MEMORY, 1997, 5 (1/2), 179–212

Extending Models of Hippocampal Function in Animal Conditioning to Human Amnesia

Mark A. Gluck and Brandon R. Ermita
Rutgers University, USA

Lindsay M. Oliver
Glasgow Caledonian University, UK

Catherine E. Myers
Rutgers University, USA

Although most analyses of amnesia have focused on the loss of explicit declarative and episodic memories following hippocampal-region damage, considerable insights into amnesia can also be realised by studying hippocampal function in simple procedural, or habit-based, associative learning tasks. Although many simple forms of associative learning are unimpaired by hippocampal damage, more complex tasks which require sensitivity to unreinforced stimuli, configurations of multiple stimuli, or contextual information are impaired by hippocampal damage. In several recent papers we have developed a computational theory of hippocampal function which argues that this brain region plays a critical role in the formation of new stimulus representations during learning (Gluck & Myers, 1993, 1995; Myers & Gluck, 1996; Myers, Gluck, & Granger, 1995). We have applied this theory to a broad range of empirical data from studies of classical conditioning in both intact and hippocampal-lesioned animals, and the model correctly accounts for these data. The classical conditioning paradigm can be adapted for use in humans, and similar results for acquisition are obtained in both normal and hippocampal-damaged humans. More recently, we have begun to address an important set of category learning studies in both normals and hippocampal-damaged amnesics. This work integrates experimental studies of amnesic category learning (Knowlton, Squire, & Gluck, 1994) with theoretical accounts of associative learning, and builds on previously established behavioural correspondences between animal conditioning and human category learning (Gluck & Bower, 1988a). Our work to

Requests for reprints should be sent to Mark A. Gluck, Center for Molecular and Behavioral Neuroscience, Rutgers University, 197 University Avenue, Newark, New Jersey 07102, USA. Email: gluck@pavlov.rutgers.edu or see web page www.archtype.com/gluck

This research was supported by an NIMH National Research Service Award 1-F32-MH10351-01 (CM), by a McDonnell-Pew Foundation for Cognitive Neuroscience Grant-in-Aid (MC & CM), and by the Office of Naval Research through the Young Investigator program (MG) and grant N00014-88-K-0112 (MG).

© 1997 Psychology Press, an imprint of Erlbaum (UK) Taylor & Francis Ltd

date illustrates some initial progress towards a more integrative understanding of hippocampal function in both animal and human learning, which may be useful in guiding further empirical and theoretical research in human memory and amnesia.

1. INTRODUCTION

In several recent papers, we have argued that the hippocampal region plays an essential role in the formation of novel stimulus representations in the formation of new associations and memories (Gluck & Myers, 1993, 1995; Myers & Gluck, 1996; Myers et al., 1995a). As shown in Fig. 1.1, the hippocampal region comprises a group of structures located deep within the brain, and includes the hippocampus itself as well as the nearby dentate gyrus, subiculum, and entorhinal cortex. The outermost of these structures, the entorhinal cortex, receives highly processed information from the entire spectrum of sensory modalities, as well as from multimodal cortical association areas. Information flows in a roughly unidirectional fashion from the entorhinal cortex to the dentate gyrus, to hippocampus, to the subiculum, and back to the entorhinal

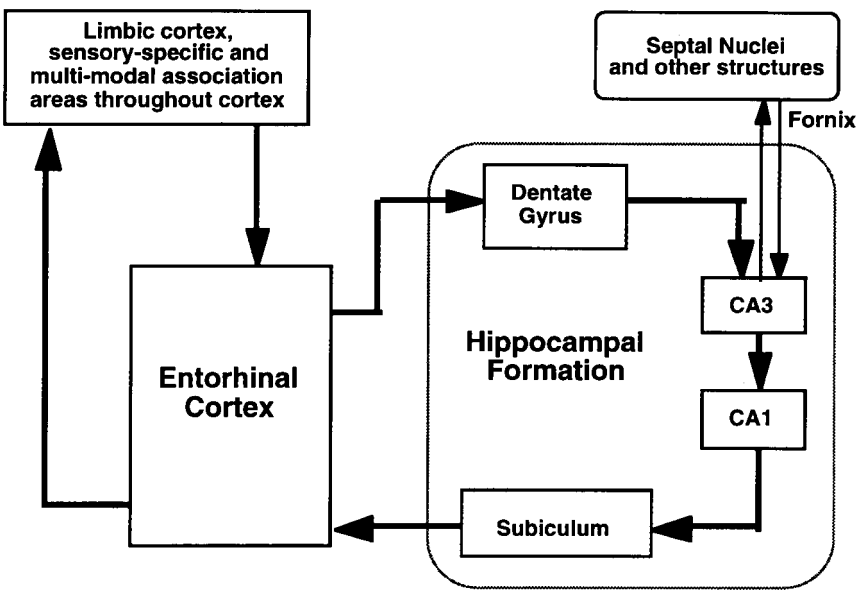

FIG. 1.1. Schematic of major information flow pathways in the hippocampal region. Highly processed, multimodal inputs enter entorhinal cortex, and proceed in a largely unidirectional pathway through the hippocampal formation (dentate gyrus, hippocampal fields CA3 and CA1 and subiculum) before returning to entorhinal cortex and thence back to the same cortical areas where they arose. There is also a bi-directional pathway through the fornix connecting the hippocampus with subcortical areas such as the septal nuclei. Many other connections exist in addition to the major ones shown here.

cortex before returning to the same sensory areas where it originally arose. In addition to this basic pathway, there are a large number of direct connections between the structures of the region. The hippocampus has another input and output pathway through the fornix, a fibre bundle connecting it with subcortical structures (e.g. basal forebrain) which modulate hippocampal functioning.

Damage to the hippocampal region in humans produces a characteristic anterograde amnesia syndrome, strongly impairing the learning of new information (Squire, 1987). Human hippocampal damage can result from a variety of etiologies, ranging from ischemia, viral encephalitis, and aneurysm/embolism to the arteries that vascularise the hippocampal region. Damage to other related structures, such as the basal forebrain, can also result in an amnesic syndrome which shares features with hippocampal amnesia, presumably because such damage indirectly interferes with normal hippocampal-region processing (see Myers et al., 1996).

The anterograde amnesia that follows human hippocampal-region damage is generally characterised by an inability to acquire new episodic or declarative information, the kind of information about individual events and experiences that is accessible to conscious control. These patients may also show some degree of retrograde amnesia, or disruption of previously acquired information, but this is usually limited to information acquired shortly before the trauma, and tends to lessen in a time-graded fashion for older information (Ribot, 1982; Squire, 1987).

In contrast to episodic or declarative memories that are often acquired in a single exposure, other kinds of memory are incrementally acquired over many exposures. These types of memory are not necessarily accessible to conscious recollection, and often involve learning skills or procedures rather than facts (Squire, 1987). This dissociation between episodic memory and procedural memory is often summarised as a difference between ''knowing that'' and ''knowing how''. Many simple forms of procedural memory survive hippocampal-region damage relatively intact. The animal learning literature is full of studies showing how animals with hippocampal-region damage can show normal acquisition of a variety of tasks such as the acquisition of classically conditioned responses to a single stimulus (Akase, Alkon, & Disterhoft, 1989; Schmajuk, 1994; Solomon, 1977; Solomon & Moore, 1975), the ability to choose the novel of a pair of objects where one object was seen immediately before (Zola-Morgan & Squire, 1992) simple discriminations of singly presented odour stimuli in an operant task (Eichenbaum, Fagan, Mathews, & Cohen, 1988), and learning to navigate to an escape platform when started from a constant location in a pool (Eichenbaum, Otto, Wible, & Piper, 1991). Similarly, human hippocampal-damaged amnesics are not impaired at learning a conditioned motor-reflex response (Daum, Channon, & Canavan, 1989; Gabrieli et al., 1995; Woodruff-Pak, 1993), learning simple classification tasks (Knowlton et al., 1994), or learning new motor skills such as mirror drawing (Cohen, 1984). All of these

spared tasks can be solved by incremental formation of habits or tendencies, without requiring episodic memories of any individual learning trial.

There are, however, other tasks that seem superficially to be just as procedural or implicit as those noted here, but which are impaired after hippocampal-region damage. For example, although the acquisition of a classically conditioned response to a single CS is not impaired by hippocampal-region damage, there may be severe impairments in classical-conditioning tasks that require learning about unreinforced stimuli (Kaye & Pearce, 1987; Solomon & Moore, 1975), configurations or combinations of stimuli (Sutherland & Rudy, 1989), contextual information (Hirsh, 1974), or relationships that span short delays (Eichenbaum, Otto, & Cohen, 1994; Moyer, Deyo, & Disterhoft, 1990; Port, Romano & Patterson, 1986; Rawlins, 1985; Zola-Morgan & Squire, 1992). These findings imply that the hippocampal-region does indeed participate in information processing during apparently procedural tasks, although this participation may not necessarily be evident from studying whether animals or people can—or cannot—acquire a simple associative response.

In Gluck and Myers (1993) we presented a computational model of hippocampal-region function in associative learning, which suggests that the hippocampal region is involved in the formation of new stimulus representations during normal procedural learning. Learning that depends on new representations is expected to be hippocampal-dependent, whereas learning that can make use of pre-existing representations may be hippocampal-independent. Section 2 of this paper presents this model. We have applied this model to a broad range of empirical data from studies of classical conditioning in intact and hippocampal-lesioned animals, as discussed in Section 3, and the model correctly accounts for these data. Additionally, the classical conditioning paradigm can be adapted for use in humans, and similar results for acquisition training appear to be obtained in both normal and hippocampal-damaged amnesic humans, much as in animals during conditioning. Finally, in Section 4, we note that the model can be easily extended to address an example of probabilistic category learning in both normals and hippocampal-damaged amnesics.

2. STIMULUS REPRESENTATION AND HIPPOCAMPAL FUNCTION

Many previous characterisations of hippocampal-region function in animal learning have been proposed which are task-centred: noting that a particular class of task is hippocampal-dependent, and then seeking to characterise the hippocampus as a processor that implements that kind of function. These task-centred descriptions have implicated the hippocampal region in spatial learning (O'Keefe & Nadel, 1978), contextual processing (Hirsh, 1974; Nadel & Willner, 1980), configural learning (Sutherland & Rudy, 1989), and more. The

hippocampal region also appears to be involved in temporal processing such as sequence learning (Buzsaki, 1989), response timing (Akase et al., 1989; Moyer et al., 1990) and intermediate-term memory (Eichenbaum et al., 1994; Olton, 1983; Rawlins, 1985; Zola-Morgan, Squire, & Amaral, 1989).

Instead of a task-specific characterisation of hippocampal function, it may be possible to try and describe an information-processing role for the hippocampal region, and, from this, derive a wider range of task-specific deficits which arise from interfering with this information-processing function. For example, it has been proposed that the hippocampal region is involved in the flexible use of learned information in novel situations (Eichenbaum & Buckingham, 1990; Eichenbaum, Cohen, Otto, & Wible, 1992); attentional control through a process of inhibiting responses to irrelevant cues (Douglas & Pribram, 1966; Schmajuk & Moore, 1985); and providing ''contextual tags'' to learned information (Penick & Solomon, 1991; Winocur, Rawlins, & Gray, 1987).

