a pattern is consolidated in a particular consolidation period. The abscissa shows the pattern presented for acquisition prior to the given consolidation period. The ordinate shows the likelihood of every pattern of being consolidated in the trials in that consolidation period. These likelihoods do not always sum to exactly 100% because in some trials more than one pattern, or none at all, are consolidated (see text).*

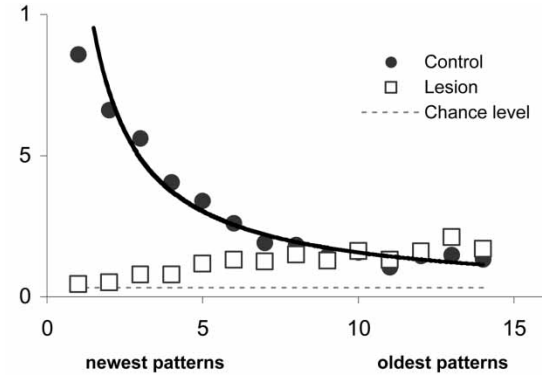

Figure 5. *Results of the basic simulation of normal learning and forgetting (filled circles), and the simulation of retrograde amnesia (open squares). Fifteen patterns were learned; the patterns on the left were the most recently learned, the patterns on the right are the oldest. The first-learned pattern is not shown. Scores are the mean proportion of nodes in the trace portion of the pattern that are active at test and not part of the cue. The continuous line is a power fit of the normal data series ($R^2 = .93$).*

retrieval later on. The cue that is presented in the trace system during the test can activate the pattern in the link system, which subsequently activates the rest of the pattern in the trace system. When a new pattern overlaps with an old one in node X, the old pattern is partly unlearned: While new connections are laid between node X and the other nodes in the new pattern, the connections from the old pattern to node X are unlearned. The node is thus effectively disconnected from the old pattern. The whole pattern is gradually unlearned when more and more nodes are disconnected from the pattern due to overlap with newer ones. Since there is more overlap in the link system than in the trace system, the link portion of a pattern is lost relatively rapidly, and the trace portion more slowly. By the time 15 patterns have been learned, many nodes in the first pattern have been part of at least one other pattern and are thus lost for the

first pattern; on average this is the case for 54% of its trace nodes and 94% of its link nodes. Since old patterns decay rapidly in the link system, patterns in the link system quickly lose their ability to be activated by a cue in the trace system, or to maintain a stable activation of the pattern in the trace system. During the time that the pattern is still strong in the link system, however, the pattern may be consolidated, and its strength continues to build up in the trace system. This enables retrieval of the older patterns from the trace layer. An old pattern may be activated on the basis of a strong trace representation alone.

Overwriting by subsequent memories seems a crude explanation for forgetting as compared to decaying connections, or forgetting through contextual cue changes. However, a recent review suggested that overwriting indeed plays an important role in the reversal of long-term potentiation in the hippocampus (Rosenzweig, Barnes, & McNaughton, 2002).

As already stated, basic results were found with many parameter sets. Some parameter sets produced undesirable side effects, however, such as sizeable primacy effects in long-term memory. This occurred because consolidation is implemented in TraceLink as a competitive process. Strong patterns have a greater likelihood of becoming active when the model cycles freely during a consolidation trial, and are thus more likely to be consolidated. Since the last pattern is the strongest in the link layer, it was usually consolidated most often (see Figure 4). However, the strength of the pattern in the trace layer also played a role in determining which pattern was consolidated. If consolidation went too fast, or if the trace layer was too strong compared to the link layer, the trace layer tended to determine which pattern was consolidated. Often one of the early patterns was then consolidated over and over again. With every consolidation trial it would become stronger and thus more likely to be consolidated in the next trial. This resulted in a high performance for the first patterns compared to later patterns, something that can be termed "runaway consolidation" (Meeter, 2003a). We avoided this artifact by choosing the parameter set in such way that the strength of patterns in the two layers was balanced.

We also varied learning schedules (scheduling consolidation periods after more than one patterns had been acquired). When consolidation occurred after two or three patterns were acquired, little change in the results was seen. However, with more patterns acquired in between consolidation trials, runaway consolidation again occurred, as the pattern acquired right before consolidation had an advantage over other patterns in the competition for consolidation (elsewhere we have suggested remedies for runaway consolidation; Meeter, 2003a).

Simulation 2: long-term retention and permastore

Though monotonically decreasing retention curves have been found with a broad range of time scales, there is evidence that, after a number of years, forgetting ceases and recall probability remains stationary (Bahrick, 1984, 1992; Bahrick et al., 1975; Conway, Cohen, & Stanhope, 1991). Bahrick has called this state "permastore." This phenomenon has been found, among others, in retention of high-school Spanish (Bahrick, 1984), memory for the faces and names of classmates (Bahrick et al., 1975), and retained knowledge from a college course in cognitive psychology (Conway et al., 1991).

To test whether the model would develop permastore, we let the model learn more patterns. Simulation 2 proceeded exactly as simulation 1, with one difference: In this simulation, the model acquired not 15 patterns but 20.

Figure 6 shows the results from the permastore simulation. We fitted a power curve on the first 15 patterns (the number of patterns in the first simulation), and found that it explained 96% of the variance. If a power curve was fitted on all patterns, the fit decreased to 87%. On the last 10 patterns, however, a flat line fitted very well. The best fitting regression line (drawn in Figure 6) had a slope of just -0.0002. The forgetting curve shown by the model thus reaches an asymptote, at which forgetting stops and performance does not further deteriorate. A balance has been reached between memory decay and memory consolidation:

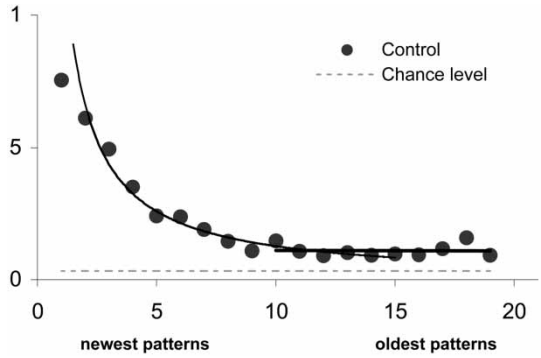

Figure 6. *Performance in the permastore simulation. A power curve was fitted on the most recent 15 patterns. A straight line was fitted on the oldest 10 patterns.*

The gains of consolidation are balanced by the forgetting caused by the acquisition of new patterns. This results in a state of permastore, where patterns remain at a constant level of retrieval performance. Perhaps permastore is thus not a state of immutable memory strength, but instead a dynamic strength in which forgetting is balanced by consolidation.

RETROGRADE AMNESIA

In the TraceLink model, retrograde amnesia can be modelled by a temporary or persistent loss of link nodes. This has different effects for recent as compared to older memories. Recent memory representations (stage 2 in Figure 2) are dependent on a functioning link system for their retrieval and internal coherence. When the link system is disabled, the trace nodes in the pattern lack the support from the link system and cannot activate the other trace nodes in the pattern. Remote memories (stage 4) have developed a supporting trace-to-trace connectivity structure, and their retrieval is independent of the link system. We postulate that this is the main mechanism underlying Ribot's Law. With intermediate memories, successful retrieval after a lesion of the link system will depend on which link nodes are unavailable and what trace-to-trace connections have been formed already (one could speculate that memories within strongly associated clusters support each other's the retrieval, which may explain the occasional isolated islands of memories preserved in an otherwise dense retrograde amnesia).

Simulation 3: Retrograde amnesia and the Ribot gradient

The learning phase in the simulation of retrograde amnesia proceeded as in simulation 1, with normal acquisition and consolidation of patterns. The only difference resided in the test: Before the test the link layer was deactivated to simulate a lesion in the medial temporal lobe. This meant that the trace layer was allowed to cycle, while all nodes in the link layer were inactive.

In the intact model, performance was highest for the most recent pattern. We obtained the opposite when the model was tested with a deactivated link layer. Though performance was relatively low for all patterns, the most recent patterns suffered more from the deactivation of the link layer than the oldest patterns (open squares in Figure 5). In that condition, the most recent patterns do not score much better than chance. This corresponds to the Ribot gradient found in patients with retrograde amnesia.

One difference between this simulation and a typical retrograde amnesia study deserves mentioning. Our items were all learned with the same strength, and our control condition shows a steep forgetting curve. Tests of retrograde amnesia, on the other hand, are typically constructed so that an equal number of items from various decades are answered correctly by normal controls (Mayes, Downes, McDonald, Rooke, Sagar, & Meudell, 1994; Meeter, 2003b; Sanders & Warrington, 1971). Since one can assume that more information from previous decades is forgotten, this equality of performance for normal controls implies that the items selected from the various decades differ in average initial learning strength. To compare data from the simulations with data from retrograde amnesia tests, one might plot the retrograde amnesia scores as a percentage of the normal control scores. Inspection of Figure 5 shows that this

would make the Ribot gradient much steeper. The retrograde amnesia score for the most recent patterns is only a small proportion of the normal control score, whereas the two scores are almost equal for older patterns.

