Memory consolidation, retrograde amnesia and the hippocampal complex
Lynn Nadel* and Morris Moscovitch†

Results from recent studies of retrograde amnesia following damage to the hippocampal complex of human and non-human subjects have shown that retrograde amnesia is extensive and can encompass much of a subject's lifetime; the degree of loss may depend upon the type of memory assessed. These and other findings suggest that the hippocampal formation and related structures are involved in certain forms of memory (e.g. autobiographical episodic and spatial memory) for as long as they exist and contribute to the transformation and stabilization of other forms of memory stored elsewhere in the brain.

Addresses
*Department of Psychology and Neural Systems, Memory and Aging Division, University of Arizona, Tucson, Arizona 85721, USA; e-mail: nadel@ccit.arizona.edu
†Department of Psychology, Erindale College, University of Toronto, Mississauga, Ontario, L5L 1C6, Canada and Department of Psychology and Rotman Institute of Baycrest Centre of Geriatric Care, 3560 Bathurst Street, North York, Ontario, M6A 2E1, Canada; e-mail: momos@credit.erin.utoronto.ca

Current Opinion in Neurobiology 1997, 7:217-227
Electronic identifier: 0959-4388
© Current Biology Ltd ISSN 0959-4388

Abbreviations
AA anterograde amnesia
PET positron emission tomography
RA retrograde amnesia

Introduction
Memory consolidation refers to the idea that neural processes transpiring after the initial registration of information contribute to the permanent storage of memory. The phenomenon of retrograde amnesia (RA) — whereby information acquired prior to a cerebral insult is lost — provides one of the major sources of evidence in support of the idea of memory consolidation because it shows that what has been learned is not instantly made permanent. The clinical syndrome and animal models of it have been used to investigate the duration of the memory consolidation process, the brain regions that are implicated, and, most recently, the forms of information involved.

In this review, we provide a synopsis of a widely accepted view of these phenomena and consider evidence gathered in the past few years that forces a reconsideration of this view. Finally, we suggest a new way to think about memory and its consolidation that may help us to understand the often confusing data generated from both clinics and experimental laboratories.

The standard model
In what we will call the 'standard' model, memory consolidation begins when information, initially registered in the neocortex, is bound into a memory trace by the hippocampus and related structures in the medial temporal lobes and diencephalon [1,2]. This initial binding of information into a memory trace involves a short-term consolidation process, or cohesion, that is believed to be completed within seconds or, at most, tens of minutes [3,4,5]. A process of long-term consolidation then begins. At first, the hippocampus and related structures are necessary for the storage and recovery of the memory trace, but their contribution diminishes as consolidation proceeds, until the neocortex alone is capable of sustaining the permanent memory trace and mediating its retrieval.

According to this model — the basic features of which have not changed substantially since Scoville and Milner's [6] report of H.M. — the hippocampal formation is viewed as a 'temporary' memory system, of use only until long-term consolidation is complete and a permanent memory is established elsewhere (in the neocortex and other structures).

For the purposes of this discussion, the hippocampal formation is taken to include the hippocampus proper, the dentate gyrus and the subiculum. Related structures in the medial temporal region that may complement the hippocampal formation in its consolidation function include the entorhinal cortex, the perirhinal cortex and the parahippocampal gyrus. Following Scoville and Milner's [6] nomenclature, we refer to this system of related medial temporal lobe structures as the hippocampal complex or the medial temporal lobe memory system.

The existence of short-term consolidation (cohesion) processes is not in dispute, although their exact nature is unclear. This review will be concerned only with long-term consolidation and, initially, with the standard model just described. Squire and Alvarez [7] have argued that two kinds of evidence support the standard model: first, demonstrations of temporally graded RA, which appear to show that after some period of time, damage to the hippocampal formation no longer causes a loss of remote memories; second, demonstrations that the severity of RA is correlated with the severity of anterograde amnesia (AA). Although Squire and Alvarez [7] noted that extensive, ungraded forms of RA can sometimes occur, they suggested that the typical finding involves temporally graded RA. They further argued that event (episode) and fact (semantic) memory are equivalently affected in RA, as both forms of memory

Addresses
*Department of Psychology and Neural Systems, Memory and Aging Division, University of Arizona, Tucson, Arizona 85721, USA; e-mail: nadel@ccit.arizona.edu
†Department of Psychology, Erindale College, University of Toronto, Mississauga, Ontario, L5L 1C6, Canada and Department of Psychology and Rotman Institute of Baycrest Centre of Geriatric Care, 3560 Bathurst Street, North York, Ontario, M6A 2E1, Canada; e-mail: momos@credit.erin.utoronto.ca

Current Opinion in Neurobiology 1997, 7:217-227
Electronic identifier: 0959-4388
© Current Biology Ltd ISSN 0959-4388

Abbreviations
AA anterograde amnesia
PET positron emission tomography
RA retrograde amnesia

Introduction
Memory consolidation refers to the idea that neural processes transpiring after the initial registration of information contribute to the permanent storage of memory. The phenomenon of retrograde amnesia (RA) — whereby information acquired prior to a cerebral insult is lost — provides one of the major sources of evidence in support of the idea of memory consolidation because it shows that what has been learned is not instantly made permanent. The clinical syndrome and animal models of it have been used to investigate the duration of the memory consolidation process, the brain regions that are implicated, and, most recently, the forms of information involved.