In our work, we have focused on a putative role for the hippocampal region in forming new stimulus representations to assist learning. This suggestion builds on many of the earlier information-processing characterisations of hippocampal-region function listed previously, and also on the psychological idea of stimulus representations (Anderson, 1977; James, 1896; Shepard, 1958). The emphasis on representation in psychological modelling has been particularly salient in recent years as simple rule-based models of cognitive function have given way to connectionist models that emphasise how associative networks re-represent stimuli within a multi-layer network (Hanson & Burr, 1990; Rumelhart, Hinton, & Williams, 1986). These connectionist networks are generally assumed to implement abstract characterisations of underlying neural processes, providing a framework for expressing theories about both animal learning (Barto & Sutton, 1982; Kehoe, 1988; Schmajuk & DiCarlo, 1990) and human learning and memory (Gluck & Bower, 1988a; Kruschke, 1992; Shanks & Gluck, 1994; McClelland, McNaughton, & O'Reilly, 1994).

In the remainder of this section, we will briefly review the idea of stimulus representation within the framework of connectionist models, and then present our own computational model of hippocampal-region function in learning.

Connectionist Models of Representation and Learning

Within a typical connectionist network (Fig. 2.1), the internal stimulus representation is characterised as a pattern of activations over a set of ''hidden nodes'' which recode a stimulus input. These hidden nodes represent the network's internal representation of the stimulus pattern, and are roughly analogous to what psychologists refer to as a psychological representation (or the ''psychological space''), as distinct from a sensory representation which primarily reflects the physical qualities of the stimulus input (Shepard, 1958).

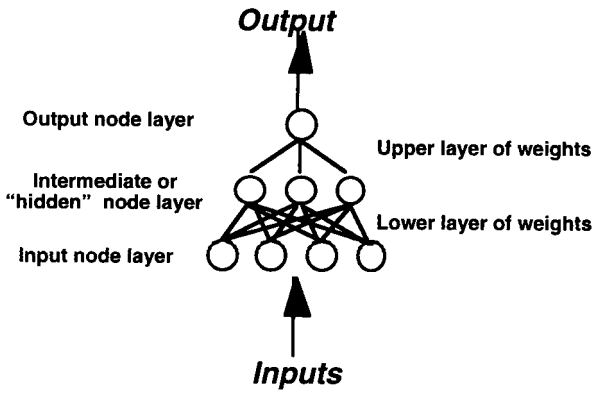

FIG. 2.1. A multi-layer network, with a ''hidden'' or intermediate layer of cells between the input and output cell layer, and weighted connections between layers. The hidden layer activations represent a re-coding or re-representation of the inputs. The output layer activations are interpreted as the behavioural response. All connection weights are adaptive.

The hidden layer activations depend on weighted connections between the input and hidden layers, and thus changes to this lower layer of weights are equivalent to changing the stimulus representation. The hidden layer activations feed through weighted connections to an output layer, and the output layer activation determines the behavioural response of the system. The network is trained by adapting the weights in both layers until the system generates the desired output in response to a given input.

The connectionist network in Fig. 2.1 can be trained to give the desired output to each of a set of different inputs. How easily (quickly) this is done depends in part on the hidden layer representations. For example, suppose that a particular input pattern generates the hidden layer activations schematised in the top of Fig. 2.2A. In order to learn the correct response to the first pattern, a set of weights will evolve mapping that hidden layer representation to the correct output activations. Now suppose a second input pattern generates the hidden layer activations schematised at the bottom of Fig. 2.2A. Note that the activation pattern at the top of Fig. 2.2A has considerable overlap (similarity) with the pattern at the bottom of Fig. 2.2A. This implies that the network in Fig. 2.1 will tend to generalise from one pattern to the other or, in other words, it will generate output to the second pattern that is similar to what is learned for the first pattern. If the desired outputs for the two patterns are the same, then this generalisation will be helpful; if the desired outputs for the two patterns are different, this generalisation will cause unwanted interference and hinder the mapping of these similar representations to different outputs. In contrast, two input patterns that evoke very different hidden layer activations (as shown in Fig. 2.2B) will engender far less generalisation, resulting in more facilitation for learning different outputs to these two patterns, but less facilitation for learning the same output to both.

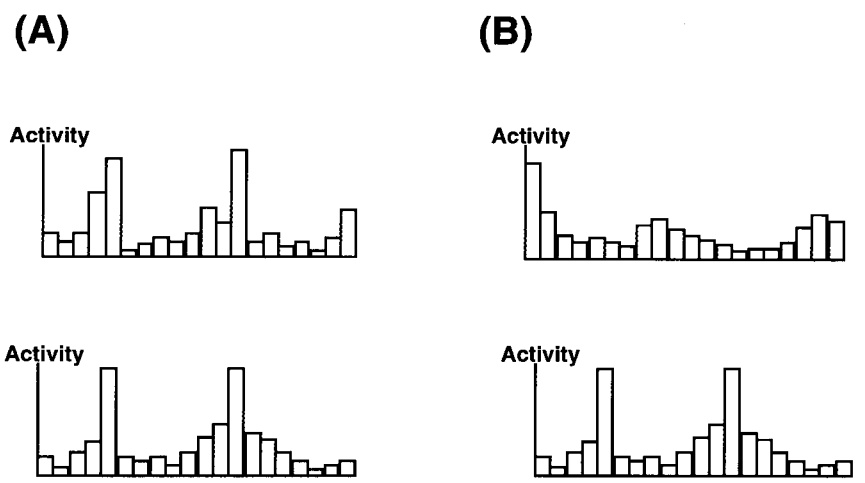

FIG. 2.2. Schematic of activations evoked in the hidden layer nodes of a connectionist network such as the one shown in Fig. 2.1, by presentation of four different input patterns. These activity patterns constitute stimulus representations in the network. (A) If the hidden-layer representations evoked by two input patterns are very similar, it will be difficult to train the network to respond differently to each. (B) If the representations evoked are very dissimilar, it will be easier to map each to a different response. Conversely, generalisation between inputs is aided by similarity in representation and made difficult by differentiated representations.

In this way, learning the correct responses to a set of input patterns can be affected greatly by changes in stimulus representations. The exact details of these representations are, in general, less important than the varying degree of similarity or overlap between representations of different input patterns.

The Cortico-hippocampal Model

The core idea of Gluck and Myers' (1993) theory of cortico-hippocampal function is that the hippocampal region facilitates learning by constructing internal stimulus representations that are biased in two ways. The first bias, *redundancy compression*, is a tendency to make more similar the representations of stimuli that co-occur or are to be associated with similar outputs; this bias ensures high generalisation between these stimuli. The second bias, *predictive differentiation*, is a tendency to differentiate, or make less similar, the representations of stimuli that are to be mapped to different outputs; this bias decreases generalisation between such stimuli. These two biases to compress and differentiate are partially opposing, and may often interact in complex ways, depending on stimulus–stimulus relationships within the training environment.

Hippocampal Processing. How can this postulated representational role for hippocampal-region function be implemented? Gluck and Myers (1993) show

how a simple connectionist architecture gives rise to just these sorts of representations. They model hippocampal-region processing as a *predictive autoencoder* (Hinton, 1989), as shown in Fig. 2.3. A predictive autoencoder maps input activations through a hidden layer to two classes of outputs: a reconstruction of the input pattern as well as a prediction for classifying the input pattern. Such a network may be trained via the error backpropagation learning algorithm (Parker, 1985; Rumelhart et al., 1986; Werbos, 1974); this algorithm allows development of representations across the internal layer of nodes that compress redundancies while emphasising predictive information—exactly as required by our theory. The details of the learning procedure are not particularly critical in that many alternative learning procedures are equally able to accomplish the required representational changes (see Myers et al., 1995a, for further discussion of this point).

Intact Cortico-hippocampal Processing. This hippocampal-region network is assumed to interact with other cortical regions which are the sites of long-term storage. One such network is shown in the intact model of Fig. 2.4A. The cortical network on the left in Fig. 2.4A represents a highly simplified model of some aspects of long-term memory in cerebral and cerebellar cortices. This network takes the stimulus input, and maps it through weighted connections to a hidden node layer and then to an output node; the activation of this output node is interpreted as the system's behavioural response to the input. This cortical network adapts its weighted connection strengths according to the LMS learning algorithm (Widrow & Hoff, 1960), which is related to psychological descriptions

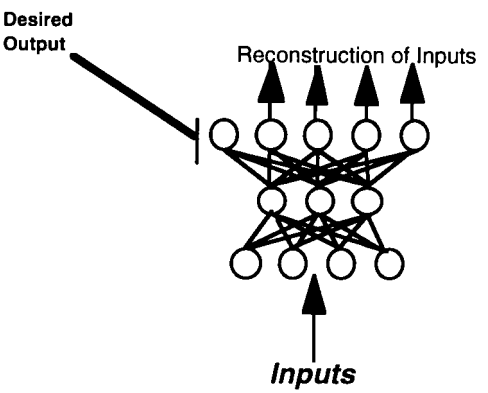

FIG. 2.3. A predictive autoencoder network, like that assumed in the cortico-hippocampal model to represent hippocampal-region processing. This network maps inputs through a hidden node layer to outputs that learn to reconstruct the input pattern as well as classifying the input. To accomplish this, representations are developed in the hidden node layer which compress redundancies in the input while maintaining and differentiating enough predictive information to allow generation of the desired outputs.

of learning (Gluck & Bower, 1988a; Rescorla & Wagner, 1972; Sutton & Barto, 1981) and also to biological mechanisms of plasticity such as LTP (Bliss & Lomo, 1973; Levy, Brassel & Moore, 1983; Stanton & Sejnowski, 1989).

The hippocampal-region autoencoder, and a simpler multi-layer network representing a cortical module, are combined into the intact model shown in Fig. 2.4A. Within this two-part model of intact cortico-hippocampal processing, new representations are developed first in the hippocampal region and then incrementally transferred, throughout the learning process, to the internal-layer nodes of cortical (or cerebellar) networks. The LMS rule calculates the error at a node as the difference between the desired and actual output at that node, and distributes this error among the weights feeding into the node. For the output layer node in the cortical network, the desired output is simply a prediction of the US arrival, and, thus, the LMS rule can be used to train the upper layer of weights in the cortical network. However, the desired output is not externally defined for the internal layer nodes of the cortical network.