SHRINKAGE OF RETROGRADE AMNESIA

Shrinkage refers to the process of recovery from retrograde amnesia. The term implies that in recovery from retrograde amnesia, older memories tend to become available before more recent memories (though, as mentioned earlier, isolated islands may become available before certain older memories; Whitty & Zangwill, 1977). TraceLink models amnesia that later resolves by a temporary unavailability of link nodes. This might be caused by, for example, a shift in the balance of inhibition and excitation through a lesion or another abnormality, resulting in a suppression of activity in link structures. Remote memories are unaffected by this unavailability, but retrieval of recent memories is impaired. If link nodes do become available again, several situations are possible. It may be that synapses on trace–link connections have deteriorated, meaning that recent memories havebeen lost. Trace–link connections may also become available again with little loss of weights. This last scenario is probable in the recovery from TGA, which can occur very rapidly (hours or less). Our simulation of shrinkage concentrates on this case; we simulated a TGA attack and its resolution.

Simulation 4: Transient global amnesia (TGA)

Evans, Wilson, Wraight, and Hodges (1993) observed medial temporal lobe hypoperfusion during a transient global amnesia (TGA) attack. Interpreting the hypoperfusion that Evans et al. found as a sign of low activity, we simulated TGA by temporarily suppressing activity in the link layer through lowering the value of k in the link layer. The parameter k is the number of nodes that are active at equilibrium. When it is lowered in a given layer, fewer nodes will generally be active in that layer.

After 14 patterns had been learned normally, k was set to zero in the link layer to simulate the TGA attack. This means that the model will try to suppress any activity in the link layer. One pattern was then learned under this simulated TGA, without any link activity. Then the first test occurred: All patterns were tested while the k parameter was still set to zero in the link layer. To simulate the gradual lifting of TGA, k was set to 3 and 5, and the model was tested a second and third time respectively. To simulate completely resolved TGA, k in the link layer was set to its normal level of 7, and the model learned five more patterns with the link layer wholly active. At the end of the simulation, the model was tested a fourth and final time, with k at its normal level.

Figure 7 shows the results for the tests during simulated TGA, during its resolution, and after the attack. In the first test, the simulated TGA resulted in a dense retrograde amnesia, and in a severe anterograde amnesia for the one pattern learned during the TGA attack (Figure 7a). There is also a Ribot gradient in the retrograde amnesia, as has been found in patients during a TGA attack (Hodges & Ward, 1989). Figure 7b and Figure 7c show how, during the resolution of the simulated TGA, more and more of the old patterns become available. The old patterns rapidly return to near-normal performance relative to the more recent ones, and even exceed their normal level (for comparison the results of simulation 1 are also drawn in the figure). We thus find the shrinkage of retrograde amnesia that was observed during recovery of TGA by Hodges and Ward, in which the amnesia first resolves for the older memories and only later for the more recent ones. Figure 7d shows performance on the fourth test, after the simulated TGA attack has resolved. Performance on all patterns is back to normal, except for the pattern that has been learned during the simulated TGA attack: Performance for this pattern is extremely low. This is typical for TGA patients, who after the attack have dense amnesia for the period of the TGA itself.

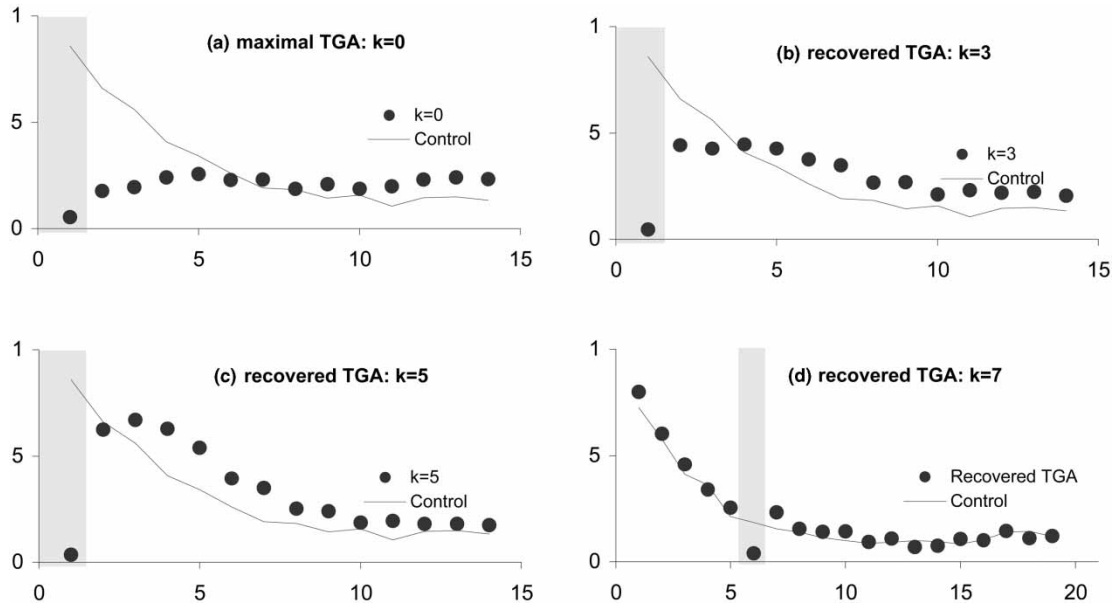

Figure 7. *Performance in the simulation of Transient Global Amnesia (TGA). A grey hatched area indicates a period of TGA. (a) Performance during the attack. The line marked "onset" indicates the onset of the attack. The model shows anterograde amnesia for the pattern learned after the onset of the TGA (pattern 1), and temporally graded retrograde amnesia for the patterns learned before the attack (patterns 2 to 14). (b) Gradual lifting of the TGA; 25% of the link nodes have become available again. (c) 50% of nodes are available. (d) Perfomance after TGA has lifted and 5 new patterns have been learned (labelled 1 to 5). There is only amnesia for the pattern learned during the attack; all other amnesia has resolved.*

The TraceLink model is thus able to simulate reversible amnesia and the shrinking of retrograde amnesia typically observed when the amnesia lifts. The mechanism simulated here may explain why patients after trauma often have an anterograde and retrograde amnesia that later partly resolves. Their lesions may upset the balance of activation and inhibition, and lead to an abnormally low activity level in link structures. The difficulties caused by this low level of activity (illustrated by the present simulation) may add to the memory deficits caused by the lesions. Though genuine restitution may also play a role, part of the resolution of amnesia may be caused by a return to normal activity levels in link structures. This entails the empirical claim that if patients suffer from temporary amnesia, one should be able to detect a pathologically low activity level within link structures (e.g., within structures of the medial temporal lobe).

ANTEROGRADE AMNESIA AND ITS CORRELATION WITH RETROGRADE AMNESIA

Anterograde amnesia can have two causes in TraceLink: (1) a lesion or dysfunction of the link system, and (2) a lesion or dysfunction of the modulatory system. These two causes of anterograde amnesia have different effects on retrograde amnesia. When the link system is lesioned, the anterograde amnesia is accompanied by retrograde amnesia. When the modulatory system is lesioned, however, there is no retrograde amnesia. The fact that two lesions can lead to anterograde amnesia, and have different consequences for retrograde amnesia, makes it possible for TraceLink to explain the correlations between anterograde and retrograde amnesia observed in different populations.

Simulation 5: Anterograde amnesia through lesioning the link layer

The first cause of anterograde amnesia in TraceLink is a lesion to the link system. This can be simulated by deactivating the link layer, followed by learning patterns with only the trace layer functioning. We simulated four degrees of lesioning of the link layer: of the entire link layer, three quarters of the link layer, one half of the link layer, and only one quarter of the link layer. In all simulations, the model learned 12 patterns before the lesion occurred. Three more patterns were learned while the link layer was lesioned, after which the model was tested.

Figure 8 shows the effects of lesioning the link layer. For comparison, the line of normal forgetting (from simulation 1) is also drawn. A lesion of the whole link layer (Figure 8a) caused an extensive retrograde amnesia with a Ribot gradient, and also a near-complete anterograde amnesia. Without a functioning link system, recently learned, not yet consolidated patterns are lost, and new patterns cannot form stable representation. These deficits are attenuated if some link nodes are still available. Lesioning three quarters of the link nodes (Figure 8b) produced a serious anterograde amnesia, and retrograde amnesia for a sizeable number of patterns preceding the lesion. Lesions of one half of the link layer (Figure 8c) caused a mild to moderate anterograde amnesia, and just a moderate loss of existing patterns. A lesion of one quarter of the link layer (Figure 8d) produced almost no retrograde amnesia, and only a mild anterograde amnesia.