In this review, we provide a synopsis of a widely accepted view of these phenomena and consider evidence gathered in the past few years that forces a reconsideration of this view. Finally, we suggest a new way to think about memory and its consolidation that may help us to understand the often confusing data generated from both clinics and experimental laboratories.

The standard model
In what we will call the 'standard' model, memory consolidation begins when information, initially registered in the neocortex, is bound into a memory trace by the hippocampus and related structures in the medial temporal lobes and diencephalon [1,2]. This initial binding of information into a memory trace involves a short-term consolidation process, or cohesion, that is believed to be completed within seconds or, at most, tens of minutes [3,4,5]. A process of long-term consolidation then begins. At first, the hippocampus and related structures are necessary for the storage and recovery of the memory trace, but their contribution diminishes as consolidation proceeds, until the neocortex alone is capable of sustaining the permanent memory trace and mediating its retrieval.

According to this model — the basic features of which have not changed substantially since Scoville and Milner's [6] report of H.M. — the hippocampal formation is viewed as a 'temporary' memory system, of use only until long-term consolidation is complete and a permanent memory is established elsewhere (in the neocortex and other structures).

For the purposes of this discussion, the hippocampal formation is taken to include the hippocampus proper, the dentate gyrus and the subiculum. Related structures in the medial temporal region that may complement the hippocampal formation in its consolidation function include the entorhinal cortex, the perirhinal cortex and the parahippocampal gyrus. Following Scoville and Milner's [6] nomenclature, we refer to this system of related medial temporal lobe structures as the hippocampal complex or the medial temporal lobe memory system.

The existence of short-term consolidation (cohesion) processes is not in dispute, although their exact nature is unclear. This review will be concerned only with long-term consolidation and, initially, with the standard model just described. Squire and Alvarez [7] have argued that two kinds of evidence support the standard model: first, demonstrations of temporally graded RA, which appear to show that after some period of time, damage to the hippocampal formation no longer causes a loss of remote memories; second, demonstrations that the severity of RA is correlated with the severity of anterograde amnesia (AA). Although Squire and Alvarez [7] noted that extensive, ungraded forms of RA can sometimes occur, they suggested that the typical finding involves temporally graded RA. They further argued that event (episode) and fact (semantic) memory are equivalently affected in RA, as both forms of memory
are subserved by the same medial temporal lobe memory system.

Others have focused their attention on patients with brain lesions that spared the medial temporal region and who present with severe RA in the absence of comparably severe AA [8,9,10*,11,12]. Although such reports are invaluable for a complete understanding of the neural basis of memory, they do not provide a direct test of the standard model, whose focus is on the medial temporal lobes.

In what follows, we address four questions raised by the standard model. First, are there dissociations within medial temporal lobe RA such that some forms of remote memory are retained while others are lost? Second, what is the extent of the RA observed after medial temporal lobe lesions? Third, is RA always, or only sometimes, temporally graded following medial temporal lobe lesions? Fourth, is there a necessary relation between the extent of RA and the severity of AA?

Subtypes of episodic and semantic memory

The standard model [2] refers primarily to declarative memory, which includes episodic memory (comprising knowledge about personal events or episodes) and a subtype of semantic memory (comprising knowledge of 'facts' about the world, such as the capital of France, the Prime Minister of the United Kingdom during World War II, details of the Civil War, and knowledge of public events and personalities, and so on [13]). For practical reasons, we will not discuss at length other aspects of semantic memory, which we term 'general semantics', that include conceptual knowledge of words, grammar and objects—these lie more properly in the domains of language and object recognition than of memory. In addition, we will not deal in any detail with procedural or implicit memory, which involve the acquisition of skills and of perceptual and general semantic knowledge without awareness. Though not the focus of this review, it is of theoretical importance that knowledge of skills, as well as perceptual and general semantic knowledge are spared in subjects with RA [14,15], even if the knowledge was acquired during a period in time for which both retrograde and anterograde deficits are seen in episodic and/or personal semantic knowledge [16*,17].

Some proponents of the standard model [1,2] suggest that episodic and semantic knowledge, though different in some respects, should nonetheless be considered as part of declarative memory, which initially is dependent on the medial temporal lobes. According to this hypothesis, both episodic and semantic memory should be subject to temporally graded RA following medial temporal lobe lesions. Others, however, have viewed episodic and semantic memory as fundamentally different (see e.g. [18–22]) and have allowed for the possibility that dissociations among these and other kinds of memory might be observed in RA.

More recently, even finer distinctions have been drawn, taking into account the precise nature of the episodic and semantic aspects of memory. In the domain of autobiographical knowledge, for example, a distinction has been drawn between personal, autobiographical episodes, which contain experiential information (i.e. where, when, and in what perceptually detailed context a personally experienced event occurred) and personal semantics, which include facts about one's life (i.e. where one was born, where one went to school, who one's friends were) and which are fundamentally no different than facts about the world [8,23–25]. Even within the subset of semantic memory with which we are dealing, there may be important distinctions to be made regarding information that is closely linked to the episode within which it was acquired (e.g. knowledge about objects that have recently been experienced) [17,26], narrative information about historical events (e.g. the French Revolution) [27], and facts about the world (e.g. the names of the longest rivers).