However, the activations of the internal layer of the hippocampal-region network can be used as the desired activations for corresponding internal layer nodes in the cortical network. Thus, a second application of the LMS rule can be used to train the lower layer of cortical weights. Over time, the internal layer of the cortical network will come to mirror the representations developed in the hippocampal network—even though the cortical network alone, using the LMS algorithm, would be incapable of devising these representations. We note that it is not necessary for there to be a one-to-one mapping between the internal-layer nodes of the hippocampal-region and cortical networks; the cortical network

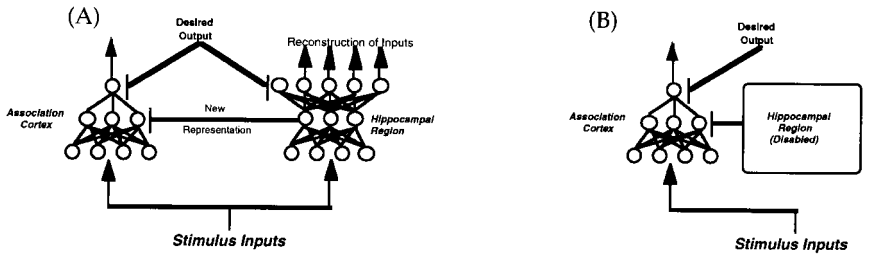

FIG. 2.4. The cortico-hippocampal model (Gluck & Myers, 1993). (A) The intact system is assumed to include a predictive autoencoder, representing hippocampal-region processing, that constructs new stimulus representations in its hidden layer which are biased to compress redundancies while differentiating predictive information. These stimulus representations are acquired by long-term storage sites in the cortex, represented as a multi-layer network which learns to predict US arrival. The cortical network uses the Rescorla-Wagner rule to map from inputs to the hippocampal-mediated internal representations, and again to map from the internal layer to output activations. (B) Hippocampal-region lesion is assumed to disable the hippocampal network, in which cases the cortical network can no longer acquire new internal representations, but can acquire new behavioural responses based on its pre-existing (and now fixed) internal representations.

nodes can be mapped to a fixed, random linear recombination of the hippocampal network activations, and this will maintain most of the relevant information, especially if the cortical hidden layer is larger than the hippocampal hidden layer.

Learning in the hippocampal-region network, and in both layers of the long-term memory cortical network, is assumed to proceed incrementally and in parallel. At this time, we have not made any attempt to clarify the relative speeds of learning in the two networks, nor the time course of transfer of information from one network to the other; however, this general two-component architecture for learning would certainly allow for a consolidation period during which information is slowly transferred from hippocampal-region to long-term memory (see Alvarez & Squire, 1994; McClelland et al., 1994; Murre, 1994; Squire, 1987; Winocur, 1990).

Although it is important to understand the aggregate functional role of the hippocampal region, ultimately we would like to know how the proposed behavioural processes map onto more detailed anatomical structures. Of particular importance is the parahippocampal region, comprising entorhinal, perirhinal, and parahippocampal cortices; these structures are the primary sites through which sensory information both enters and leaves the hippocampal formation. Later, in the Discussion section, we will briefly note more recent work by Myers, Gluck, and Granger (1995) which suggests how Gluck and Myers' (1993) aggregate hippocampal region might be distributed among hippocampal and parahippocampal regions. Other more recent work of ours (Myers, Ermita, Harris, Hasselmo, Solomon, & Gluck, 1996) extends the model to incorporate subcortical inputs, specifically the cholinergic inputs from the septum. For the moment, however, this aggregate characterisation of hippocampal-region function will suffice for describing the relevant mappings from animal learning to human amnesia.

Hippocampal-lesioned Processing. The lesioned model shown in Fig. 2.4B represents the processing that remains after total removal of all hippocampal-region mediated processes. In this case, the cortical modules are assumed to be unable to acquire new representations, although they can still learn new behavioural responses based on their pre-existing (and now fixed) stimulus representations. Our lesioned model therefore predicts that hippocampal-region damage will be most deleterious to those tasks that require new stimulus representations, but less evident in those tasks for which pre-existing or random stimulus representations suffice.

As described in the next section, the Gluck and Myers' model of Fig. 2.4 can be applied to classical conditioning, and captures many of the trial-level aspects of the behaviour of intact and hippocampal-lesioned animals. Later, in Section 4, we will illustrate how the model can also be applied to some aspects of normal and amnesic human learning.

3. APPLICATION OF MODEL TO CLASSICAL CONDITIONING

Classical Conditioning

Classical Pavlovian conditioning is one of the simplest forms of associative learning: a previously neutral cue (the conditioned stimulus or CS) is repeatedly paired with a response-evoking cue (the unconditioned stimulus or US) and comes itself to evoke an anticipatory response (the conditioned response or CR). Since Pavlov's seminal studies, almost 80 years of detailed behavioural studies and theoretical analyses of classical conditioning have been accumulated. There now exist several detailed mathematical and computational models for classical conditioning which illustrate how a wide range of conditioning behaviours could emerge from a small set of underlying processing and representational assumptions (Mackintosh, 1983; Pearce & Hall, 1980; Rescorla & Wagner, 1972; Sutherland & Mackintosh, 1971). No form of animal or human learning has been better characterised, or more successfully modelled, than classical conditioning. These behavioural models provide an important starting point for understanding and characterising what is missing or different in hippocampal lesioned animals.

Pavlovian conditioning techniques have been applied to several natural reflexes including salivation, knee-jerking, galvanic skin response, and heart rate conditioning. The most popular and widely adopted preparation has been the eyeblink reflex generated in response to an aversive corneal airpuff or paraorbital shock (Fig. 3.1A). If the airpuff or shock US is reliably preceded by a tone or light CS, subjects learn to generate a protective eyeblink to the CS, and time their blinks so that the eye is maximally closed at the time of expected US arrival (Fig. 3.1B; Gormezano, Kehoe, & Marshall, 1983). Eyeblink conditioning has been studied most extensively in rabbits, but also in frogs, rats, cats, and dogs. The neural circuits underlying the mammalian eyeblink response have been delineated in greatest detail in the rabbit (Thompson, 1986, 1990), suggesting that the long-term storage of CS–US associations occurs in the cerebellar cortex and underlying nuclei. Cerebellar lesion permanently abolishes new conditioned eyeblink learning in animals (Clark, McCormick, Lavond, & Thompson, 1984; Glickstein, Hardiman, & Yeo, 1983).

Hippocampal lesion, however, does not impair the acquisition of simple delay-conditioning in animals (Akase et al., 1989; Berger & Orr, 1983; Port & Patterson, 1984; Schmaltz & Theios, 1972; Solomon & Moore, 1975). However, during learning of the conditioned response, hippocampal activity does change, as indicated through neurophysiological recordings (Berger, Rinaldi, Weisz & Thompson, 1983; Disterhoft, Coulter & Alkon, 1988) and human positron emission tomography (PET) scans (Blaxton et al., 1996; Molchan et al., 1994). This implies that, although the hippocampus may not be strictly required for conditioned learning, it is normally active during such learning. Furthermore, in

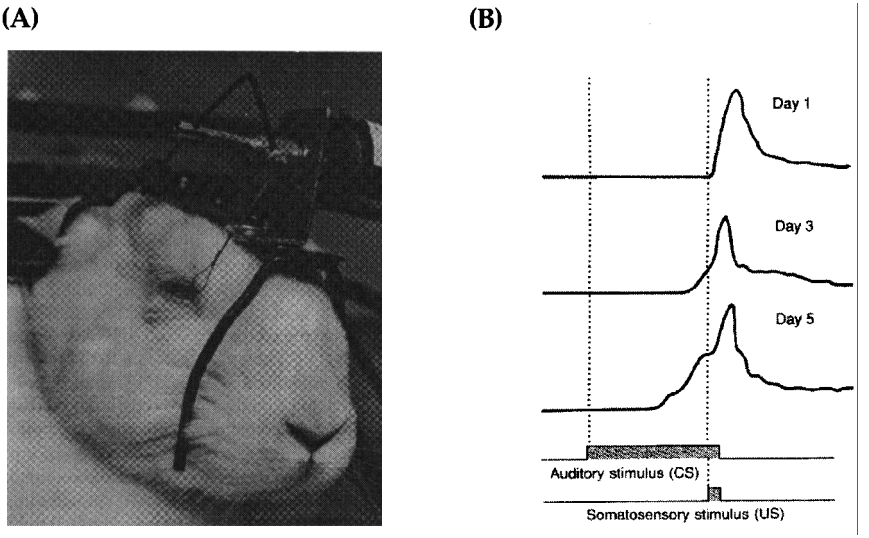

FIG. 3.1. (A) Eyeblink conditioning in rabbits. The unconditioned stimulus (US) is a blink-evoking corneal airpuff or paraorbital shock. If the US is reliably preceded by a tone or light conditioned stimulus (CS), the animal comes to generate a conditioned blink to the CS alone. Conditioned responding is recorded by noting the displacement of a wire sutured to the subject's eyelid. (B) A CS–US pairing trial consists of a presentation of the CS, followed by US presentations; CS and US co-terminate. Initially (day 1), there is blink responding to the US but not the CS; with repeated CS–US pairings (day 5), a conditioned blink is evoked in response to the CS which anticipates US arrival. Figures reprinted from Carlson, 1991.

conditioning paradigms where there are complex temporal relationships between the CS and US, or complex relationships among multiple CSs and the US, eyeblink conditioning is generally disrupted by hippocampal-region damage (e.g. Moyer et al., 1990; Solomon & Moore, 1975).

Given the extensive available data on classical conditioning for both intact and hippocampal-lesion animals, together with the existence of formal models which characterise normal classical conditioning behaviours, it would seem that a mechanistic interpretation of hippocampal function might be most clearly revealed within this elementary form of associative learning. As such, we began our computational modelling of hippocampal function in a ''top-down'' fashion, with a well-defined computational model that accounts for a wide range of behavioural phenomena in classical conditioning, and accurately predicts how behaviour should be affected by hippocampal-region damage (Gluck & Myers, 1993; Myers & Gluck, 1994). In particular, we have suggested that the hippocampus is required to form new stimulus representations to facilitate learning, but not necessarily for simpler stimulus-response learning (Gluck & Myers, 1993). The most salient result from our analyses has been a formal characterisation of how lesioned animals solve tasks differently from normal

intact animals, even if this is not always evident from comparing performance on the initial learning task. Later we will note how this account of hippocampal function is broadly compatible with the previous suggestion by Eichenbaum and Cohen that animals with hippocampal lesions are characterised by an inability to apply learned knowledge flexibly in a new situation (Eichenbaum et al., 1992).

Application to Animal Data

Evidence for Redundancy Compression. Our theory's first bias, redundancy compression, is a bias to compress—or make more similar—the representations of co-occurring or redundant stimuli. Perhaps the simplest paradigm in which redundancy compression is expected in *sensory preconditioning.* Consider two stimulus cues, A and B, perhaps a light and a tone. If these two cues are highly salient, then we would expect the representations they evoke in the hippocampal network hidden layer to be highly distinct—like the examples shown in Fig. 2.2A. As such, there should be very little generalisation between A and B; if A is subsequently paired with the US (''A+ training''), a test presentation of B should evoke very little response. However, if prior to the A + training, there are repeated nonreinforced trials pairing A and B (''AB-pre-exposure''), the situation changes. Redundancy compression makes the representations of co-occurring cues A and B more similar, as in the example of Fig. 2.2B. This will increase generalisation between A and B, so that subsequent A + training will transfer partially to B, and, thus, a test presentation of B will evoke some conditioned responding. Figure 3.2 shows this sensory preconditioning effect in the intact model (Gluck & Myers, 1993). Sensory preconditioning is likewise seen in intact animals (Thompson, 1972). Because our model assumes that sensory preconditioning arises from hippocampal-dependent representational compression during the pre-exposure phase, the enhanced generalisation transfer is not seen in the hippocampal lesioned model (Fig. 3.2). Similarly, hippocampal damage through fornix-fimbrial lesion eliminates sensory preconditioning in rabbits (Port & Patterson, 1984).