TraceLink thus predicts a correlation of anterograde amnesia and retrograde amnesia when the cause of the anterograde amnesia is a lesion of the link system. Also, Ribot gradients only appear in TraceLink when there is a near-total lesion of the link system. If the lesion is less than complete, no absolute Ribot gradient appears, in the sense that recent memories are worse than equivalent remote memories. This can be seen as a prediction for the animal literature, in which equivalent memories can be used. In patient studies, however, retrograde amnesia tests are used in which items are of equivalent recall probability. As described above, this means that remote memories must actually have been stronger at encoding. Patient data can therefore best be compared to the results of the lesioned model divided by the results of the control simulations. As inspection of Figure 8 will show, TraceLink therefore still predicts a relative Ribot gradient in scores on retrograde amnesia tests if link lesions are sizeable but not complete (75%, 50%).

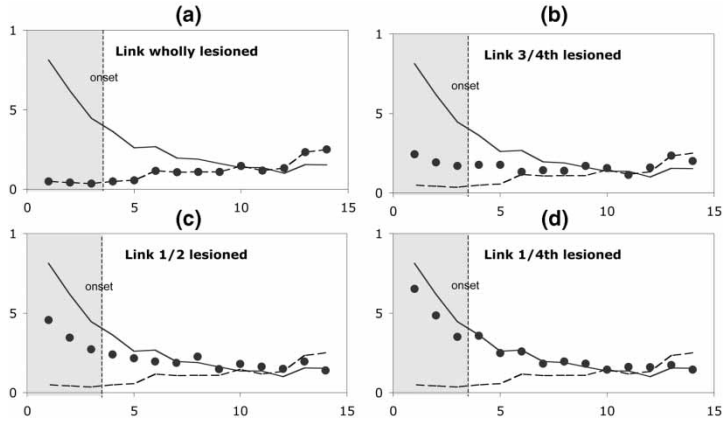

Figure 8. *Results of the simulation of anterograde amnesia through lesioning of the link layer. In all simulations, the model learns 12 patterns, labelled 4 to 15, before the lesion occurs. After the lesion (marked by the line labelled "onset"), the model learns 3 more patterns. (a) Performance after deactivation of all of the link layer; (b) deactivation of 75% of the link layer; (c) 50% of the link layer; (d) 25% of the link layer. For comparison, the continuous line gives performance of the intact model, the dashed one of the whole lesion condition.*

Moreover, our model predicts that small link lesions produce mild anterograde amnesia accompanied by hardly noticeable retrograde amnesia. This may explain the deficits of a category of patients that is usually referred to as "showing isolated anterograde amnesia." Patients with lesions limited to field CA1 of the hippocampus tend to show mild to moderate anterograde amnesia, and only mild retrograde amnesia for a limited period before the lesion (Rempel-Clower et al., 1996; Zola-Morgan, Squire, & Amaral, 1986). This is similar to the performance of the model when only one quarter of the link layer is lesioned.

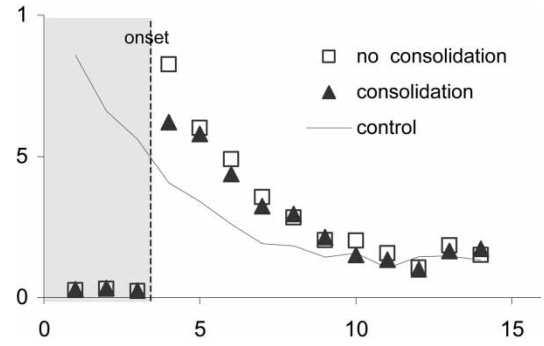

Figure 9. *Results of the simulation of anterograde amnesia through lesioning of the modulatory system. After 12 patterns have been learned, the modulatory system was lesioned. "Control" refers to the control simulation without anterograde amnesia (simulation 1). "Consolidation" refers to the simulation in which the model continued consolidating patterns after the lesion, "No consolidation" to the simulation in which that did not occur. After the lesion (marked by the line labeled "onset"), the model learned 3 more patterns.*

Simulation 6: Anterograde amnesia through lesioning the modulatory system

Another way anterograde amnesia can occur in the TraceLink model is by a lesion of the modulatory system. Since the modulatory system is not implemented in these simulations, we simulated the effects of such a lesion by a decrease in the plasticity of the link system. The learning rate in the link layer was, after the lesion, set at a base rate value equal to the learning rate in the trace layer (see Table 1). The simulation started with the normal acquisition and consolidation of 12 patterns. Then the simulated lesion of the modulatory system was applied. Three more patterns were presented to the model after the lesion, after which the model was tested.

Because the modulatory system in TraceLink has a role in regulating plasticity during consolidation, it is unclear whether after a lesion of this system consolidation continues in patients. We therefore simulated two conditions. In the first condition, consolidation continued after the lesioning of the modulatory system; between acquisition of two patterns, there was a normal consolidation phase. In the second condition, there was no more consolidation after the lesion.

Figure 9 shows the effect of lesioning the modulatory system. For both conditions, the lesion produced very dense anterograde amnesia, but no retrograde amnesia. In fact, performance for patterns learned before the lesion was better than in the simulation of normal forgetting (which is drawn in the figure for comparison). Comparison of the two conditions, with and without consolidation after the lesion, shows that improvement in pattern recall is weaker when there is consolidation after the lesion has occurred. Learning during consolidation may have an unlearning effect in the second condition. Nevertheless, the effect is there in both conditions. It also has a simple explanation. In the normal simulation, the model learns new patterns that interfere with the older ones, which induces forgetting and a lower performance for the older patterns. When the modulatory system is lesioned, however, new patterns are learned only faintly in the link layer, so that these patterns interfere little with the patterns that are already engrained in the link layer. Therefore little forgetting occurs in the link layer for the patterns learned before the lesion, and these patterns remain strong.

TraceLink thus makes the rather counterintuitive prediction that there is a class of patients who have dense anterograde amnesia and no retrograde amnesia, but instead have better-than-normal memory for the events in the time right before the lesion. These would be patients in which the structures

that correspond with the modulatory system are lesioned, but none of those that correspond to the link system. Up to this moment, no such patient has been reported in the literature: Every patient with enduring moderate to severe anterograde amnesia has at least a short period—sometimes just a few weeks—of retrograde amnesia. Possibly, lesions limited to the modulatory system do not occur because link structures and modulatory system structures are overlapping. However, the predictions given here could be tested using scopolamine, a drug that produces anterograde amnesia without retrograde amnesia (though scopolamine would have to be administered for a longer period than it typically is) or in a study with experimental animals.

The correlation between retrograde and anterograde amnesia

Two mechanisms can lead to anterograde amnesia in the TraceLink model: a lesion or dysfunction of the link system, and a lesion or dysfunction of the modulatory system. Together, these mechanisms can explain the whole spectrum of anterograde–retrograde correlations. Loss in the link system, on the one hand, will cause both retrograde and anterograde amnesia, and their severity will show a correlation. Dysfunction of the modulatory system, on the other hand, makes it hard to form new representations, but has no effect on the existing ones.

For lesions limited to structures corresponding to the link system (such as the parahippocampal region), TraceLink predicts a strong correlation between anterograde amnesia and retrograde amnesia, and also between the latter two and the size of the lesion. Diffuse lesions that affect the link, trace, and modulatory systems equally will show more variability in the pattern of amnesia, but in general anterograde and retrograde amnesia would still be expected to correlate with each other, and with the severity of the lesion. This may explain the correlations found with closed-head injury.

If lesions in the modulatory system are disproportionally large, TraceLink predicts a large anterograde relative to retrograde amnesia. This may be the case in Alzheimer's disease, in which the hippocampus and basal forebrain is strongly affected relative to other brain structures (Hyman, Hoesen, Damasio, & Barnes, 1984; Van Hoesen, 1990; Whitehouse, Price, Struble, Clark, Coyle, & DeLong, 1982). Indeed, disproportionate anterograde amnesia is often a feature of the first stages of Alzheimer's disease (Spaan, 2003). From this analysis it follows that the correlations between anterograde and retrograde amnesia should be much lower in patients with Alzheimer's disease than in cases of isolated hippocampal damage.

ISOLATED RETROGRADE AMNESIA

Isolated retrograde amnesia, also known as focal retrograde amnesia (Kapur, 1993), is retrograde amnesia without accompanying anterograde amnesia. We simulated it in TraceLink by a rupture or deterioration of the connections between the link system and the trace system. These connections code for the learned patterns, and when they are lost, retrograde amnesia occurs. If, after such loss, it is possible for the system to form new connections between the trace and link systems, new memories can still be formed and little anterograde amnesia need occur. It is unlikely, however, that immediately after a rupture of old connections new connections can be formed, as it may take some time before enough synapses are available (Robertson & Murre, 1999). This can explain the initial anterograde amnesia that later resolves, leaving isolated retrograde amnesia.