Ultimately, these distinctions will be useful if they help us understand how normal memory works and how brain damage fractionates what, in most instances, appears to be a unified domain. A number of recently reported studies on brain damage have provided support for the view that episodic and semantic memory can be separately disturbed and, further, that autobiographical content can also be relevant in determining the effect of such damage on memory (see below).

Memory dissociations: extent and gradients of retrograde amnesia

Because we are focusing on the medial temporal lobe, we will examine primarily data from patients with damage exclusive to, or primarily in, that region. The relevant studies are summarized in Table 1. The data show that different patterns of RA can be observed when one is assessing the status of distinct forms of memory.

Autobiographical episodic memory

The duration of RA for autobiographical episodes is long-lasting, often extending as much as 25–40 years, and, in some cases, encompasses the subject's entire lifetime. Two subjects, each with hippocampal damage confined only to one subfield, are the only known exceptions to this scenario [28**,29]. Consistent with evidence from other studies, the data described in Table 1 strongly suggest that the extent of RA, and whether or not it is graded, depends on the amount of medial temporal lobe damage.

Although Rempel-Clower et al. [28**] have argued that RA for autobiographical episodes is temporally graded, their own data indicate that patients with large hippocampal lesions display no discernible gradient for the most recent three decades and have very little or no autobiographical memory; except for events that occurred before 1950, the most remote time period they tested. It is difficult to conceive of an adaptive basis for a consolidation process.
Table 1

Retrograde amnesia in humans with hippocampal complex damage. *

<table>
<thead>
<tr>
<th>Study</th>
<th>Nature of the RA</th>
<th>Extent of the RA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scoville and Milner, 1957 [6] (see also</td>
<td>Loss of autobiographical episodes,</td>
<td>Graded loss for 15 years</td>
</tr>
<tr>
<td>Corkin, 1984 [73] and Corkin, Amaral,</td>
<td>public events and personalities</td>
<td></td>
</tr>
<tr>
<td>Johnson and Hyman, in press [74+])</td>
<td>General semantics relatively spared</td>
<td></td>
</tr>
<tr>
<td>Cermak and O’Connor, 1990 [20] (see also</td>
<td>Loss of autobiographical episodes,</td>
<td>Flat loss of autobiographical episodes</td>
</tr>
<tr>
<td>O’Connor, Cermak and Seidman, 1995 [75])</td>
<td>personal semantics, public events and personalities</td>
<td>Graded, extensive loss (30–40 years) of public events and personalities</td>
</tr>
<tr>
<td>Damasio, Eisinger, Damasio and Van Hooseen,</td>
<td>Loss of autobiographical episodes</td>
<td>Flat loss of autobiographical episodes</td>
</tr>
<tr>
<td>1985 [76]+</td>
<td>Semantic memory intact</td>
<td></td>
</tr>
<tr>
<td>Zola-Morgan, Squire and Amaral, 1985 [29]</td>
<td>No loss of either autobiographical episodes or public events</td>
<td>Minor</td>
</tr>
<tr>
<td>Tulving, Schacter, McLachlan and Moscovitch,</td>
<td>Loss of autobiographical episodes</td>
<td>Flat loss of autobiographical episodes</td>
</tr>
<tr>
<td>1988 [58]+</td>
<td>Loss of public faces</td>
<td>Extended loss of public faces</td>
</tr>
<tr>
<td>Barr, Goldberg, Wasserstein and Novelly,</td>
<td>Loss of autobiographical episodes,</td>
<td>Flat loss of autobiographical episodes</td>
</tr>
<tr>
<td>Victor and Agamanolis, 1990 [33]</td>
<td>Loss of autobiographical episodes</td>
<td>Graded loss of events and personalities</td>
</tr>
<tr>
<td>McCarthy and Warrington, 1992 [77]+ (see</td>
<td>Loss of both autobiographical episodes and public events</td>
<td>extends for at least 40 years</td>
</tr>
<tr>
<td>also Warrington and McGuity, 1990 [17])</td>
<td>General semantic memory intact</td>
<td></td>
</tr>
<tr>
<td>Yoneda, Mori, Yamashita and Yamadori, 1994</td>
<td>Loss of personal and public events</td>
<td>Flat loss of both autobiographical episodes and public events</td>
</tr>
<tr>
<td>[78]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kartsounis, Hudge and Stevens, 1995 [79]</td>
<td>Loss of autobiographical memory, personal semantics and public events</td>
<td>Flat loss of both autobiographical episodes and public events</td>
</tr>
<tr>
<td>Rempel-Clower, Zola-Morgan, Squire and</td>
<td>Patient G.D.: intact autobiographical episodes; mild loss of public events</td>
<td>Flat loss of both autobiographical episodes and public events</td>
</tr>
<tr>
<td>Amaral, 1996 [28++] (see references herein</td>
<td>Patient L.M.: loss of both autobiographical episodes and public events; some</td>
<td>Patient G.D.: No discernible loss</td>
</tr>
<tr>
<td>earlier publications regarding these</td>
<td>loss of general semantics</td>
<td>Patient L.M.: virtually flat loss of autobiographical memory for over 25 years,</td>
</tr>
<tr>
<td>patients)</td>
<td>Patient W.H.: loss of both autobiographical episodes and public events</td>
<td>graded thereafter; for public events and faces, graded loss for at least 10</td>
</tr>
<tr>
<td></td>
<td></td>
<td>years; Patient W.H.: virtually flat loss of autobiographical episodes for about</td>
</tr>
<tr>
<td></td>
<td></td>
<td>30 years, graded thereafter; graded loss of public events and faces for over 25</td>
</tr>
<tr>
<td></td>
<td></td>
<td>years</td>
</tr>
</tbody>
</table>