Sensory preconditioning illustrates how the increased similarity of stimulus representations can enhance stimulus generalisation. Earlier we noted another implication of increased similarity of stimulus representations: an expected impairment for discriminability. To the extent that two stimuli have similar representations, we expect that discriminating between these patterns will be slower and more difficult than between stimuli with more distinct representations.

We can test this expectation in the following simple variation on sensory preconditioning. We expect that exposure to the compound AB will retard later learning to discriminate A and B in the intact model, as generalisation between the compressed representations will make it harder to map the stimuli to opposite responses. Figure 3.2B indicates that the intact model indeed shows this

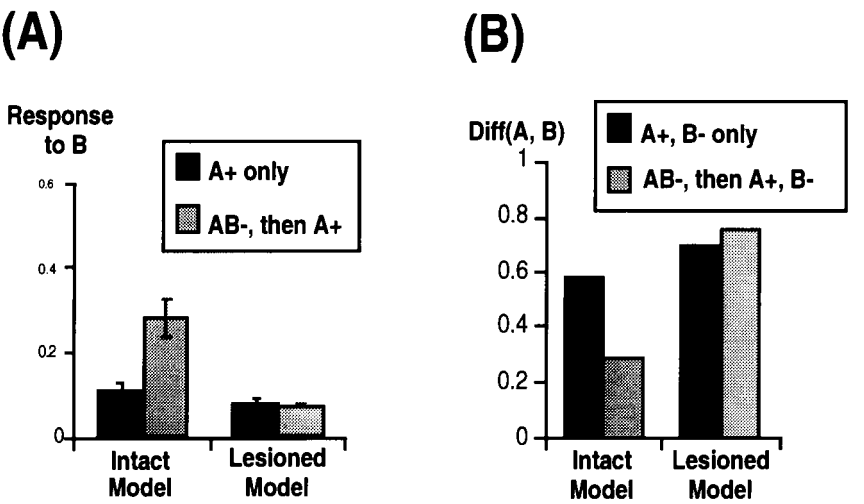

FIG. 3.2. Simulations results with the intact and lesioned models. (A) *Sensory preconditioning*: unreinforced pre-exposure to a compound of two stimuli, AB, followed by training to respond to A, results in stronger responding to B alone than in a control condition with no pre-exposure. The intact but not lesioned model shows this effect; fornix lesion similarly eliminates sensory preconditioning in rabbits (Port & Patterson, 1984). (B) *Compound preconditioning*. Unreinforced pre-exposure to AB slows later training to discriminate A and B, as shown by less relative difference in responding to A and B (Diff{A,B}) after, in the intact model. Intact rats show compound preconditioning (Lubow, et al, 1976); the model predicts hippocampal-region lesions should eliminate the effect.

compound preconditioning effect: the relative difference in responding to A and B, *Diff{A,B}*, following 100 training blocks is reduced if preceded by 20 AB pre-exposure trials, relative to a control condition with no pre-exposure (Gluck & Myers, 1993). In contrast, the lesioned model, with no redundancy compression, shows no difference between the two conditions. This effect is known as the *compound preconditioning* effect, and has been reported both for intact rodents and normal human children (Lubow, Rifkin, & Alek, 1976); this effect has not been tested in hippocampal-lesioned animals (or amnesic humans) and, thus, the elimination of compound preconditioning by hippocampal damage remains a novel prediction of the model.

Note that this is a particularly interesting prediction because we expect a relative enhancement of learning for the lesion group relative to the control group. This is in contrast to the predictions of an expected lesion-related deficit in many other tasks. Such lesion-related deficits can be difficult to interpret because lesion-related deficits can often result from other confounding factors. Thus, a prediction of a lesion-related facilitation is generally more powerful than a predicted lesion-related deficit.

Evidence for Predictive Differentiation. The second representational bias assumed to depend on hippocampal mediation is predictive differentiation—a bias to decrease the similarity of the representations of stimuli that are to be mapped to different outputs. The simplest paradigm in which differentiation is expected is in a simple discrimination task in which two stimuli, A and B, are associated with different responses (e.g. A+, B– training where A predicts the US but B does not). In the intact model, the hippocampal-region network constructs new internal representations which facilitate this simple discrimination by decreasing the similarity between A and B. These new differentiated representations are acquired by the cortical network's hidden layer, which can then easily map these to different responses, as the task requires. Interestingly, the lesioned model shows no particular deficit on this task, and learns as quickly at the intact model (Fig. 3.3A; Gluck & Myers, 1993). This is because, for this simple task, the pre-existing (fixed) hidden layer representations in the lesioned model's cortical network are likely to at least partially distinguish A and B, and so all the network must do is map these representations to the correct responses.

Consistent with this behaviour, hippocampal lesions do not impair learning a simple discrimination in a variety of preparations (e.g. Good & Honey, 1991; Jarrard, 1993; Jones & Mishkin, 1972; Ross, Orr, Holland, & Berger, 1984; Silveira & Kimble, 1968; Zola-Morgan & Squire, 1986; Zola-Morgan et al., 1992). In some cases, hippocampal-region damage has even been shown to facilitate learning (e.g. Eichenbaum, Fagan & Cohen, 1986; Eichenbaum et al., 1988; Eichenbaum et al., 1991; Port, Mikhail, & Patterson, 1985; Schmaltz & Theios, 1972). These data led to early speculations that the intact hippocampus impaired simple classical conditioning (e.g. Port, Romano & Patterson, 1986). In fact, our lesioned model does show a slight facilitation relative to the intact model (Fig. 3.3A). However, this paradoxical facilitation of simple discrimination learning in the lesioned model results from the fact that the lesioned model learns less than the intact model, which is forming new stimulus representations.

The additional time and effort spent by the intact model (and, presumably, by intact animals) in constructing these new and differentiated stimulus representations can be very helpful if task demands change in a way that preserves which cues are relevant, even if the associations to these relevant cues are different. A simple example of this occurs in the *easy–hard transfer* paradigm, in which animals are first trained on an ''easy'' discrimination (e.g. black vs. white) and then transferred to a ''hard'' discrimination along the same stimulus continuum (e.g. dark grey vs. light grey). This transfer facilitates learning even more than an equivalent amount of pre-training on the hard discrimination itself (e.g. Lawrence, 1952; Riley, 1968; Terrace, 1963). The intact model correctly shows this effect (Fig. 3.3b; Gluck & Myers, 1993): in terms of the difference in responding, *Diff(H+, H–)*, between the two hard stimuli in the hard task, pre-training on the easy task leads to better performance

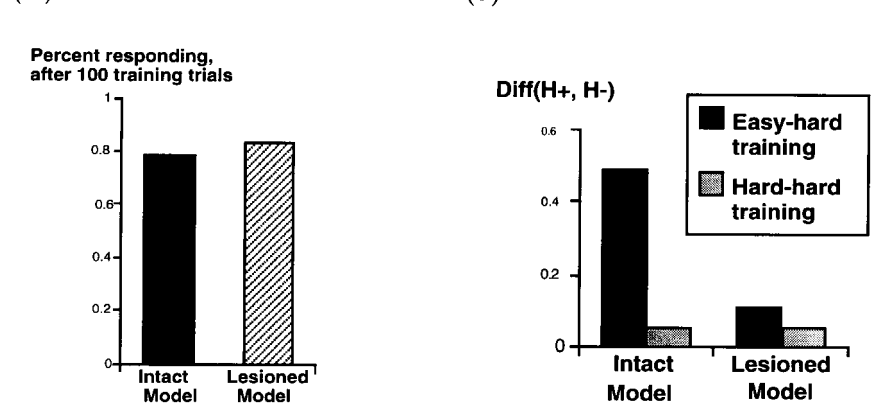

FIG. 3.3. Simulation results with the intact and lesioned models. (A) *Stimulus discrimination*: training to respond to A but not to B. During this task, the intact model forms new stimulus representations, which differentiate A and B, and then maps them to opposite responses. The lesioned model simply maps from pre-existing (fixed) representations in the cortical network to the correct responses. As a result, there is no impairment in conditioned discrimination learning in the lesioned model. Similarly, there is generally no impairment in simple discrimination learning in hippocampal-damaged animals. Plotted from data presented in Myers et al., 1996. (B) *Easy–hard transfer*: learning a hard discrimination between A and B is facilitated by prior training on an easier discrimination along the same stimulus continuum, in the intact but not lesioned model. This is shown by less relative difference in responding to A and B (Diff(A,B)) after a fixed amount of training in the lesioned model than in the intact model. Intact animals show this effect (e.g. Lawrence, 1952; Riley, 1968; Terrace, 1963); the prediction of hippocampal-dependence remains to be tested in animals. Reprinted from Myers et al., 1995.

than pre-training on the hard task. During pre-training on the easy task, the hippocampal-region network is assumed to differentiate the representations of the two stimuli, which predict different outcomes. As these two stimuli are assumed to differ on only a single feature dimension (such as brightness), that dimension will be differentiated, so that there is decreased generalisation for stimuli with differing values of this feature dimension. This differentiation will help discriminate the stimuli in the subsequent hard task, because they differ along the same dimension. In the control condition, with pre-training on the hard task, the same basic mechanisms operate, but they are slower because the stimuli are harder to distinguish, and therefore the pre-training is not so effective. In the lesioned model, with no differentiation mechanisms, the easy–hard transfer effect is not obtained. This leads us to predict that hippocampal-lesioned animals will not show easy–hard transfer, a novel prediction that remains to be tested.

Contextual Processes. Many of the learning deficits associated with hippocampal damage can be described as context deficits, as they suggest an inability to incorporate information about the environmental conditions under

which an event occurs (cf Hirsh, 1974). In studies of classical conditioning it is possible to manipulate the experimental setup—consisting of the sight, smell, and texture of the training chamber—to explicitly test for the effects of contextual information. The simplest of contextual manipulations involves training an animal respond to a given CS in one context and then testing for responding to the CS in a novel context (Fig. 3.4A). Under these conditions, a normal animal will generally respond less to the CS in the novel context than in the training context (e.g. Hall & Honey, 1989). This context-sensitive decrement in responding suggests that the animal has included some information about the context in the association formed between the experimentally defined CS and the US. In contrast, a hippocampal-lesioned animal does not show this decrement in responding after a context shift (Honey & Good, 1993; Penick & Solomon, 1991). It should be noted that this does not reflect a general inability to perceive contextual cues in the hippocampal-lesioned animal, as lesioned animals can still learn to discriminate contexts (e.g. Good & Honey, 1991; Phillips & Le Doux, 1994). What seems to be disrupted in the lesioned animal is the ability to use context to interpret the meaning of conditioned cues (Myers & Gluck, 1994).