Simulation 7: Lesioning the connections between trace and link

The model initially learned 12 patterns, after which the connections between the link layer and the trace layer were severed. To simulate a partial lesion, all weights of these connections were multiplied by a random factor varying with a uniform distribution function between 0 and .20. After the lesion, weights were thus only between 0 and 20% of their initial value (such a connectivity lesion is appropriate if each connection models a number of axons, of which a certain percentage is cut). After

this, the model learned four more patterns. To simulate the gradual reappearance of connections, the learning rate for the connections from trace to link and vice versa was set to half its usual value, and moved back to its usual value with an exponential function ($1-0.5^x$, where x is the number of the pattern after the lesion). Finally, the model was tested.

Figure 10 gives the result of the simulation and, as a comparison, the control simulation. As can be seen, the lesion resulted in a dramatic loss of prelesion patterns. The first patterns learned after the lesion were not learned well due to the lower learning rate in the connections between the trace and link layer. This corresponds to retrograde amnesia and anterograde amnesia immediately after the lesion. For the most recent patterns, performance was normal. The model did not suffer from any residual anterograde amnesia. The retrograde amnesia therefore corresponds to an isolated retrograde amnesia.

Because cases of isolated retrograde amnesia are rare, there is currently little knowledge about the precise neurological basis of the syndrome. It is thus not yet possible to validate our approach toward isolated retrograde amnesia against neurological data. The structures implicated in isolated retrograde amnesia, anterior structures of the temporal lobe such as the parahippocampal region and the temporal pole, are structures that link the hippocampus with areas in the neocortex. Squire and Alvarez (1995) favour the theory that with isolated retrograde amnesia the knowledge base itself is damaged, leading to dense retrograde amnesia, with flat temporal gradients. There is good evidence that this pattern occurs in some patients (Kapur, 1993). In other cases, however, the isolated retrograde amnesia shows evidence of a Ribot gradient (Kapur, 1993). A combination of Squire's hypothesis and the one offered above may be needed to explain all cases of isolated retrograde amnesia.

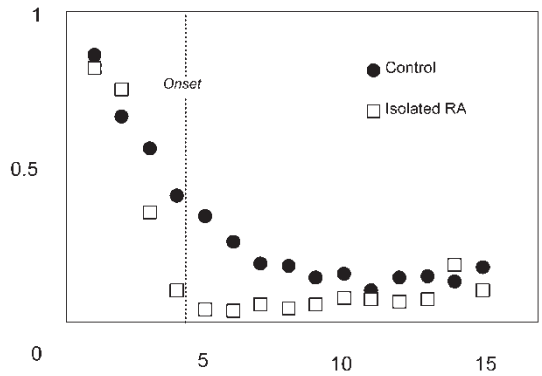

Figure 10. *Results of the simulation of isolated retrograde amnesia. After 12 patterns had been learned, the connections between the trace layer and the link layer were disturbed by multiplying them with a number between 0 and 0.20. This is marked by the line labeled "onset." After the lesion, patterns were learned with a lower learning parameter for the connections between trace and link. The learning parameter moved back to its normal value with an exponential function (see text for details).*

IMPLICIT LEARNING IN AMNESIA

When a new pattern is learned, strong connections are formed between activated nodes in the trace system and the link system, and among the activated nodes in the link system. Some learning also occurs between the nodes in the trace system, though at a much lower rate than for the connections within link or between trace and link. This learning alone is not enough to sustain a new memory after a single trial, as was shown in the simulations of anterograde amnesia.

When trace nodes are activated that already have connections between them, however, these existing connections will be strengthened. Such strengthening of existing memories in the trace system, as opposed to forming new memories via the link system, is how TraceLink models implicit learning. The increment of the connection strengths in the trace system may also offer a way to model the acquisition of new skills and of low-level knowledge in amnesic patients. Every presentation of a pattern leads to a small increment in the strength of connections in the trace system. Slowly these increments may accrue until a strong bond is formed between the nodes in a pattern. This approach to implicit learning is similar to the way that connectionist models have tackled implicit tasks such as gradual build-up of prototypes

(McClelland & Rumelhart, 1985), artificial grammar learning (Cleeremans & McClelland, 1991), and repetition priming (Stark & McClelland, 2000).

The link system and modulatory system are both crucial for the formation of new connections, but need not be involved in further strengthening of existing connections. If implicit learning were dependent on the modulatory system, we would expect implicit memory tasks to be sensitive to modulating factors such as arousal, which they are not (Gold, 1995; Jacoby & Dallas, 1981). The link system and modulatory system can nevertheless be active during this process, but this merely implies that implicit and explicit memories typically form together in normal subjects (Jacoby, 1991).

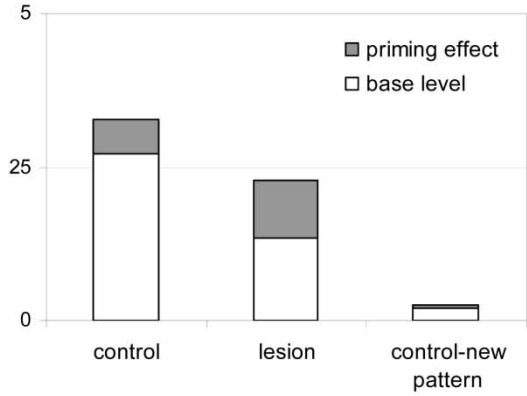

Figure 11. *Results of the simulation of implicit memory. Plotted is the enhancement in performance after an implicit learning trial (priming effect), on top of base level performance. The control-new pattern refers to a new unlearned pattern receiving an implicit learning trial. Base level here is random activity of the pattern nodes.*

Simulation 8: Implicit learning

Our simulation of implicit learning was kept relatively simple. We simulated implicit learning as learning without involvement of either the link layer or the modulatory system. First, 15 patterns were learned with normal involvement of all systems in the model to simulate the memory of a subject who comes into an experiment. After that, two random patterns were given an additional simulated implicit learning trial. In this trial, only the trace portion of the pattern was activated. Learning occurred within the trace layer, with the normal trace learning parameter. The model was tested both before and after the implicit trial. To generate enough data points per pattern, this simulation was replicated 1350 times.

Figure 11 shows the results of the simulation, contrasting performance before and after an item had received an implicit learning trial. As a comparison, the learning of a new pattern via an implicit learning trial is also shown. The implicit learning trial had a substantially greater impact for an already learned pattern than for the new pattern. Recent patterns benefited more than old patterns, though the effect was small. Furthermore, the impact of the implicit learning trial was greater in the condition with a deactivated link layer than in the normal condition. This reflects the fact that in the normal condition performance is partly a function of the (unchanging) strength of the pattern in the link layer. However, the most important conclusion is that, in TraceLink, implicit learning benefits both simulated normal controls and simulated retrograde amnesia patients. With a failing link system or modulatory system, implicit learning is still intact because strengthening of existing connections is independent of the link system and the modulatory system. Implicit learning is, therefore, preserved in amnesia.

VARIATIONS IN ENCODING

In all simulations already discussed, the model learned equivalent, equally strong patterns. Variations in strength emerged during the simulation because some patterns had more overlap with the rest of the patterns than others (all patterns were random) or because some patterns were chosen more often in the random consolidation process, and not because of differences in the original learning conditions. This is not a very realistic assumption. We therefore investigated variations in initial pattern strength, which can be seen as reflecting the effects of arousal on memory.

In TraceLink, variations in the strength of acquisition result from the action of the modulatory system, which modulates learning in the link system. High activation of this system results in a higher learning rate in the link system, and thus stronger connections between the nodes in the link part of a pattern. Activation of the modulatory system does not change the learning rate in the trace system. This choice was motivated by parsimony, but there is also empirical evidence for a lesser modulation of learning in the neocortex compared to learning in the hippocampus. Neuromodulators that enhance explicit memory, for example, have no detectable influence on performance in implicit memory tasks (Gold, 1995).

Simulation 9: The effect of arousal

We simulated the functioning of the modulatory system by manually setting the learning rate in the link layer to a higher value for some patterns, and to a lower one for others. The learning rate in the link layer varied with a continuous distribution function around the link learning rate given in Table 1. In the low variance simulation, the learning rate varied between 90% and 110% of that value; in the high variance simulation it varied between 50% and 150% of the generic rate. For reasons of computational economy, 18 patterns were learned in a simulation instead of the usual 15. The model was tested twice, once with a functioning link layer, and once with a deactivated link layer to investigate the effect of variance in pattern strength on the Ribot gradient.

To summarise the influence of the variation in strength, we did an analysis of variance on the results of the simulation. We calculated the percentage of the variance in the results of the normal test that was explained by two variables: the variations in the learning rate, and the order in which the patterns are learned (i.e., the age of a pattern at test). When all patterns are learned with equal strength (i.e., simulation 1), a large part of the variance in the simulations is explained by whether a pattern is learned recently or some time ago (see Table 2). When patterns are not all equally strong, however, much of the test variance is explained by the variation in initial strength, and only little by the order in which the patterns were learned. This translates into a much flatter forgetting curve (see Figure 12a, open triangles): Strongly learned old patterns do not necessarily perform worse than more recent but weaker patterns. The Ribot gradient, shown in Figure 12b, is not flatter but instead slightly steeper when variance is introduced in the strength of patterns.