*In all these cases, the major lesion included bilateral damage to the hippocampal complex. See Neurocase 1996, 2:259–298 for additional references and brief summaries of many articles on confirmed hippocampal complex lesions in humans, but with poorer anatomical and/or psychological documentation than for the studies listed in this table. †In some cases, damage extended beyond the medial temporal lobe. ‡A recent, careful magnetic resonance imaging (MRI) study (S Kohler, R Habib, SB Black, C Szekely, M Sinden, E Tulving, abstract, Brain Cogn 1997, in press) has shown that this patient had extensive bilateral hippocampal lesions.

that is almost as long as the average human life span throughout much of history. According to this notion, most people would not have consolidated any autobiographical memories before they died. The estimates of the time needed to consolidate autobiographical memories is so long that, for all practical purposes, the hippocampal complex would need to be involved in recovering most of them throughout one’s life.
Public events and personalities
When assessed with questionnaires about public events and tests requiring the identification of personalities from pictures of faces, RA, in some cases, is not as extensive as for autobiographical episodes and is more likely to be graded. Nonetheless, even for this type of memory, RA can extend to between 10 and 30 years. For those people and events that had only a brief stay in the public eye, RA may be as extensive and as ungraded as autobiographical memory in subjects with large medial temporal lobe lesions [30].

Personal semantics and general semantics
The evidence regarding RA for personal and general semantic knowledge is less clear. For many of the subjects, there are insufficient data about memory for either kind of knowledge to draw any firm conclusions. When it is tested, RA for personal semantic knowledge is extensive and graded but less severe than RA for autobiographical memory (see Table 1). With respect to general semantics—insofar as most subjects do not show any obvious loss of vocabulary, grammar, or object recognition—most studies report no RA for general semantic knowledge. It should be noted, however, that detailed tests designed to pick up such deficits are lacking in most reports.

A fascinating finding has been reported by Warrington and McCarthy [17], who specifically tested their subject’s general semantic knowledge for words that came into use during the period for which he has a dense RA for autobiographical episodes and public events. He knew the meaning of words such as AIDS, telecom, shuttle (referring to space) and provisionals, and could give detailed definitions of them.

This finding has been corroborated by a single-case study [31], by Barr et al. [13] in subjects with unilateral temporal lobectomies, and by Verfaellie et al. [16*], who found a similar effect in a group of non-Korsakoff amnesic patients who showed a slight, ungraded, but not significant, loss of memory for words that entered the vocabulary in the past 25 years.

By contrast, Korsakoff amnesic patients, whose lesions do not include the medial temporal lobes, did show a temporally graded amnesia for new words. To complicate matters further, two other patients with medial temporal lesions (L.M. of Rempel-Clover et al. [28**], initially reported in Beatty et al. [32], and Victor and Agamanolis' [33] patient) and one Korsakoff amnesic (see Butters and Cermak [34]) exhibited RA for terms related to their profession. One could argue that in contrast to general semantic knowledge, such terms are recovered by way of the episodic memory system [26], as no other semantic memory loss has been reported in these patients. Until evidence on RA for general semantics is gathered as systematically as that for other types of memory, the prudent conclusion is that if RA exists for general semantics, it is less severe than that for other types of memory.

Although we have confined this survey to studies with confirmed medial temporal lobe lesions, the same pattern is evident in subjects in which the medial temporal lobes are temporarily rendered dysfunctional. Most cases of transient global amnesia are associated with hypo-perfusion of the medial temporal lobes, often with relative or complete sparing of cognitive functions other than memory. While in the amnesic state, AA is accompanied by a RA that resembles in all ways the RA observed in people with very large hippocampal lesions (see [35,36*] and references therein).

Summary of the evidence
On the basis of the human data we have just reviewed, the four questions raised above can be addressed.

First, are there dissociations within medial temporal lobe RA such that some forms of remote memory are retained while others are lost? It appears that there are dissociations within RA among different types of memory. Memory for autobiographical episodes appears to be the more severely affected, followed by memory for personal semantics, public events, and persons, which seem to be equally affected. The evidence regarding general semantic memory is more variable, but a safe conclusion is that it is less severely affected than other types of memory in non-Korsakoff amnesic patients, whose lesions spare the lateral temporal cortex of the left hemisphere [37].

Second, what is the extent of the RA observed after medial temporal lobe lesions? RA is very extensive following substantial medial temporal lobe damage, ranging from 10–30 years, and sometimes longer, for facts about the world to an entire lifetime for autobiographical memory.

Third, is RA always, or only sometimes, temporally graded following medial temporal lobe lesions? The dissociations noted above are also reflected in the temporal gradients of RA. More than for other types of memory, the temporal gradient for autobiographical memory is either entirely absent or very shallow, rising only at the more remote periods; that is, extensive loss is observed for all, or almost all, time periods. Although a noticeable temporal gradient can often be observed for memory of public events and persons, RA can also be shallow or flat if the material chosen was only briefly held in the public eye.