The cortico-hippocampal model provides a framework for understanding why this pattern of preserved and impaired contextual processing might emerge in lesioned animals (Myers & Gluck, 1994). Input to the hippocampal-region autoencoder is assumed to represent not only the presence of any CSs and USs, but also the complete set of background or contextual cues that are present. These contextual cues are ''tonic'' or unchanging throughout the experiment,

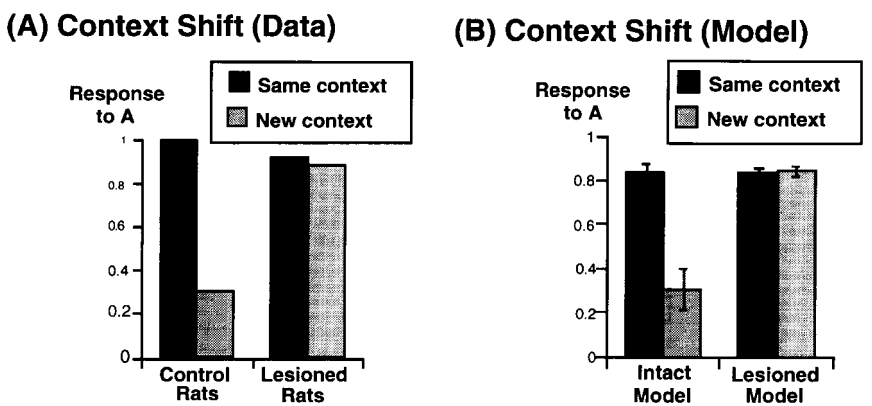

FIG. 3.4. (A) In normal animals, a conditioned response to A may show a decrement if A is then presented in a new context (Hall & Honey, 1989); hippocampal-lesioned animals do not show this response decrement after a context shift (Honey & Good, 1993; Penick & Solomon, 1991; Figure replotted from data presented in Penick & Solomon, 1991.) (B) The intact but not lesioned model correctly shows this response decrement with context shift (Myers & Gluck, 1994; Figure reprinted from Myers et al., 1995a).

while the CSs and USs are ''phasic'' and only occur sporadically during training.

As the hippocampal-region autoencoder learns to predict CS–US associations, it also learns to reconstruct its input, including these contextual cues (Fig. 3.5). Thus, contextual information is included in the representations formed in

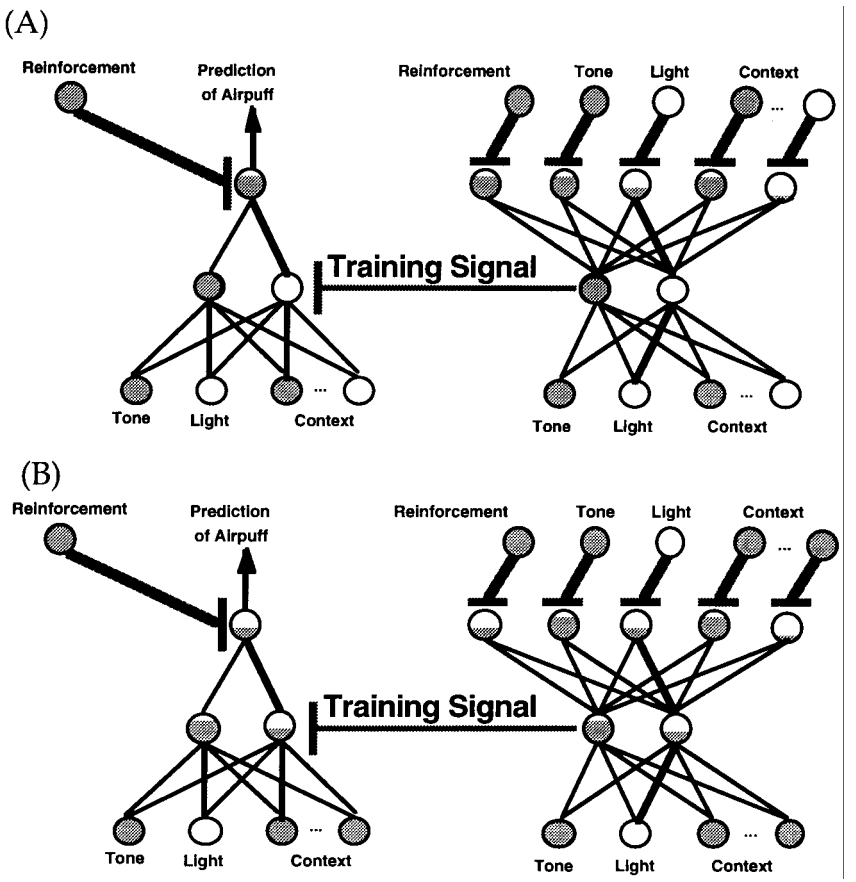

FIG. 3.5. (A) Schematic of activations in an intact model trained to respond to the presence of a tone stimulus in a particular context (represented by the pattern of input activations). The network has learned to output a strong response at the cortical network output nodes, and a good reconstruction of the input at the hippocampal-region network output nodes. The hidden layer representation that has evolved involves strong activation of one hidden-layer node, and weak activation of the other. (B) Schematic of activations when the tone stimulus is presented in a new context (represented by a different pattern of input activations on the nodes that represent contextual inputs. The hidden layer activation in both networks is more weakly activated than in the trained pattern shown in (A), and the behavioural response output from the cortical network is weaker as well. Thus, the model shows a decremented response when a trained stimulus is presented in a novel context.

the autoencoder's internal layer. As a result, the context influences the internal layer representations in such a way that, if the CS is later presented in a new context, the representation of that CS will be more weakly activated than it was in the training context. Figure 3.4B shows this response decrement after context shift in the intact version of the model, just as observed in the intact animal data of Fig. 3.4A. The model also correctly expects that with extended training the hippocampal representation will exclude irrelevant contextual information and become more and more context-dependent (Myers & Gluck, 1994); there is some evidence of the same kind of time-dependence of contextual sensitivity in animals (Hall & Honey, 1990; see also Myers & Gluck, 1994, for review). In the lesioned version of the model, however, no new internal representations are formed, and so there is no means for incorporating contextual information into the representations of a CS. As a result, the lesioned model is relatively insensitive to context (Fig. 3.4B). This is consistent with the strong responding generally shown by hippocampal-lesioned animals in a new context (Fig. 3.4A). The cortico-hippocampal model can similarly account for results from a range of context studies (Myers & Gluck, 1994) and provides a computational instantiation and elaboration of several existing qualitative theories which have implicated the hippocampus in context learning (Hirsh, 1974; Penick & Solomon, 1991).

Another context-related behavioural phenomenon is *latent inhibition* (Fig. 3.6A) in which unreinforced pre-exposure to a cue slows the rate of subsequent conditioning to that cue (Lubow, 1973). Within the cortico-hippocampal model, latent inhibition can be understood as a variation of compound preconditioning, in which the context is considered as another cue. If we represent the context as a single, compound cue, X, and the CS as another cue A, then conditioning is expressible as learning the discrimination AX+, X–, that is, respond when A is present in context X, but not to X alone. Latent inhibition involves a prior phase of AX–, X– training followed by the AX+, X– training. In the initial phase, where neither AX nor X predicts the US, redundancy compression in the intact model results in an increase in similarity between the representations of A and X, making it more difficult to discriminate AX from X alone. As a result of this, the later phase of learning to respond to AX but not X is made more difficult, which is the observed latent inhibition effect.

This is exactly what we see in simulations of the intact model (Fig. 3.6B). The lesioned model, with no representational changes, does not perform redundancy compression during the pre-exposure phase and therefore learning in the subsequent acquisition phase is not slowed down (Fig. 3.6B). Consistent with the model, broad hippocampal-region damage eliminates latent inhibition (Fig. 3.6A), and these lesioned animals learn about the pre-exposed cue as quickly as a nonpre-exposed cue (Kaye & Pearce, 1987; Solomon & Moore, 1975).

It should be noted that a smaller lesion, limited to the hippocampus but not including surrounding cortices, does not eliminate latent inhibition (Honey &

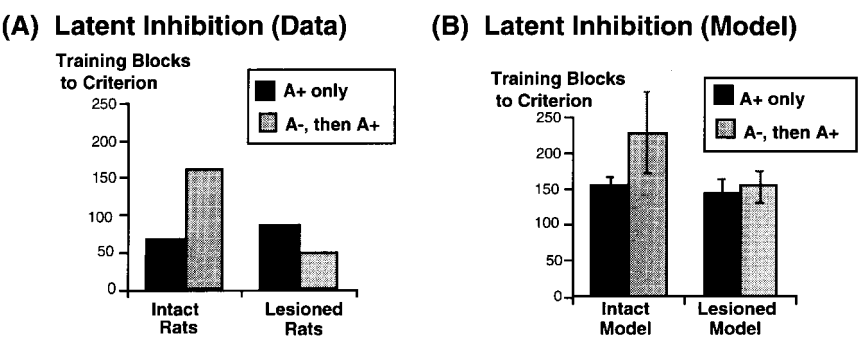

FIG. 3.6. Latent inhibition, in which unreinforced pre-exposure to a cue A slows later acquisition of conditioned responding to A. (A) Intact animals show this effect (Lubow, 1973), reflected in longer training time to learn a strong response to A. Broad hippocampal-region lesion eliminates latent inhibition (Kaye & Pearce, 1987; Solomon & Moore, 1975). Figure plotted from data presented in Solomon & Moore (1975). (B) The intact model correctly shows latent inhibition while the lesioned model does not. Figure reprinted from Myers et al. 1995.

Good, 1993). In more recent work, Myers et al. (1995) have suggested that this result is consistent with a mapping of function to anatomy in which stimulus compression may occur in the entorhinal cortex without requiring the hippocampus proper. Eichenbaum and Bunsey (1995) have argued for a similar functional interpretation of the cortical areas overlaying the hippocampus, especially the entorhinal cortex, in their analyses of intact vs. lesioned animals in olfactory simple discrimination learning.

Summary of Model Performance in Animal Studies. Given the extensive available data on classical conditioning for both intact and hippocampal-damaged animals, together with the existence of formal models which characterise normal classical conditioning behaviours, we have shown how a mechanistic interpretation of hippocampal function can be clearly revealed within this most elementary form of associative learning. The most salient result from our analyses is a formal characterisation of how lesioned animals solve tasks differently from normal intact animals, even if this is not always evident from comparing performance on the initial learning task. Analyses of sensory preconditioning and compound preconditioning are interpreted as reflecting the hippocampal-dependent bias for *redundancy compression*, while stimulus discrimination and easy–hard transfer are interpreted as reflecting *predictive differentiation.* Further evidence for these hippocampal-dependent representational biases comes from analyses of contextual effects seen in studies of context shifts and latent inhibition.

Human Hippocampal Amnesia and Conditioning

What can all these studies of hippocampal-lesion effects in animals tell us about human medial-temporal lobe (especially hippocampal-based) amnesia? One of the most important reasons for studying classical conditioning is that the neural structures and behavioural properties are very similar across a wide range of species, including humans. Classical conditioning has been extensively studied in humans, using a variety of preparations (cf Daum et al., 1993; Lye, O'Boyle, Ramsden, & Schady, 1988; Solomon, Stowe, & Pendlebury, 1989; Topka, Valls-Sole, Masaquoi, & Hallet, 1993). As with animals, the most popular preparation is the conditioned eyeblink response. Subjects typically wear a hat-type device fitted with tubing to allow light airpuffs to be directed at the cornea; the hat also produces an infrared beam which is reflected off the white of the eyeball and interruption of this beam by an eyeblink is recorded as a response. Just as in animals, repeated pairing of a previously neutral tone or light CS with the blink-evoking airpuff US results in an anticipatory conditioned eyeblink response to the CS alone. At the neural level of analysis, the cerebellum seems to be the necessary substrate for conditioned eyeblink responding in humans as well as in infra-human species (Daum et al., 1993; Lye et al., 1988; Solomon et al., 1989; Topka et al., 1993).