In the high variance simulation, we divided the patterns into strong ones, learned with an initial

Table 2. *Percentages of the variance in the results of the arousal simulations explained by the age of a pattern (determined by the order in which patterns were learned), the variance in the learning rate during initial acquisition, by the interaction between the two, and the remaining variance, shown under "error"*

Variance explained by . . .	*No variance in strength*[a]	*Small variance in strength*[b]	*Strong variance in strength*[c]
"Age" of a pattern	38%	26%	6.5%
Strength of a pattern		8%	33%
Interaction between "age" and strength		5%	6.5%
Error	62%	61%	54%

[a] Simulation in which there was no variance in learning rate (i.e., simulation 1).

[b] Simulation where learning rate during acquisition varied between 90% and 110% of its standard value.

[c] Simulation where learning rate during acquisition varied between 50% and 150% of its standard value.

strength between 120% and 150% of the normal value, medium ones, learned with between 80–120% of the normal learning rate, and weak patterns, with strength between 50% and 80% of the normal learning rate (see Figure 12c and Figure 12d). This showed that the strong patterns were not only better recalled immediately after learning, but that the effect persists for old patterns. For the normal simulation, high initial strength of patterns results in a long-term primacy effect, in which the oldest patterns have an advantage over less remote ones (see Figure 12c). This effect is reminiscent of the primacy effect sometimes observed in remote memory. Sehulster (1989) examined his memory for opera performances in his 25 years as a season ticket holder, and found a typical serial position curve with recency and primacy effects: His memory was best for the most recent ones, but it was also good for the first seasons in which he held season tickets. For the Ribot curve, in the high variance simulation it can be attributed almost completely to the strong patterns (Figure 12d, filled triangles). There was no Ribot gradient in the performance of the weak patterns.

The results thus show that the variations in strength had a prolonged influence on the retrievability of the pattern. This is surprising, since only

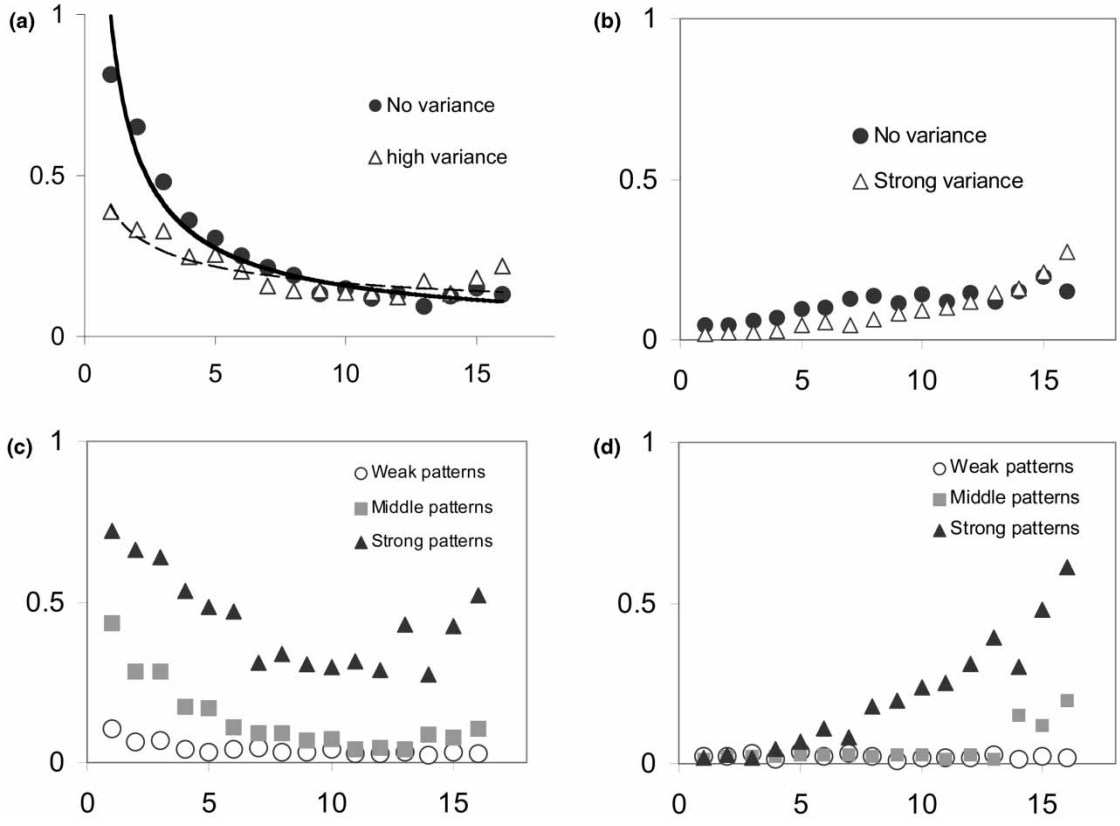

Figure 12. *Results of the simulation of the effects of arousal on memory. (a) Contrast between results with strong variance in learning rate and with no variance in the learning rate. (b) Same as panel (a) for the lesioned model. (c) Simulation with strong variance in learning rate, with patterns subdivided into those learned either with a high, low, or medium learning rate (strong = learned with 150% to 120% of the normal learning rate, middle = 80% to 120%, weak = 50% to 80%). (d) Same as panel (c) for the lesioned model.*

the initial strength of the link portion of the patterns was varied, and this portion is unlearned relatively fast. The explanation for this finding is that strong patterns tend to monopolise consolidation resources. If weak patterns are learned after a strong one, the strong one tends to be the one that is consolidated. Though the strong pattern is subsequently unlearned in the link system, its trace portion tends to become so strong that the pattern remains a likely winner in the competition for consolidation resources. We checked which patterns were consolidated in the simulation with high variance, and found that 85% of the consolidated patterns were strong patterns, 14% medium patterns, and only 1% consisted of weak patterns. Through consolidation, the patterns with strong connections in the link system thus also become strong in the trace system.

TraceLink thus predicts that stronger patterns, e.g., emotional ones, are consolidated more often, and thus are forgotten more slowly than weaker, e.g., nonemotional, patterns. Slower forgetting for episodes learned under high arousal has indeed been observed (Burke, Heuer, & Reisberg, 1992; LaBar & Phelps, 1998). The finding that strong patterns remain strong through more-than-average consolidation may hint at the process behind flashbulb memories, memories of gripping events that years later seem to be remembered with as much detail and clarity as on the first day (R. Brown & Kulik, 1977). If consolidation indeed works as a process in which patterns compete for resources, then TraceLink predicts that strong patterns will win this Darwinian competition and that the difference between strong and weak representations will grow with time.

GENERAL DISCUSSION

We have presented in this paper a model of normal memory and amnesia. The model is able to simulate the normal forgetting curve, and several forms of amnesia. It does so on the basis of a few assumptions shared with many other models: (1) that there is a consolidation process in long-term memory; (2) that there is a hierarchy in the brain, with a link system that helps retrieve patterns in the trace system; (3) that some brain structures are part of a modulatory system that influences link plasticity, making it higher if the circumstances suggest advantages to rapid new storage.

The consolidation assumption has been defended or used by many theoreticians. This prompted Nadel and Moscovitch (1997) to call consolidation theory the "Standard Model" of retrograde amnesia. TraceLink comes close to the core of that standard model, and thus has strong similarities with other implementations of the theory. At least two other models have explored consolidation in the context of hippocampal–cortical interactions (Alvarez & Squire, 1994; McClelland et al., 1995). Moreover, our implementation of consolidation—as the strengthening of attractors surfacing out of noise—has been explored in several more theoretical simulation studies (Ans & Rousset, 2000; Robins, 1995; see Meeter, 2003a, for more discussion of this work).

However, several features make TraceLink stand out among other incarnations of consolidation theory. TraceLink is the first model to combine a study of amnesia with an automatic consolidation process with random cueing of remote memories. If consolidation indeed takes place during rest or during certain sleep stages, then consolidation must occur without outside steering. McClelland et al. (1995) concentrated on the consequences of consolidation for the neocortical store, and thus did not implement how consolidation is steered in a neural network (instead, an algorithm determined which pattern was consolidated). Consolidation was also not completely autonomous in the model by Alvarez and Squire (1994): A random hippocampal representation was activated that then activated the neocortical memory to consolidate. Moreover, the model of Alvarez and Squire (1994) is very small: Just two nonoverlapping memories are stored in the system. This makes it impossible for that model to carry out many of the simulations reported here.

Another novel element of TraceLink is the inclusion of a modulatory system, which allows for a better explanation for patterns of correlation between anterograde and retrograde amnesia.

Without it, the two forms of amnesia can only correlate perfectly. Moreover, it allowed for the exploration of the effect of arousal on consolidation, and the discovery of the strong interaction between memory strength and consolidation that may explain the time course of the effect of emotional content on memories.