Fourth, is there a necessary relation between the extent of RA and the severity of AA? It appears that the extent of RA following medial temporal damage is related to the extent of AA insofar as the patients with the least damage are also the ones whose amnesia, both anterograde and retrograde, is least severe. The upper limit for AA, however, may be reached quickly so that no further deficits
are observed with increasing lesion size; by contrast, RA may continue to increase with increasing lesion size. The most conservative conclusion would seem to be that the extent of RA and the slope of the RA gradient are very likely related to the extent of medial temporal lobe damage and, to a lesser degree, to the extent of AA. Extensive RA with relative sparing of AA (e.g. [38]) shows that there is no necessary linkage between these two forms of amnesia (but see De Renzi and Lucchelli [39]).

**Current status of the standard model**

On a number of important issues, the standard model is not supported by the data. There are sharp dissociations in RA between different aspects of declarative memory. The temporal gradient for autobiographical memory is shallow or absent, and RA can extend to a lifetime. Although the gradient is steeper, and the period of RA not as extensive, for personal semantics, public events, and public persons, here too the deficit can reach back as far as 30 years or more. The model would have to assume that consolidation processes last on the order of decades, or indeed the entire lifetime, given that RA lasts that long in patients with extensive damage to the medial temporal lobes.

It is unreasonable to suppose that the initial proponents of the standard model (e.g. [6]) had such long intervals in mind for the consolidation process when they proposed that memory could be subserved by extra-hippocampal structures—including most prominently the neocortex—after consolidation was complete (neither Nadel nor his collaborators had any such concept in mind in their formulation of this model [40]; see also NJ Cohen, Soc Neurosci Abstr 1987, 13:205; nor did Moscovitch subscribe to this view [4*], when he endorsed the model).

In light of current data, a more reasonable proposal is that the hippocampus and related memory structures are required for recovering even remote memories. To account for the dissociations observed between autobiographical memory and various kinds of semantic memory and for the kinds of temporal gradients that are observed, it is necessary to propose a model of how memory is represented in the hippocampal complex and neocortex throughout life and how such memory is recovered during retrieval. Before doing so, however, we think it is worth reviewing relevant literature from the study of memory consolidation in animal models.

**Animal models of memory consolidation**

The use of animal models in the study of memory consolidation has had a long tradition. Throughout the 1950s and 1960s, hundreds of studies explored the effects of a variety of disruptive agents (e.g. drugs and electrical stimulation) on memory formation. This enterprise was almost entirely focused on what we have termed 'short-term consolidation'—that is, processes lasting seconds, minutes or hours. The use of animal models to study longer term consolidation processes is a relatively recent event. Table 2 lists studies investigating such long-term consolidation after experimentally induced lesions in parts of the medial temporal lobe. In all these cases, an attempt was made to determine whether damage caused at various intervals after learning had different effects, such as whether there was some sort of gradient of RA. Once again, the standard model of consolidation outlined above has been applied to this area, the assumption being that damage in the hippocampal formation would cause RA at some, but not all, intervals after learning. As in the work with humans, the data must be viewed with an eye towards the kind of learning that is being assessed.

Although the data are not entirely consistent, it is possible to draw some conclusions from these studies. When the task is one for which the hippocampus is absolutely essential during learning (e.g. spatial tasks, such as the water maze, the radial maze or scene discrimination), there is most often a flat RA. That is, no matter when the lesions are made in the hippocampus, deficits ensue. This is comparable to the result seen in humans for autobiographical memory (see Table 1), and suggests that in some domains at least, the hippocampus is involved at all stages of learning and remembering. When the task is one for which the hippocampus is not essential (e.g. object discrimination), the typical finding is either no RA at all or a graded RA. It is important to note that in the latter case and in those few cases in which more severe or even flat RA gradients are observed with object discriminations [41,42], there was extensive damage to structures beyond the hippocampus that might have played a critical role in object discrimination. Finally, there are reasons to suspect that the severity of RA and the likelihood of obtaining a flat gradient depend upon the extent of the lesion in the hippocampus. If the lesion is relatively complete, a flat RA will result, at least in those cases where tasks requiring the hippocampus are employed.

The data from animal studies are largely consistent with the conclusion we have already drawn from the human studies: the standard model of memory consolidation is vulnerable in the face of the evidence and needs either to be reformulated or replaced. The extent of RA and the slope of the RA gradient seem to depend on the size of the lesion to the hippocampal formation and adjacent, temporal lobe structures: the greater the lesion, the more extensive the RA, and the more shallow the slope.