Intact Delay Conditioning in Amnesia

Four studies have examined classical conditioning and human hippocampal amnesia. Weiskrantz and Warrington (1979) first examined eyeblink conditioning in two amnesia patients. One patient had hippocampal damage as characterised by an episode of viral encephalitis, while the other was an alcoholic Korsakoff patient. Although details indicating damage specific to a neuroanatomical focus was not provided, both patients neuropsychologically demonstrated memory defects that are commonly attributable either to pathology to the medial temporal lobe or diencephalic damage. Both amnesic individuals showed clear evidence of conditioned learning without impairment. Interestingly, when later asked to verbalise a declarative account about the experiment they were unaware of their performance.

Daum et al. (1989) investigated eyelid conditioning in three amnesic patients with bilateral temporal damage. Patient 1 had a history of epilepsy and alcohol abuse, patient 2 had a history of encephalitis and epileptic seizures, and patient 3 suffered from epileptic seizures. All three patients showed significant impairments on standard neuropsychological memory tests and indicated hippocampal lesions as supported by EEG or CT measures. As in the amnesics of Weiskrantz and Warrington (1979) all three of these patients were able to condition without impairment using the delay paradigm. However, when assessed on a more complex two-CS, conditional discrimination and reversal experiment, all three failed to effectively learn the task.

Woodruff-Pak (1993) conducted a series of eyeblink studies in the well-studied amnesic, HM. As a result of bilateral removal of medial-temporal lobe structures, including the hippocampus, to relieve the patient's epilepsy, HM because densely amnesic for learning and remembering new information. A second amnesic with medial-temporal-lobe lesions resultant from encephalitis was included for comparison, as were two normal adult controls. It took HM longer (473 trials) to reach learning criterion than a control subject (315 trials). The encephalitic amnesic took 119 trials and his control took 15. In the case of HM, although the extent of his hippocampal removal lesion is well known, some cerebellar atrophy in the vermis and cerebellar hemispheres has recently been noted. Damage to the cerebellum, more so than hippocampal damage, may have been more important as the cause of his slower impairment (yet not abolishment) in conditioning.

Gabrieli et al. (1995) recently tested seven well-characterised amnesic subjects and seven well-matched normal controls. In reviewing the experimental details of the three previous amnesic papers, the studies may be criticised for using inferior eyeblink measurement apparatus and technique, too few amnesic subjects, poor quality of patient diagnoses and characterisation, and not including control subjects for comparisons. In this paper, the amnesic etiologies ranged from encephalitis (n = 2), anoxia (n = 2), aneurysm (n = 1), status epilepticus (n = 1), and closed head injury (n = 1). Five of the patients had neuroimaging data confirming bilateral hippocampal damage. All were densely amnesic in terms of standard neuropsychological memory tests. The subjects were assessed in terms of pseudoconditioning (no pairing of CS and US), conditioning (CS–US pairings) and extinction contingencies (CS alone). In support of the three previous amnesia and delay conditioning studies, no statistical differences were found, with amnesics learning the conditioned response as well as the normal controls (see Fig. 3.7).

Most of the other conditioning paradigms in which hippocampal-lesioned animals are affected, such as simple discrimination reversal, sensory preconditioning, latent inhibition, etc., remain to be tested in human hippocampal amnesics. For example, normal intact humans show latent inhibition of the conditioned eyeblink response (Hulstijn, 1978). It remains to be seen if human hippocampal amnesics will show a similar disruption in latent inhibition as do hippocampal-damaged animals. Similarly, we expect that human hippocampal amnesics may show disruptions in other paradigms (simple discrimination reversal, latent inhibition, etc.) that parallel those seen in hippocampal-lesioned animals.

Having reviewed the evidence for our theory within studies of animal conditioning, and described some preliminary studies of its relevance to human eyeblink conditioning in amnesia, we turn now to consider a more clearly cognitive form of human learning: category learning.

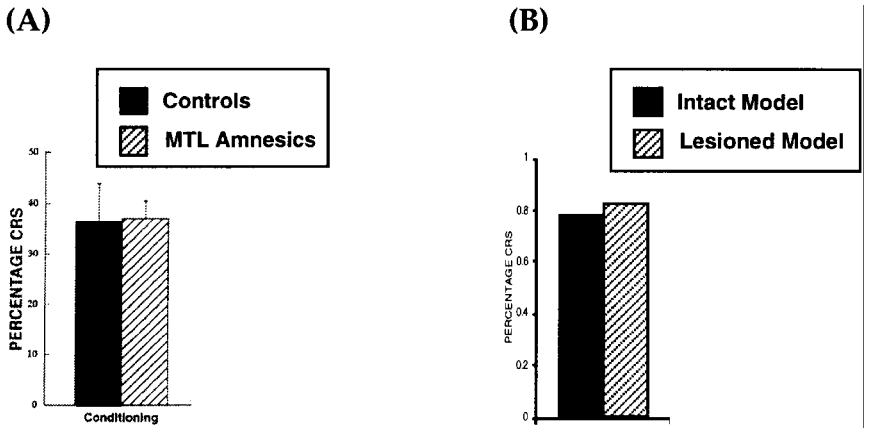

FIG. 3.7. (A) Eyeblink conditioning in humans is not disrupted by hippocampal damage in human amnesics. Figure reprinted from Gabrieli et al. (1995). (B) Likewise, the lesioned model predicts no deficits. Figure reprinted from Myers et al. (1995b).

4. APPLICATION OF MODEL TO CATEGORY LEARNING

In the previous section of the paper we discussed how the predictions of the cortico-hippocampal model have been tested in both animal and human eyeblink conditioning studies. In this section our focus shifts to human category learning, as we explore the possibility of extending the cortico-hippocampal model further to aid our understanding of higher cognitive processes. Category learning plays a pivotal role in human cognition and has generated a substantial body of empirical and theoretical work (Estes, 1993). By extending a neurobiologically based model of associative learning to account for category learning data we hope to start to establish a bridge over which we can begin to integrate animal and human data from studies of the neural and behavioural bases of cognition.

In a typical category learning experiment, subjects are presented with a series of multidimensional stimuli. The subject's task is to determine to which of two or more categories each stimulus belongs. The subject is then provided with corrective feedback, indicating the correct response.

The behavioural correspondences between animal conditioning and category learning were described by Gluck and Bower (1988a,b) who demonstrated that the Rescorla-Wagner model of conditioning (Rescorla & Wagner, 1972), a simple associative model, can be extended to account for a variety of category learning phenomena in humans. Gluck and Bower's connectionist extension of the Rescorla-Wagner model can account for many fundamental aspects of

human category learning (Gluck, Bower, & Hee, 1989) including base-rate neglect (Gluck & Bower, 1988a,b), the interaction of exemplar similarity and linear separability in determining categorisation difficulty (Medin & Schwanenflugel, 1981), the relationship between classification and recognition memory for instances (Hayes-Roth & Hayes-Roth, 1977), and the impact of correlated attributes on classification (Medin, Altom, Edelson, & Freko, 1982). The modelling of Gluck and Bower as well as other researchers working in a similar vein (Estes, 1993; Shanks, 1991) provide us with an *a priori* basis for believing that there are significant behavioural correspondences between animal conditioning and human category learning. This leads us to ask whether or not there might be neurobiological correspondences, too.

Some preliminary progress in this direction was made by Knowlton et al. (1994) who addressed this issue of neurobiological correspondences between conditioning and category learning in a series of studies in which amnesic patients and control subjects were given a series of multidimensional stimuli and told to classify them into one of two categories. For example, in one task, there were four ''tarot'' cards, each with a unique set of abstract geometric patterns. The stimuli were a drawing of up to four of these cards, and subjects were asked to predict whether there would be ''good'' or ''bad'' weather based on these cards. The actual weather was determined according to a probabilistic rule based on the cards, and each card was therefore a partial predictor of the weather. The probabilistic rule ensured that it was impossible for the subjects to learn the prediction with complete certainty, although it was possible to use the probabilistic card–weather relationships to achieve significantly better-than-chance performance. Two equivalent tasks were also used, with similar underlying logic. These tasks are equivalent to those used by Gluck and Bower (1988b) in an investigation of probabilistic category learning in intact subjects, with the exception that the categories used in the Knowlton et al. study occurred with equal frequency.

In each task, the amnesic patients initially learned to associate the cues with the appropriate outcome at the same rate as control subjects, improving from chance performance (50% correct) to approximately 65% correct (Knowlton et al., 1994). With extended training, however, control subjects eventually outperformed amnesic patients (see Fig. 4.1A). This finding that amnesics outperformed controls only late in training mirrors a similar finding that hippocampal amnesics are initially unimpaired but later impaired at cognitive skill learning (Squire & Frambach, 1990).

As shown in Fig. 4.1B, the cortico-hippocampal model produces learning curves analogous to those observed by Knowlton et al. (Gluck, Oliver, & Myers, in press). In the intact model, the hippocampal-region network develops new stimulus representations to aid learning; these representations may develop over many trials as the model is exposed to a representative subset of training patterns. Early in learning, before these representations develop, the intact

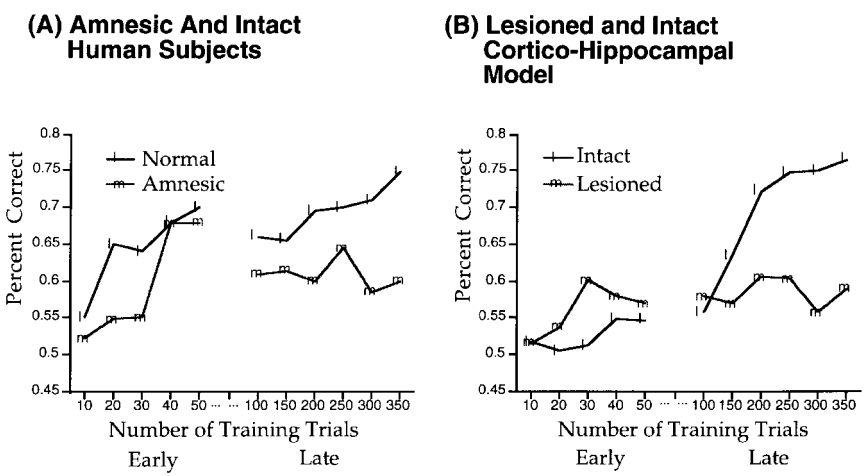

FIG. 4.1. (A) On a probabilistic category learning task (Knowlton et al., 1994), hippocampal amnesics initially learn at the same rate as controls; later in training, the controls continue to improve performance more than the amnesics. Figure reprinted from Knowlton et al. (1994). (B) The cortico-hippocampal model shows a similar effect: the lesioned model initially performs as well as the intact model, but later in training, the lesioned model shows a deficit. Figure reprinted from Gluck and Oliver (in prep.).

model depends on the pre-existing representations in the cortical network, and maps these to (an approximation of) the correct response. Later in training, as the hippocampal representations become available, these new representations are acquired by the cortical network and allow further improvement in performance. In the lesioned model, only the prior cortical representations are ever available. Thus, early in training, performance is similar in the lesioned and intact models; as training progresses, however, and no new representations become available to the lesioned cortical network, the lesioned model shows a deficit relative to the intact model.