The connectionist simulations themselves were kept on a qualitative level: We did not attempt to quantitatively fit concrete data sets. This is mainly because the data in the neuropsychological literature are generally noisy due to the low numbers of patients available for testing. This problem is further aggravated by an understandable tendency among neuropsychologists to focus on interesting single cases. Instead of fitting concrete data sets, we therefore tried to simulate the direction and shape of some of the principal effects reported in the literature. We have also developed a mathematical theory of learning, memory, and amnesia that is more amenable to quantitative fits of noisy and distorted data sets (Murre, Chessa, & Meeter, 2004).

The breadth of phenomena to which TraceLink has been applied also sets it apart from other models. TraceLink shows how the "standard view" can account for anterograde amnesia, correlational patterns, shrinkage, intact implicit memory, and other neuropsychologically relevant characteristics, while remaining consistent with characteristics of normal forgetting. Elsewhere, we show how the TraceLink model without any modifications can also simulate some of the principal characteristics of semantic dementia (Meeter & Murre, 2004).

The simulations presented here and elsewhere by no means prove that remote memories are consolidated to the neocortex, as there are at least two cogent alternative explanations of the Ribot gradient (Meeter & Murre, in press). However, the simulations do show that consolidation offers a coherent explanation for a wide variety of findings in the neuropsychology of memory. Neither multiple trace theory (Nadel & Moscovitch, 1997; Nadel, Samsonovitch, Ryan, & Moscovitch, 2000) nor what one could call the semantisation hypothesis (Cermak, 1984; Meeter & Murre, in press; Rosenbaum, Winocur, & Moscovitch, 2001) has been worked out in much detail. Neither hypothesis has been applied to anything more than the Ribot gradient, and in the case of multiple-trace theory, to semantic dementia (Moscovitch & Nadel, 1999). Until these theories are proven to be able to account, for example, for shrinkage, isolated anterograde and retrograde amnesia, and implicit memory in amnesia, the current simulations give consolidation theory one leg up.

Manuscript received 20 May 2003
Revised manuscript received 18 February 2004
Revised manuscript accepted 27 May 2004
PrEview proof published online 10 January 2005

REFERENCES

Abeles, M. (1991). *Corticonics: Neural circuits of the cerebral cortex*. Cambridge: Cambridge University Press.

Alvarez, R., & Squire, L. R. (1994). Memory consolidation and the medial temporal lobe: A simple network model. *Proceedings of National Academy of Sciences (USA), 91,* 7041–7045.

Amaral, D. G., Ishizuka, N., & Claiborne, B. (1990). Neurons, numbers and the hippocampal network. In J. Storm-Mathisen, J. Zimmer, & O. P. Otterson (Eds.), *Understanding the brain through the hippocampus, Vol. 83* (pp. 1–11). Amsterdam: Elsevier.

Anderson, J. R., & Schooler, L. J. (1991). Reflections of the environment in memory. *Psychological Science, 2,* 396–408.

Ans, B., & Rousset, S. (2000). Neural networks with a self-refreshing memory: Knowledge transfer in sequential learning tasks without catastrophic interference. *Connection Science, 12,* 1–19.

Baddeley, A. D. (1990). *Human memory: Theory and practice*. Hove, UK: Lawrence Erlbaum Associates Ltd.

Bahrick, H. P. (1984). Semantic memory content in permastore: Fifty years of memory for Spanish learned in school. *Journal of Experimental Psychology: General, 113,* 1–27.

Bahrick, H. P. (1992). Stabilized memory of unrehearsed knowledge. *Journal of Experimental Psychology: General, 121,* 112–113.

Bahrick, H. P., Bahrick, P. O., & Wittlinger, R. P. (1975). Fifty years of memory for names and faces: A cross-sectional approach. *Journal of Experimental Psychology: General, 104,* 54–75.

Braitenberg, V., & Schüz, A. (1991). *Anatomy of the cortex: Statistics and geometry*. Berlin: Springer Verlag.

Brown, A. S. (2002). Consolidation theory and retrograde amnesia in humans. *Psychonomic Bulletin and Review, 9,* 403–425.

Brown, R., & Kulik, J. (1977). Flashbulb memories. *Cognition, 5,* 73–99.

Burke, A., Heuer, F., & Reisberg, D. (1992). Remembering emotional events. *Memory and Cognition, 20,* 277–290.

Burwell, R. D. (2000). The parahippocampal gyrus: Corticocortical connectivity. *Annals of the New York Academy of Sciences, 911,* 25–42.

Cabeza, R., & Nyberg, L. (2000). Imaging cognition II: An empirical review of 275 PET and fMRI studies. *Journal of Cognitive Neuroscience, 12,* 1–47.

Cahill, C., & Frith, C. (1995). Memory following electroconvulsive therapy. In A. D. Baddeley, B. A. Wilson, & F. N. Watts (Eds.), *Handbook of memory disorders* (pp. 319–335). Chichester, UK: John Wiley & Sons.

Carpenter, G. A., & Grossberg, S. (1988). The ART of adaptive pattern recognition by a self-organizing neural network. *Computer, 21,* 77–88.

Cermak, L. S. (1984). The episodic/semantic distinction in amnesia. In L. R. Squire & N. Butters (Eds.), *The neuropsychology of memory*. New York: Guilford Press.

Cleeremans, A., & McClelland, J. L. (1991). Learning the structure of event sequences. *Journal of Experimental Psychology: General, 120,* 235–253.

Conway, M. A., Cohen, G., & Stanhope, N. (1991). On the very long term retention of knowledge acquired through formal education: Twelve years of cognitive psychology. *Journal of Experimental Psychology: General, 120,* 395–409.

Doya, K. (2000). Complementary roles of basal ganglia and cerebellum in learning and motor control. *Current Opinion in Neurobiology, 10,* 732–739.

Doya, K., Dayan, P., & Hasselmo, M. E. (2002). Computational models of neuromodulation. Special Issue. *Neural Networks, 15,* 475–774.

Ebbinghaus, H. (1885). *Über das Gedächtnis.* Leipzig, Germany: Dunker.

Eichenbaum, H., Cohen, N. J., Otto, T., & Wible, C. G. (1992). Memory representation in the hippocampus: Functional domain and functional organization. In L. R. Squire, G. Lynch, N. M. Weinberger, & J. L. McGaugh (Eds.), *Memory: Organization and locus of change*. Oxford: Oxford University Press.

Evans, J., Wilson, B. A., Wraight, E. P., & Hodges, J. R. (1993). Neuropsychological and SPECT scan findings during and after global amnesia: Evidence for the differential impairment of remote episodic memory. *Journal of Neurology, Neurosurgery and Psychiatry, 56,* 1227–1230.

Ezrokhi, V. L., Zosimovskii, V. A., Korshunov, V. A., & Markevich, V. A. (1999). Restoration of decaying long-term potentiation in the hippocampal formation by stimulation of neuromodulatory nuclei in freely moving rats. *Neuroscience, 88,* 741–753.

Felleman, D. J., & Van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex, 1,* 1–47.

French, R. M. (1999). Catastrophic forgetting in connectionist networks. *Trends in the Cognitive Sciences, 3,* 128–135.

Gabrieli, J. D. E. (1998). Cognitive neuroscience of human memory. *Annual Review of Psychology, 49,* 87–115.

Gold, P. E. (1995). Modulation of emotional and nonemotional memory: Same pharmacological systems, different neuroanatomic systems. In J. L. McGaugh, N. M. Weinberger, & G. Lynch (Eds.), *Brain and memory: Modulation and mediation of neuroplasticity* (pp. 41–74). New York: Oxford University Press.

Greilich, H. (1984). *Quantitative Analyse der cortico-corticalen Fernverbindungen bei der Maus*. Unpublished thesis. Tuebingen, Germany: University of Tuebingen.

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

Grossberg, S. (1987). *The adaptive brain. Volume I: Cognition, learning, reinforcement, and rhythm. Volume II: Vision, speech, language, and motor control*. Amsterdam: North-Holland.

Hasselmo, M. E. (1994). Runaway synaptic modification in models of the cortex: Implications for Alzheimer's disease. *Neural Networks, 7,* 13–40.

Hasselmo, M. E. (1995). Neuromodulation and cortical function: Modeling the physiological basis of behavior. *Behavioural Brain Research, 67,* 1–27.

Hasselmo, M. E., Bodelón, C., & Wyble, B. P. (2002). A proposed function for hippocampal theta rhythm: Separate fases of encoding and retrieval enhance reversal of prior learning. *Neural Computation, 14,* 793–817.

Hasselmo, M. E., Schnell, E., & Barkai, E. (1995). Dynamics of learning and recall at excitatory recurrent synapses and cholinergic modulation in rat hippocampal region CA3. *Journal of Neuroscience, 15,* 5249–5262.

Hasselmo, M. E., & Wyble, B. P. (1997). Free recall and recognition in a network model of the hippocampus: Simulating effects of scopolamine on human memory function. *Behavioural Brain Research, 89,* 1–34.