**Inadequacy of traditional models**

The standard model and variants of it (e.g. [43,44]) cannot readily account for evidence of very extensive and often flat RA gradients, especially for autobiographical memory, nor for evidence of relatively preserved general semantics (vocabulary) acquired during periods for which subjects have RA. The retrieval model initially proposed by Sanders and Warrington [45] also cannot account adequately for the results because it assumes that the RA gradient of amnesic patients should always be parallel.
### Table 2

**Retrograde amnesia in animals with hippocampal complex damage.**

<table>
<thead>
<tr>
<th>Study</th>
<th>Animal</th>
<th>Lesion</th>
<th>Task</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Salmon, Zola-Morgan and Squire, 1985 [42]</td>
<td>Monkey</td>
<td>Hippocampus Amygdala</td>
<td>Object discrimination</td>
<td>Flat RA up to 8 months</td>
</tr>
<tr>
<td>Zola-Morgan and Squire, 1990 [41]</td>
<td>Monkey</td>
<td>Hippocampus Subiculum Posterior entorhinal cortex Parahippocampus</td>
<td>Object discrimination</td>
<td>Graded RA (deficit up to 4 weeks, no deficit beyond)</td>
</tr>
<tr>
<td>Winocur, 1990 [80]</td>
<td>Rat</td>
<td>Hippocampus</td>
<td>Food preference</td>
<td>Graded RA (deficit at 2 days, no deficit beyond)</td>
</tr>
<tr>
<td>Kim and Fanselow, 1992 [81]</td>
<td>Rat</td>
<td>Hippocampus Cortex</td>
<td>CER</td>
<td>Tone: no RA Context: graded RA (deficit at 1 day, no deficit at 28 days)</td>
</tr>
<tr>
<td>Cho, Beracochea and Jaffard, 1993 [82]</td>
<td>Mouse</td>
<td>Entorhinal cortex</td>
<td>Radial maze</td>
<td>Graded RA, but performance not normal, even after 6 weeks</td>
</tr>
<tr>
<td>Gaffan, 1993 [83]</td>
<td>Monkey</td>
<td>Fornix</td>
<td>Scene discrimination</td>
<td>Flat RA up to 6.5 months</td>
</tr>
<tr>
<td>Vnek and Rothblat, 1993*</td>
<td>Rat</td>
<td>Hippocampus</td>
<td>Object discrimination</td>
<td>Graded RA (deficit at 27 days, no deficit at 56 days)</td>
</tr>
<tr>
<td>Bolhuis, Stewart and Forrest, 1994 [84]</td>
<td>Rat</td>
<td>Hippocampus or subiculum</td>
<td>Water maze</td>
<td>Flat RA up to 14 weeks</td>
</tr>
<tr>
<td>Kim, Clark and Thompson, 1995 [85]</td>
<td>Rabbit</td>
<td>Hippocampus Cortex</td>
<td>Trace eyeblink</td>
<td>Graded RA (deficit at 1 day, no deficit at 1 month)</td>
</tr>
<tr>
<td>Cho, Kesner and Brodaire, 1995 [86**]</td>
<td>Rat</td>
<td>Hippocampus or entorhinal cortex</td>
<td>Radial maze</td>
<td>Graded RA, but performance not normal</td>
</tr>
<tr>
<td>Wiig, Cooper and Bear, 1996 [87]</td>
<td>Rat</td>
<td>Fornix or perirhinal cortex, or both</td>
<td>Visual discrimination</td>
<td>Graded RA in all groups (deficit up to 3 weeks)</td>
</tr>
<tr>
<td>Weisend, Astur and Sutherland, 1996*</td>
<td>Rat</td>
<td>Hippocampus</td>
<td>Water maze, CER and operant</td>
<td>Flat gradient in all cases up to 36 weeks</td>
</tr>
<tr>
<td>Koerner, Thomas, Weisend and Sutherland, 1994†</td>
<td>Rat</td>
<td>Hippocampus</td>
<td>Water maze</td>
<td>Flat RA up to 10 weeks</td>
</tr>
<tr>
<td>Bobbot, Liu, Thurm, Nadel and Bures, 1996§</td>
<td>Rat</td>
<td>Hippocampus</td>
<td>Water maze</td>
<td>Flat RA up to 20 weeks</td>
</tr>
<tr>
<td>Francis, Glenn and Mumby, 1996§</td>
<td>Rat</td>
<td>Perirhinal cortex or hippocampal formation</td>
<td>Water maze and object discrimination</td>
<td>Hippocampal lesions: flat RA (up to 72 hours) on the water maze Perirhinal lesions: transient RA on object discrimination</td>
</tr>
<tr>
<td>Anagnostaras, Maren and Fanselow, 1996**</td>
<td>Rat</td>
<td>Hippocampus</td>
<td>CER (signalled or unsignalled)</td>
<td>Unsignalled CER: flat RA up to 40 days Signalled CER: graded RA, deficit at short delay, no deficit at 50 days</td>
</tr>
</tbody>
</table>


...to that of controls; however, it is often not the case, especially for events and personal semantics. The most parsimonious account of the data would be to assume that the hippocampal complex and neocortex continue to be involved in both the storage and the retrieval of episodic memory traces throughout life. General semantics, on
the other hand, are learned and represented outside the hippocampal complex (or medial temporal lobes) [37,46].

Below, we sketch a model that provides a parsimonious account of the evidence without resorting to the notion that the hippocampal formation and adjacent medial temporal lobe structures play a very long, but still temporary, role in all forms of episodic and semantic memory.

The hippocampus and episodic memory: the multiple trace theory

We propose that the formation, maintenance and recovery of episodic memory involves the nine processes listed below. The first four points are shared with the standard model, whereas the others distinguish our model from it.

1. The hippocampal complex rapidly (and obligatorily) encodes all information that is attended or consciously apprehended [4*,7*,47]. This process involves what we have called short-term consolidation, or cohesion.