The simulation study reported by Gluck, Oliver, and Myers (in press) should be viewed only as preliminary evidence in favour of the cortico-hippocampal model as a model of human category learning. Category learning research has generated a considerable body of data which must be addressed before a model can claim to account for this pervasive aspect of human cognition. Nevertheless, these preliminary analyses point to the possible potential of the cortico-hippocampal model for capturing subtle aspects of amnesic learning which go beyond simply classifying their learning as intact or impaired relative to normals. Moreover, they suggest that careful comparisons of early versus late training difference in learning may be an important factor in understanding the differences between intact and amnesic learning.

5. DISCUSSION

Although most analyses of amnesia have focused on the loss of explicit declarative and episodic memories following hippocampal-region damage, we have tried to show here how insights into amnesia can also be realised by studying hippocampal function in simple procedural, or habit-based, associative learning tasks. Although many simple forms of associative learning are unimpaired by hippocampal damage, more complex tasks that require sensitivity to unreinforced stimuli, configurations of multiple stimuli, or contextual information are impaired by hippocampal damage. We have reviewed here the results of several of our recent papers in which we have argued that these animal conditioning data imply that the hippocampal region plays a critical role in the formation of new stimulus representations during learning (Gluck & Myers, 1993, 1995; Myers & Gluck, 1996; Myers et al., 1995).

Dissociating Hippocampal and Parahippocampal Function

More recently we have begun to extend our modelling efforts to seek a closer rapprochement with the underlying biology. Although the intact model of Fig. 2.4A adopts the simplifying assumption that the hippocampal region functions as a unitary processor, the hippocampal region is, in fact, anatomically divided into several distinct structures—including hippocampal fields CA1 through CA4, dentate gyrus, subiculum, and entorhinal cortex. There are also several inputs to the hippocampal region, including perirhinal and parahippocampal cortices, which provide highly processed polymodal sensory information, as well as a direct connection from unimodal piriform (olfactory) cortex (see Fig. 5.1). In Myers et al. (1995) we hypothesised that the representational function computed in our intact model's hippocampal region may be subdivided, and the subfunctions localised in different anatomical sites around the region; in particular, we proposed that stimulus–stimulus redundancy compression could emerge from the anatomy and physiology of superficial entorhinal cortex. The net result of both entorhinal and hippocampal processing would then be a new stimulus representation constrained by both compression and differentiation biases which could be transferred to long-term storage sites in the cortex.

This hypothesis assumes that a selective hippocampal-lesion (the H-lesion), which does not otherwise damage entorhinal cortex (cf Jarrard & Davidson, 1991), might allow redundancy compression processes to survive. Behaviours that depend mainly on these processes should continue to be exhibited after H-lesion, whereas behaviours that require other representational processes such as predictive differentiation should be disrupted. We can study this situation by constructing a ''Selective H-lesioned'' model, in which the full hippocampal-region network of Fig. 2.4 is reduced to an enthorhinal network only. The long-

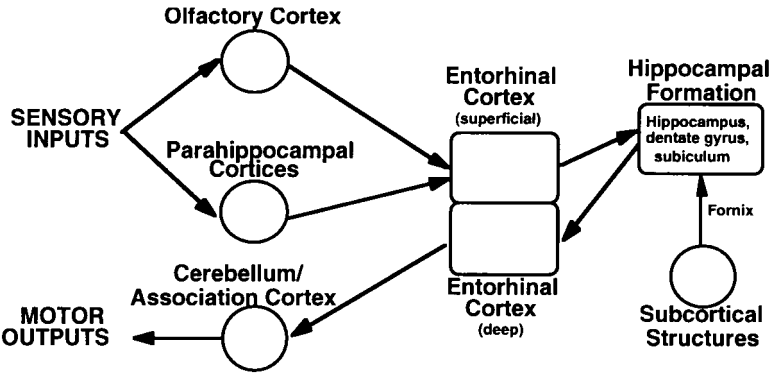

FIG. 5.1. Schematic of major information flow pathways in the hippocampal region. Inputs are provided directly from unimodal olfactory cortex, as well as polymodal cortices including parahippocampal and perirhinal cortex. Information travels through superficial entorhinal cortex, through the hippocampal formation (including hippocampal fields CA1–CA4, dentate gyrus, and subiculum), and returns through deep entorhinal cortex to other cortical areas for long-term storage. There is also a bi-directional pathway through the fornix which provides connection to subcortical structures. Many other connections exist, which are not shown here for simplicity.

term memory network continues to operate as in the intact model, except that the new representations provided by the entorhinal network are biased only by stimulus–stimulus redundancy compression, but not by the other representational biases attributed to the hippocampal region as a whole.

A simple example of a compression-based task is latent inhibition, in which unreinforced pre-exposure to a cue slows later acquisition of a conditioned response to that cue (Lubow, 1973). In our intact model of Fig. 2.4A, latent inhibition is caused by compression of the pre-exposed cue with co-occurring and equally nonreinforced contextual cues. The subsequent increase in learning time results because the model must first redifferentiate the cue from the context before a response can be selectively associated with the cue (Fig. 3.6). Because this effect is assumed to depend primarily on redundancy compression, the entorhinal network in the H-lesioned model is sufficient to produce latent inhibition. Consistent with this model behaviour, rats with selective hippocampal lesions but no entorhinal damage show normal or even enhanced latent inhibition (Honey & Good, 1993; Reilly, Harley & Revusky, 1993). Obviously, more experimental tests need to be performed, but this is at least initial support for the Myers et al. (1995) proposed dissociation of hippocampal and entorhinal function.

Eichenbaum and Bunsey (1995) have recently addressed this same issue and have suggested that the parahippocampal region mediates the ''fusion'' of co-occurring or nearly coincident stimuli; this process is essentially identical to the redundancy compression function we have proposed. It is interesting to note that whereas their fusion theory derives from behavioural observations comparing

paired-associate learning in intact, hippocampal-lesioned and parahippocampal-lesioned animals, our similar compression theory arises from an integration of both physiologically based and behaviourally based computational models of hippocampal-region function. We hope the convergence of these two widely different approaches to theory development is a sign that stimulus compression is a useful and accurate description of parahippocampal-region function.

Modelling Cholinergic Modulation in Septohippocampal Pathways

In other, more recent work, we have also shown that a simple extension of the model can successfully account for the retarded classical conditioning that occurs after septohippocampal disruption via anticholinergic drugs (Myers et al., 1996). This last is important because septohippocampal pathways may be damaged in a variety of ways. Basal forebrain damage may occur in cases of Alzheimer's Dementia, but more commonly occurs in cases of aneurysm rupture of the anterior communicating artery (ACoA). The ACoA aneurysm survivors are an especially interesting group because they often manifest an anterograde amnesia syndrome which is superficially very similar to that shown by hippocampal amnesics (DeLuca & Diamond, 1995)—although the ACoA amnesics have no direct lesion to the hippocampus. The problem in these patients may result from basal forebrain damage, indirectly interfering with proper hippocampal functioning. Neuroimaging to directly assess basal forebrain lesion extent is difficult because ACoA aneurysm is often treated by surgical placement of ferromagnetic clips on the ruptured artery. MRI is therefore contraindicated in these cases as it may result in displacement and heating of the metallic clip which could kill the patient (Klucznik, Carrier, Pyka, & Haid, 1993). Alternate neuroimaging techniques such as angiography and computerised tomography (CT scan) may be used in these amnesic patients recovering from stroke; however, the placement of the metallic clip often obscures precise imaging of the basal forebrain region. Therefore, an interesting potential application of this work is to develop behavioural tests that are sensitive to the damage of particular neural structures—thus aiding in the assessment of lesion extent in these patients.

Summary

Our long-term goal is to use computational models to develop a comprehensive view of hippocampal-region function that illuminates common behavioural functions, and underlying neural mechanisms, in both animal and human memory. Our work to date illustrates some initial progress to this end, and may help provide predictions and directions to guide further empirical and theoretical research in human memory and amnesia. Our hope is that the insights gained from these initial analyses of a simple learning behaviour will lead to a deeper

understanding of the hippocampal region in more complex forms of learning and memory, in both animals and humans.

REFERENCES

Akase, E., Alkon, D.L., & Disterhoft, J.F. (1989). Hippocampal lesions impair memory of short-delay conditioned eyeblink in rabbits. *Behavioral Neuroscience*, *103*(5), 935–943.

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

Anderson, J. (1977). Neural models with cognitive implications. In D. LaBerge & S. Samuels (Eds.), *Basic processes in reading: Perception and comprehension* (pp.27–90). Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Barto, A., & Sutton, R. (1982). Simulation of anticipatory responses in classical conditioning by a neuron-like adaptive element. *Behavioral Brain Research*, *4*, 221–235.

Berger, T., & Orr, W. (1983). Hippocampectomy selectively disrupts discrimination reversal learning of the rabbit nictitating membrane response. *Behavioral Brain Research*, *8*, 49–68.

Berger, T.W., Rinaldi, P., Weisz, D.J., & Thompson, R.F. (1983). Single-unit analysis of different hippocampal cell types during classical conditioning of the rabbit nictitating membrane response. *Journal of Neurophysiology*, *50*(5), 1197–1219.

Blaxton, T.A., Zeffiro, T.A.,Gabrieli, J.D.E., Bookheimer, S.Y., Carrillo, M.C., Theodore, W.H., & Disterhoft, J.F. (1995). Mapping the neuroanatomical correlates of human learning: A PET activation study of eyeblink conditioning. Manuscript submitted for publication.

Bliss, T., & Lomo, T. (1983). Long-lasting potentiation of synaptic transmission in the dentate area of anaesthetized rabbit following stimulation of the perforant path. *Journal of Physiology*, *232*, 331–356.

Buzsaki, G. (1989). Two-stage model of memory-trace formation: A role for ''noisy'' brain states. *Neuroscience*, *31*(3), 551–570.

Carlson, N.R. (1991). *Physiology of behavior* (4th edn.). Needham Heights, MA: Allyn & Bacon.

Clark, G.A., McCormick, D.A., Lavond, D.G., & Thompson, R.F. (1984). Effects of lesions of cerebellar nuclei on conditioned behavioral and hippocampal neuronal responses. *Brain Research*, *291*, 125–136.

Cohen, N. (1984). Preserved learning capacity in amnesia: Evidence for multiple learning systems. In L. Squire & N. Butters (Eds.), *Neuropsychology of memory* (pp.83–103). New York: Guilford Press.

Daum, I., Channon, S., & Canavan, A. (1989). Classical conditioning in patients with severe memory problems. *Journals of Neurology, Neurosurgery and Psychiatry*, *52*, 47–51.

Daum, I., Schugens, M.M., Ackermann, H., Lutzenberger, W., Dichgans, J., & Birbaumer, N. (1993). Classical conditioning after cerebellar lesions in humans. *Behavioral Neuroscience*, *107*(5), 748–756.