Hebb, D. O. (1949). *The organization of behavior*. New York: Wiley.

Hodges, J. R., & Ward, C. D. (1989). Observations during transient global amnesia. *Brain, 112,* 595–620.

Hyman, B. T., Hoesen, G. W. V., Damasio, A. R., & Barnes, C. L. (1984). Alzheimer's disease: Cell-specific pathology isolates the hippocampal formation. *Science, 225,* 1168–1170.

Izquierdo, I., Quillfeldt, J. A., Zanatta, M. S., Quevedo, J., Schaeffer, E., Schmitz, P. K., & Medina, J. H. (1997). Sequential role of hippocampus and amygdala, entorhinal cortex and parietal cortex in formation and retrieval of memory for inhibitory avoidance in rats. *European Journal of Neuroscience, 9,* 786–793.

Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional use of memory. *Journal of Memory and Language, 30,* 513–541.

Jacoby, L. L., & Dallas, M. (1981). On the relationship between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General, 110,* 306–340.

Jensen, O., Idiart, M. A., & Lisman, J. E. (1996). Physiologically realistic formation of autoassociative memory in networks with theta/gamma oscillations: The role of fast NMDA channels. *Learning and Memory, 3,* 243–256.

Johnson, L. L., & Moberg, G. P. (1980). Adrenocortical response to novelty stress in rats with dentate gyrus lesions. *Neuroendocrinology, 30*, 187–192.

Kapur, N. (1993). Focal retrograde amnesia in neurological disease: A critical review. *Cortex, 29,* 217–234.

Knight, R. T., & Nakada, T. (1998). Cortico-limbic circuits and novelty: A review of EEG and blood flow data. *Review of Neuroscience, 9,* 57–70.

Kopelman, M. D., & Corn, T. H. (1988). Cholinergic "blockade" as a model for cholinergic depletion: A comparison of the memory deficits with those of Alzheimer-type dementia and the alcoholic Korsakoff syndrome. *Brain, 111,* 1079–1110.

Kopelman, M. D., Wilson, B. A., & Baddeley, A. D. (1989). The Autobiographical Memory Interview: A new assessment of autobiographical and personal semantic memory in amnesic patients. *Journal of Clinical and Experimental Psychology, 11,* 724–744.

Kritchevsky, M. (1992). Transient global amnesia. In L. R. Squire & N. Butters (Eds.), *The neuropsychology of memory* (pp. 147–155). New York: Guilford Press.

LaBar, K. S., & Phelps, E. A. (1998). Extra consolidation for arousing memories. *Psychological Science, 9,* 490–493.

Lopes da Silva, F. H., Witter, M. P., Boeijinga, P. H., & Lohman, A. H. (1990). Anatomic organization and physiology of the limbic cortex. *Physiological Reviews, 70,* 453–511.

Marr, D. (1971). Simple memory: A theory for archicortex. *Philosophical Transactions of the Royal Society, London: Biological Sciences, 262,* 23–81.

Mayes, A. R., Downes, J. J., McDonald, C., Rooke, S., Sagar, H. J., & Meudell, P. R. (1994). Two tests for assessing remote public knowledge: A tool for assessing retrograde amnesia. *Memory, 2,* 183–210.

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. *Psychological Review, 102,* 419–457.

McClelland, J. L., & Rumelhart, D. E. (1985). Distributed memory and the representation of general and specific information. *Journal of Experimental Psychology: General, 114,* 159–188.

McCloskey, M., & Cohen, N. J. (1989). Catastrophic interference in connectionist networks: The sequential learning problem. In G. H. Bower (Ed.), *The psychology of learning and motivation* (pp. 109–164). New York: Academic Press.

McGaugh, J. L. (1990). Significance and remembrance: The role of neuromodulatory systems. *Psychological Science, 1,* 15–25.

Meeter, M. (2003a). Control of consolidation in neural networks: Avoiding runaway effects. *Connection Science, 15,* 45–61.

Meeter, M. (2003b). *Long-term memory disorders: Measurement and modeling*. Unpublished Doctoral dissertation, University of Amsterdam.

Meeter, M., & Murre, J. M. J. (in press). Consolidation of long-term memory: Evidence and alternatives. *Psychological Bulletin*.

Meeter, M., & Murre, J. M. J. (2004). Simulating episodic memory deficits in semantic dementia with the TraceLink model. *Memory, 12,* 272–287.

Meeter, M., Talamini, L. M., & Murre, J. M. J. (in press). Mode shifting between storage and recall based on novelty detection in oscillating hippocampal circuits. *Hippocampus*.

Milner, P. M. (1989). A cell assembly theory of hippocampal amnesia. *Neuropsychologia, 6,* 215–234.

Minai, A. A., & Levy, W. B. (1994). Setting the activity level in sparse random networks. *Neural Computation, 6,* 85–99.

Miyashita, Y. (1993). Inferior temporal cortex: Where visual perception meets memory. *Annual Review of the Neurosciences, 16,* 245–263.

Montag-Sallaz, M., Welzl, H., Kuhl, D., Montag, D., & Schachner, M. (1999). Novelty-induced increased expression of immediate-early genes c-fos and arg 3.1 in the mouse brain. *Journal of Neurobiology, 38,* 234–246.

Moscovitch, M., & Nadel, L. (1999). Multiple-trace theory and semantic dementia: Response to K. S. Graham (1999). *Trends in the Cognitive Sciences, 3,* 87–89.

Mumby, D. G., Gaskin, S., Glenn, M. J., Schramek, T. E., & Lehmann, H. (2002). Hippocampal damage and exploratory preferences in rats: Memory for objects, places, and contexts. *Learning and Memory, 9,* 49–57.

Murre, J. M. J. (1992). *Categorization and learning in modular neural networks*. Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Murre, J. M. J. (1994). A model for categorization and recognition in amnesic patients. In M. Gazzaniga (Ed.), *Proceedings of the Cognitive Neuroscience Meeting 1994* (p. 38). Andover, MA: Cognitive Neuroscience Society.

Murre, J. M. J. (1996). TraceLink: A model of amnesia and consolidation of memory. *Hippocampus, 6,* 675–684.

Murre, J. M. J. (1997). Implicit and explicit memory in amnesia: Some predictions by the TraceLink model. *Memory, 5,* 55–82.

Murre, J. M. J., Chessa, A. G., & Meeter, M. (2004). *Quantitative modeling of amnesia: A unified account*. Manuscript submitted for publication.

Murre, J. M. J., & Sturdy, D. P. F. (1995). The connectivity of the brain: Multi-level quantitative analysis. *Biological Cybernetics, 73,* 529–545.

Nadel, L., & Moscovitch, M. (1997). Memory consolidation, retrograde amnesia and the hippocampal complex. *Current Opinion in Neurobiology, 7,* 217–227.

Nadel, L., Samsonovitch, A., Ryan, L., & Moscovitch, M. (2000). Multiple trace theory of human memory: Computational, neuroimaging and neuropsychological results. *Hippocampus, 10,* 352–368.

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

O'Reilly, R. C., & McClelland, J. L. (1994). Hippocampal conjunctive encoding, storage, and recall: Avoiding a trade-off. *Hippocampus, 4,* 661–682.

O'Reilly, R. C., & Rudy, J. W. (2000). Computational principles of learning in the neocortex and hippocampus. *Hippocampus, 10,* 389–397.

Ratcliff, R. (1990). Connectionist models of recognition memory: Constraints imposed by learning and forgetting functions. *Psychological Review, 97,* 285–308.

Reed, J. M., & Squire, L. R. (1998). Retrograde amnesia for facts and events: Findings from four new cases. *Journal of Neuroscience, 18,* 3943–3954.

Rempel-Clower, N. A., Zola, S. M., Squire, L. R., & Amaral, D. G. (1996). Three cases of enduring memory impairment after bilateral damage limited to the hippocampal formation. *Journal of Neuroscience, 16,* 5233–5255.

Ribot, T. (1881). *Les maladies de la memoire*. Paris: Germer Baillare.

Robertson, I. H., & Murre, J. M. J. (1999). Rehabilitation from brain damage: Brain plasticity and principles of guided recovery. *Psychological Bulletin, 125,* 544–575.

Robins, A. V. (1995). Catastrophic forgetting, rehearsal, and pseudorehearsal. *Connection Science, 7,* 123–146.

Rosenbaum, R. S., Winocur, G., & Moscovitch, M. (2001). New views on old memories: Re-evaluating the role of the hippocampal complex. *Behavioural Brain Research, 127,* 183–197.

Rosenzweig, E. S., Barnes, C. A., & McNaughton, B. L. (2002). Making room for new memories. *Nature Neuroscience, 5,* 6–8.

Rubin, D. C., & Wenzel, A. E. (1996). One hundred years of forgetting: A quantitative description of retention. *Psychological Review, 103,* 734–760.

Russel, W. R., & Nathan, P. W. (1946). Traumatic amnesia. *Brain, 69,* 280–300.