2. This information is sparsely encoded in a distributed ensemble of hippocampal complex neurons [48-50].

3. This ensemble acts as a pointer, or index, to those neocortical (or other) neurons that represent the attended information [4*,50-55] and serves as a mechanism to bind them into a coherent memory trace.

4. The entire hippocampal-neocortical ensemble constitutes the memory trace for the episode. The feature information contained in neocortical circuits is bound into an episode by the spatial context provided from the hippocampus.

5. Each re-activation of this memory trace occurs in an altered neuronal and experiential context.

6. Because the hippocampal complex obligatorily encodes all information that is attended, the re-activation of a memory trace results in the creation of a newly encoded hippocampal trace, which, again, is sparse and distributed.

7. By virtue of indexing a similar set of neocortical neurons encoding the features of the episode, each such trace shares some or all of the information about the initial episode.

8. The creation of multiple, related traces facilitates the 'extraction' of factual information from an episode and its integration with pre-existing semantic memory stores. Facts about the world (e.g. the Eiffel Tower is in Paris, apples are round, mangos are a type of fruit, etc.) that are acquired in the context of a specific episode are separated from that episode and ultimately stored independently of it. This process could be viewed as one of the consequences of 'consolidation'.

9. The spatial and temporal contextual information that conveys the episodic quality to memory, however, depends upon the continuing involvement of the hippocampal complex (for spatial context) and the frontal cortex (for temporal context) [56,57]. For example, remembering where and when you first tasted a mango always demands the participation of these systems.

Note that in this model, there is no assumption of a long-term consolidation process slowly strengthening neocortical neurons, resulting in the transfer of information initially stored in the hippocampal complex to the neocortex. Instead, it is assumed that the hippocampal complex and neocortex interact over time, in ways that can influence the content of knowledge in both systems [43], but that autobiographical memory always depends upon the hippocampal complex and its provision of contextual information. (For earlier, similar views, see [21,22,39,58-60].) As episodic memories age, they would either be forgotten or would have benefitted from the formation of multiple traces in the hippocampal complex and neocortex. Thus, older episodic memories would be associated with a greater number of traces. Retrieval would become easier as the number of traces, and the number of access routes to them, proliferate. Because memory traces are sparse and distributed, even minimal damage at any location in the hippocampal formation could affect acquisition, retention, and recovery of any single one of them. Newly acquired traces would be particularly vulnerable, but older memories, which are multiply represented, would be able to withstand the loss of more hippocampal complex tissue.

As a result, partial damage to the hippocampal complex would affect remote episodic memory in relation to its age—that is, one will observe a graded RA. According to our model, the extent of RA and the shape of the gradient would depend upon the size of the hippocampal lesion. Complete damage to the hippocampal complex should lead to a flat gradient for autobiographical episodic information that extends throughout life. Differences in the extent and temporal gradient of RA among the different types of memory would be determined by the complexity or richness of the trace that is to be recovered. Because the full details of an autobiographical episode are not likely to be re-activated often, these details would be the most vulnerable to disruption. On the other hand, the gist of an episode, partial information about it, or facts about one’s life—those things that comprise personal semantic memory—are more likely to be re-activated and hence would be multiply represented.

The same would be true of semantic memory for public events and personalities as for personal semantic memories. Moreover, such memories need not necessarily
incorporate information about the place and time in which they were acquired. As a result, RA for both types of semantic memory, personal and public, would be typically less extensive and with a steeper gradient than for autobiographical memory, unless extremely detailed information about semantic memory were required [27].

On a related issue, concepts and words appear to be stored in circuits in the posterior and lateral temporal cortex, and if these are spared, RA for these forms of information is not observed. Damage primarily to that system, as in ‘semantic dementia’, leads to conceptual/word loss with relative sparing of episodic memory and of words dependent on that has also been noted in normal aging (M Moscovitch, R Ladowsky, abstract 318, Psychonomic Soc Bull 1988, 26:521; see also [61]).

Thus, our model accounts for differences in RA among different types of memory and explains reports of graded RA lasting for over 25 years, or longer, without recourse to extremely long-lasting (and unlikely) physiological mechanisms and without recourse to the inherently unsatisfactory notion that the hippocampal complex is a 25-year or longer ‘temporary’ memory system. In short, according to our model, long-term episodic memory does not result from continuous, gradual strengthening (either endogenously or by repeated external stimulation) of neural connections in the neocortex that were formed initially (but very weakly) at encoding. The decreased vulnerability observed as memories age owes to their multiple representation. It does not owe to a consolidation-induced shift from a reliance on both the hippocampal complex and the neocortex to exclusive reliance on the neocortical system.

Effects of lesions to different regions of the hippocampal complex and to structures outside the medial temporal lobes

Little is known with certainty about the functions of the various subregions of the hippocampal formation and the surrounding structures in the medial temporal lobe. Each subregion probably has its own function, either with respect to the type of memory that it mediates or with respect to the type of memory process it subserves [62*]. For example, there is growing evidence that memory for dynamic spatial information requires the hippocampal formation, that memory for objects depends on the peri-rhinal cortex, and that memory for the conjunction of objects and locations is mediated by the parahippocampal gyrus (see review by Tulving and Markowitsch, in this issue, pp 209–216). If recovery of remote memories always depends on re-activation of the hippocampal–neocortical complex that constitutes the memory trace, it follows that damage to any part of that complex will lead to a remote memory loss. Recent reports of focal RA are consistent with our model [8,9,10**,11,63,64*]. Even when damage largely spares the medial temporal lobes, remote memory loss will occur if lesions affect neocortical structures that represent features of the trace (see point 5 of our model) or that serve as transmission routes or convergence zones [65] between the neocortex and the medial temporal lobes (see [8,10**,11,63,64*65,66*] for further discussions regarding this idea). Consistent with our model, a recent PET study has shown that medial temporal lobe structures are activated during recovery of both recent and remote memories [67*].

Frontal lobe damage or dysfunction [25,68,69] or lesions to regions and pathways connecting the prefrontal cortex to the medial temporal lobes [33,59,63] also lead to remote memory loss that can be as severe as that which occurs following medial temporal lobe damage. The loss, however, does not occur because the memory trace or aspects of it are degraded or lost; rather, RA is observed because strategic retrieval processes, mediated by the frontal lobes and necessary for gaining access to the trace, are either malfunctioning or inoperative [4*,27,70*,71]. Indeed, because one can envisage situations in which recovery of remote memories requires more strategic retrieval than recovery of recent memories (i.e. what one did yesterday versus last summer versus the summer of one’s 14th year), it may be possible to obtain a temporal gradient opposite to that observed following medial temporal lobe lesions.

This prediction has recently been supported by a study of subjects with semantic dementia, typically assumed to involve fronto-temporal damage [72]. In contrast to patients with dementia of the Alzheimer’s type, who have extensive medial temporal lobe damage, these subjects showed a reversed temporal gradient such that recent events were remembered better than remote ones.

Conclusions

It is important to distinguish between the core of the model and the particular way we have chosen to implement it. The core notion is that episodic memory traces consist of linked ensembles of the hippocampal complex and neocortical neurons and that the hippocampal component remains a necessary part of that ensemble as long as the memory traces are viable.

To account for the relative sparing of remote memories (temporally graded RA) that sometimes, but not always, follows medial temporal lobe damage, we have proposed a multiple trace theory of memory, primarily to show that there are alternatives to the standard model, which is challenged by available data.

There may be other ways to account for the existing data without assuming multiple traces: for example, it might be the case that new nodes or retrieval routes are added to the initial memory trace every time it is...
re-activated. Whichever of these alternatives proves to be correct, it is our hope that by advancing them, we will encourage investigators to re-examine the notion of long-term consolidation and to consider alternatives that are more biologically and psychologically plausible.

Acknowledgements

This review was prepared while M. Moscovitch was a visiting professor at the University of Arizona and was supported by the Flinn Foundation Cognitive Neuroscience Program. Our research is supported by the McDonnell Foundation (grant JS-8154 to I. Nadel) and the Natural Science and Engineering Research Council of Canada (Grant A8347 to M Moscovitch). We thank Mary Peterson, Elizabeth Warrington, Rob Naderlander, Peter Milner, Larry Squire and Gordon Winocur for their help, comments and criticisms. The authors contributed equally to the manuscript; the order of authors is arbitrary.

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

• of special interest
•• of outstanding interest


This paper provides a succinct account of a neuropsychological model of memory that the author develops in more detail in some of the other papers cited [3,14,47]. The essence of the model, which provides some of the basis of the present review, is that the hippocampus and related structures bind into a memory trace those neural elements that mediated the conscious experience of an event. 'Consciousness' is an integral part of the memory trace. The frontal lobes act as 'working-memory' structures that contribute to encoding and that mediate strategic retrieval processes that are needed to locate and verify the memory traces recovered from the hippocampal system and that place them in the proper temporal and thematic context.


A succinct presentation of the traditional model of memory consolidation and a computational connectionist model of it.


This author (see also [9]) was among the first to recognize the importance of identifying and studying patients with focal RA. In this paper, the author presents some new data, summarizes his findings and offers a model to account for focal RA.


This paper provides corroborating evidence from a group of non-Korsakoff amnesic patients of a finding first reported by Warrington and McCarthy [17] in a single-case study. The finding is that premorbidly acquired vocabulary is relatively preserved in these patients, even if it was acquired during the period for which the patient has RA. This is not true of Korsakoff amnesic patients, who show a temporal gradient for those words.


Reports the postmortem examination of the brains of three patients whose AA and RA had been documented extensively. All three patients had lesions confined almost exclusively to the hippocampal formation. Although the authors do not stress this point, two of the three patients had a history of alcoholism and one of them had epileptic seizures accompanied by blackouts.


The reported cases of transient global amnesia during and after the attack. Behav Neurol 1995, 8:93-101.


This paper contains a case report and a set of annotated references for all the reported cases of transient global amnesia.


Critique of the view that all portions of the medial temporal lobe play the same role in declarative memory.


A thorough review of the literature on RA and a good introduction to the author's neuroanatomically based theory of memory.


Demonstration that even when animals can perform in spatial tasks after a long learning–surgery interval, there are critical differences between normal and lesioned animals. The results suggest that performance is not based on a shift of learned information from hippocampus to neocortex as much as on a shift of behavioral strategy.

87. Wiig KA, Cooper LN, Bear MF: Temporally graded retrograde amnesia following separate and combined lesions of the perirhinal cortex and fornix in the rat. Learn Mem 1996, 3:313–325.