Sanders, H. I., & Warrington, E. K. (1971). Memory for remote events in amnesic patients. *Brain, 94,* 661–668.

Schacter, D. L. (1992). Priming and multiple memory systems: Perceptual mechanisms of implicit memory. *Journal of Cognitive Neuroscience, 4,* 244–256.

Schmidtke, K., & Vollmer, H. (1997). Retrograde amnesia: A study of its relation to anterograde amnesia and semantic memory deficits. *Neuropsychologia, 35,* 505–518.

Sehulster, J. A. (1989). Content and temporal structure of autobiographical knowledge: Remembering twenty-five seasons at the Metropolitan Opera. *Memory and Cognition, 17,* 590–606.

Shimamura, A. P., & Squire, L. R. (1986). Korsakoff's syndrome: A study of the relation between anterograde amnesia and remote memory impairment. *Behavioral Neuroscience, 100,* 165–170.

Slamecka, N. J., & McElree, B. (1983). Normal forgetting of verbal lists as a function of their degree of learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 9,* 384–397.

Spaan, P. E. J. (2003). *Early assessment of dementia: The contributions of different memory components.* Unpublished dissertation, Universiteit van Amsterdam, Amsterdam.

Squire, L. R. (1992). Memory and the hippocampus: A synthesis from findings with rats, monkeys, and humans. *Psychological Review, 99,* 195–231.

Squire, L. R., & Alvarez, P. (1995). Retrograde amnesia and memory consolidation: A neurobiological perspective. *Current Opinion in Neurobiology, 5,* 169–175.

Squire, L. R., Cohen, N. J., & Nadel, L. (1984). The medial temporal region and memory consolidation: A new hypothesis. In H. Weingarter & E. Parker (Eds.), *Memory consolidation* (pp. 185–210). Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Squire, L. R., Haist, F., & Shimamura, A. P. (1989). The neurology of memory: Quantitative assessment of retrograde amnesia in two groups of amnesic patients. *Journal of Neuroscience, 9,* 828–839.

Stark, C. E. L., & McClelland, J. L. (2000). Repetition priming of words, pseudowords, and nonwords. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 945–972.

Trepel, C., & Racine, R. J. (1998). Long-term potentiation in the nocortex of the adult, freely moving rat. *Cerebral Cortex, 8,* 719–729.

Treves, A., & Rolls, E. T. (1994). Computational analysis of the role of the hippocampus in memory. *Hippocampus, 4,* 374–391.

Van Hoesen, G. W. (1990). The dissection of Alzheimer's disease of cortical and limbic neural systems relevant to memory. In J. L. McGaugh, N. M. Weinberger, & G. Lynch (Eds.), *Brain organization and memory: Cells, systems, and circuits* (pp. 234–261). Oxford: Oxford University Press.

Whitehouse, P. J., Price, D. L., Struble, R. G., Clark, A. W., Coyle, J. T., & DeLong, M. R. (1982). Alzheimer's disease and senile dementia: Loss of neurons in the basal forebrain. *Science, 215,* 1237–1239.

Whitty, C. W. M., & Zangwill, O. L. (1977). Traumatic amnesia. In C. W. M. Whitty & O. L. Zangwill (Eds.), *Amnesia* (pp. 118–135). London: Butterworths.

Wickelgren, W. A. (1974). Single-trace fragility theory of memory dynamics. *Memory and Cognition, 2,* 775–780.

Wickelgren, W. A. (1979). Chunking and consolidation: A theoretical synthesis of semantic networks, configuring in conditioning, S-R versus cognitive learning, normal forgetting, the amnesic syndrome, and the hippocampal arousal system. *Psychological Review, 86,* 44–60.

Witter, M. P., Wouterlood, F. G., Naber, P. A., & Van Haeften, T. (2000). Anatomical organization of the parahippocampal-hippocampal network. *Annals of the New York Academy of Sciences, 911,* 1–24.

Wixted, J. T., & Ebbesen, E. B. (1991). On the form of forgetting. *Psychological Science, 2,* 409–415.

Zola-Morgan, S., Squire, L. R., & Amaral, D. G. (1986). Human amnesia and the medial temporal region: Enduring memory impairment following a bilateral lesion limited to field CA1 of the hippocampus. *Journal of Neuroscience, 6,* 2950–2967.

APPENDIX A

Implementation of the model

The model is based on binary, stochastic nodes that fire synchronously. The thresholds of the nodes in a module are controlled by an inhibition mechanism: Inhibition in a module is diminished if too few nodes are active (i.e., less than a target k) and increased if too many are. At each iteration, a learning rule is applied to all connections after node activations have been updated. The details of these mechanisms are described below. All simulations were done with the Nutshell simulator (available free of charge via http://nutshell.neuromod.org).

Activation rule

A node i has an activation a_i that can take on either of two values: 0 or 1. The probability that node i will "fire" (i.e., that its activation becomes 1) increases with its net input, as follows:

$$p_i = \frac{1}{1 + e^{-\frac{net_i}{temp}}} \quad (1)$$

where net_i is the total input activation to node i, or the weighted input to node i minus inhibition:

$$net_i = \sum_{j=1}^{n} w_{ij} a_j - inhibition \quad (2)$$

where w_{ij} is the connection weight from node j to node i, a_j is the activation value of node j, and n is the number of nodes in the model (if there is no connection between j and i, w_{ij} is zero by default). Inhibition is discussed in the next paragraph. The temperature parameter *temp* in Equation 1 controls the degree of randomness of the nodes: If *temp* is near zero the nodes behave as simple threshold devices; if *temp* is high the role of the net input is limited and the node takes on values 0 or 1 randomly. We used a temperature of 0.2 in all simulations.

Threshold control

Inhibition is constantly adjusted to ensure that the total number of activated nodes in a module (called A) does not wander too far from the target number k. Each module has its own k, and inhibition control is separate for each module. To keep the number of activated nodes at time t, A_t, as close as possible to the target number k, two thresholds are constantly adapted. Inhibition is the sum of a fast-changing threshold parameter T multiplied by A_t, and a slow moving threshold τ:

$$inhibition = TA_t + \tau \quad (3)$$

TA_t, fast inhibition, may reflect the excitability of the basket cells by the excitatory neurons. Slow inhibition, τ, may reflect the autonomous activity of inhibitory cells. Fast inhibition reacts rapidly to a departure from equilibrium, and slow inhibition moves gradually to establish the equilibrium anew.

The control of T is straightforward: If A_t is higher than k, T is increased (more inhibition); if A_t is lower it is decreased. In particular, if A_t is larger than k, T is increased a lot; if A_t is only a bit larger, T is increased a little:

$$\begin{array}{l} \text{if } A_t > (1 + crit)k \\ \quad T = T + \Delta_t \\ \text{if } A_t < (1 - crit)k \\ \quad T = T - \Delta_t \end{array} \quad (4)$$

where *crit* is the criterion for deciding whether A_t is much larger or smaller, and Δ_T is the change made to T (*crit* = 0.20, and $\Delta_T = 0.01$). If A_t is only a little bit larger or smaller than k (e.g., $k < A_t < (1 + crit) * k$), then one third of Δ_T is added to or subtracted from T. To prevent violent oscillations in activity, A_t is a moving average. When A^*_t is the current level of activation, the value used to compute both the level of inhibition TA_t and the change in T is:

$$A_t = 0.5A_{t-1} + 0.5A^*_t \quad (5)$$

This precedes calculation of the new value of T (Eq. 4).

The slow inhibition process aims to keep the "slow threshold" τ equal to TA_t. When the equilibrium is disturbed, for example, if the activation is diminished due to a lesion, τ slowly decreases to a new equilibrium value. The speed of this change is determined by the parameter Δ_τ, which is chosen low (0.001). The expression for calculating τ_{t+1} at $t + 1$ is

$$\tau_{t+1} = (1 - \Delta_\tau)\tau_t + \Delta_\tau TA_t \quad (6)$$

The amount of "fast" inhibition is bounded by a minimum value T^{min} and a maximum value T^{max}. If $T < T^{min}$ it is set to T^{min}, and if $T > T^{max}$ it is set to T^{max}. Similarly, τ is also kept between upper and lower bounds: if $\tau < \tau^{min}$, τ is τ^{min}; if $\tau > \tau^{max}$, τ is τ^{max}. T^{min} and τ^{min} were set to 0. T^{max} and τ^{max} were set to such high values that they were never reached in the simulations.

Learning rule

The learning rule is a simple Hebbian rule that also allows decreases in weight. The change in weight Δw_{ij} on each time step is equal to:

$$\Delta w_{ij} = \mu^+ a_i a_j - \mu^- a_i (1 - a_j), \quad (7)$$

where μ^- and μ^+ represent the learning rates. Both μ^- and μ^+ must be larger than 0. The weights w_{ij} are kept within the interval [0, 1] by setting $w_{ij} = 1$ if $w_{ij} > 1$, and $w_{ij} = 0$ if $w_{ij} < 0$.