DeLuca, J., & Diamond, B. (1995). Aneurysm of the anterior communicating artery: A review of neuroanatomical and neurophysiological sequelae. *Journal of Clinical and Experimental Neuropsychology*, *17*(1), 100–121.

Disterhoft, J.F., Coulter, D.A., & Alkon, D.L. (1988). Conditioning-specific biophysical alterations in rabbit hippocampus. In C.D. Woody, D.L. Alkon, & J.L. McGaugh (Eds.), *Cellular mechanisms of conditioning and behavioral plasticity* (pp.89–104). New York: Plenum Publishing Corporation.

Douglas, R., & Pribam, K. (1966). Learning and limbic lesions. *Neuropsychologia*, *4*, 192–222.

Eichenbaum, H., & Buckingham, J. (1990). Studies on hippocampal processing: Experiment, theory and model. In M. Gabriel & J. Moore (Eds.), *Learning and computational neuroscience: Foundations of adaptive networks* (pp.171–231). Cambridge, MA: MIT Press.

Eichenbaum, H., & Bunsey, M. (1995). On the binding of associations in memory: Clues from studies on the role of the hippocampal region in paired associated learning. *Current Directions in Psychological Science*, *4*.

Eichenbaum, H., Cohen, N.J., Otto, T., & Wible C. (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* (pp.163–204). Oxford: Oxford University Press.

Eichenbaum, H., Fagan, A., & Cohen, N.J. (1986). Normal olfactory discrimination learning set and facilitation of reversal learning after medial-temporal damage in rats: Implication for an account of preserved learning abilities in amnesia. *Journal of Neuroscience*, *6*(7), 1876–1884.

Eichenbaum, H., Fagan, A., Mathews, P., & Cohen, N.J. (1988). Hippocampal system dysfunction and odor discrimination learning in rats: Impairment or facilitation depending on representational demands. *Behavioral Neuroscience*, *102*(3), 331–339.

Eichenbaum, H., Otto, T., & Cohen, N. (1994). Two functional components of the hippocampal memory system. *Behavioral and Brain Sciences*, *17*(3), 449–518.

Eichenbaum, H., Otto, T., Wible, C., & Piper, J. (1991). Building a model of the hippocampus in olfaction and memory. In J. Davis & H. Eichenbaum (Eds.), *Olfaction as a model for computational neuroscience* (pp.167–210). Cambridge, MA: MIT Press.

Estes, W.K. (1993). Models of categorization and category learning. *The Psychology of Learning and Motivation*, *29*, 15–56.

Gabrieli, J., McGlinchey-Berroth, R., Carrillo, M., Gluck, M., Cermak, L., & Disterhoft, J. (1995). Intact delay-eyeblink classical conditioning in amnesia. *Behavioral Neuroscience*, *109*(5), 819–827.

Glickstein, M., Hardiman, M.J., & Yeo, C.H. (1983). The effects of cerebellar lesions on the conditioned nictitating membrane response of the rabbit. *Journal of Physiology*, *341*, 30–31.

Gluck, M., & Bower, G. (1988a). From conditioning to category learning: An adaptive network model. *Journal of Experimental Psychology: General*, *117*(3), 225–244.

Gluck, M., & Myers, C. (1993). Adaptive stimulus representations: A computational theory of hippocampal-region function. In S. Hanson, J. Cowen, & C. Giles (Eds.), *Advances in neural information processing systems 5* (pp.937–944). San Mateo, CA: Morgan Kaufmann.

Gluck, M., & Myers, C. (1995). Representation and association in memory: A neurocomputational view of hippocampal function. *Current Directions in Psychological Science*, *4*(1), 23–29.

Gluck, M.A., & Bower, G.H. (1988b). Evaluating an adaptive network model of human learning. *Journal of Memory and Language*, *27*, 166–195.

Gluck, M.A., Bower, G.H., & Hee, M.R. (1989). A configural-cue network model of animal and human associative learning. In *Proceedings of 11th Annual Conference of Cognitive Science Society, Ann Arbor: MI* (pp.323–332).

Gluck, M.A., & Oliver, L.M. (1995). Early vs. late performance in normal and hippocampal-impaired learning: A neurocomputational perspective on Knowlton, Squire and Gluck (1994). Manuscript in preparation.

Gluck, M.A., Oliver, L.M., & Myers, C.E. (in press). Late-training amnesic deficits in probabilistic category learning: A neurocomputational analysis. *Learning and Memory*.

Good, M., & Honey, R.C. (1991). Conditioning and contextual retrieval in hippocampal rats. *Behavioral Neuroscience*, *105*(4), 499–509.

Gormezano, I., Kehoe, E.J., & Marshall, B.S. (1983). Twenty years of classical conditioning research with the rabbit. *Progress in Psychobiology and Physiological Psychology*, *10*, 197–275.

Hall, G., & Honey, R. (1989). Contextual effects in conditioning, latent inhibition, and habituation: Associative and retrieval functions of contextual cues. *Journal of Experimental Psychology: Animal Behavior Processes*, *15*(3), 232–241.

Hall, G., & Honey, R. (1990). Context-specific conditioning in the conditioned-emotional-response procedure. *Journal of Experimental Psychology: Animal Behavior Processes*, *16*(3), 271–278.

Hanson, S.J., & Burr, D.J. (1990). What connectionist models learn: Learning and representation in connectionist networks. *Behavioral and Brain Sciences*, *13*, 471–518.

Hayes-Roth, B., & Hayes-Roth, F. (1977). Concept learning and the recognition and classification of exemplars. *Journal of Verbal Learning and Verbal Behavior*, *16*, 321–338.

Hinton, G. (1989). Connectionist learning procedures. *Artificial Intelligence*, *40*, 185–234.

Hirsh, R. (1974). The hippocampus and contextual retrieval of information from memory: A theory. *Behavioral Biology*, *12*, 421–444.

Honey, R., & Good, M. (1993). Selective hippocampal lesions abolish the contextual specificity of latent inhibition and conditioning. *Behavioral Neuroscience*, *107*(1), 23–33.

Hulstijn, W. (1978). The orienting reaction during human eyelid conditioning following preconditioning exposures to the CS. *Psychological Research*, *40*, 77–88.

James, W. (1896). *The principles of psychology*. New York: Henry Holt & Co.

Jarrard, L. (1993). On the role of the hippocampus in learning and memory in the rat. *Behavioral and Neural Biology*, *60*, 9–26.

Jones, B., & Mishkin, M. (1972). Limbic lesions and the problem of stimulus-reinforcement association. *Experimental Neurology*, *36*, 362–377.

Kaye, H., & Pearce, J. (1987). Hippocampal lesions attenuate latent inhibition and the decline of the orienting response in rats. *Quarterly Journal of Experimental Psychology*, *39B*, 107–125.

Kehoe, E.J. (1988). A layered network model of associative learning. *Psychological Review*, *95*(4), 411–433.

Klucznick, R.P., Carrier, D.A., Pyka, R., & Haid, R.W. (1993). Placement of ferromagnetic intracerebral aneurysm clip in a magnetic field with a fatal outcome. *Radiology*, *187*, 855–856.

Knowlton, B., Squire, L., & Gluck, M. (1994). Probabilistic classification learning in amnesia. *Learning and Memory*, *1*, 106–120.

Krushke, J.K. (1992). ALCOVE: An exemplar-based connectionist model of category learning. *Psychological Review*, *99*(1), 22–44.

Lawrence, D.H. (1952). The transfer of a discrimination along a continuum. *Journal of Comparative and Physiological Psychology*, *45*, 511–516.

Levy, W.B., Brassel, S.E., & Moore, S.D. (1983). Partial quantification of the associative synaptic learning rule of the dentate gyrus. *Neuroscience*, *8*(4), 799–808.

Lubow, R. (1973). Latent inhibition. *Psychological Bulletin*, *79*, 398–407.

Lubow, R.E., Rifkin, B., & Alek, M. (1976). The context effect: The relationship between stimulus pre-exposure and environmental pre-exposure determines subsequent learning. *Journal of Experimental Psychology: Animal Behavior Processes*, *2*(1), 38–47.

Lye, R.H., O'Boyle, D.J., Ramsden, R.T., & Schady, W. (1988). Effects of a unilateral cerebellar lesion on the acquisition of eye-blink conditioning in man. *Journal of Physiology (London)*, *403*, 58P.

Mackintosh, N. (1983). *Conditioning and associative learning*. Oxford: Oxford University Press.

McClelland, J., McNaughton, B., & O'Reilly, R. (1994). *Why we have complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory*. Technical Report No. PDP.CNS.94.1. Carnegie Mellon University, Pittsburgh.

Medin, D.L., Altom, M.W., Edelson, S.M., & Freko, D. (1982). Correlated symptoms and simulated medical classification. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *8*, 37–50.

Medin, D.L., & Schwanenflugel, P.J. (1981). Linear separability in classification learning. *Journal of Experimental Psychology: Human Learning and Memory*, *7*, 355–368.

Molchan, S.E., Sunderland, T., McIntosh, A.R., Hersovitch, P., & Schreus, B.G. (1994). A functional anatomical study of associative learning in humans. *Proceedings of the National Academy of Sciences of the United States of America*, *91*, 8122–8126.

Moyer, J., Deyo, R., & Disterhoft, J. (1990). Hippocampectomy disrupts trace eye-blink conditioning in rabbits. *Behavioral Neuroscience*, *104*(2), 243–252.

Murre, J. (1994). *A model for amnesia*. Unpublished manuscript.

Myers, C., & Gluck, M. (1994). Context, conditioning and hippocampal re-representation. *Behavioral Neuroscience*, *108*(5), 835–847.

Myers, C., & Gluck, M. (1996). Cortico-hippocampal representations in simultaneous odor discrimination learning: A computational interpretation of Eichenbaum, Mathews & Cohen (1989). *Behavioral Neuroscience*, *110*(4), 1–22.

Myers, C., Gluck, M., & Granger, R. (1995). Dissociation of hippocampal and entorhinal function in associative learning: A computational approach *Psychobiology*, *23*(2), 116–138.

Myers, C.E., Ermita, B., Harris, K., Hasselmo, M., Solomon, P., & Gluck, M.A. (1996). A computational model of the effects of septohippocampal disruption on classical eyeblink conditioning. *Neurobiology of Learning and Memory*, *66*, 51–66.

Nadel, L., & Willner, J. (1980). Context and conditioning: A place for space. *Physiological Psychology*, *8*(2), 218–228.

O'Keefe, J., & Nadel, L. (1978). *The hippocampus as a cognitive map*. Oxford: Clarendon University Press.

Oliver, L.M., & Gluck, M.A. (1995). A hippocampal-based model of human learning: implications for the distinction between integral and separable dimensions. Manuscript in preparation.

Olton, D. (1983). Memory functions and the hippocampus. In W. Seifert (Ed.), *Neurobiology of the Hippocampus* (pp.335–373). London: Academic Press.

Parker, D. (1985). *Learning logic*, Technical Report, Center for Computational Research in Economics and Management Science, MIT.

Pearce, J., & Hall, G. (1980). A model for pavlovian learning: Variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychological Review*, *87*, 532–552.

Penick, S., & Solomon, R. (1991). Hippocampus, context and conditioning. *Behavioral Neuroscience*, *105*(5), 611–617.

Phillips, R. & LeDoux, J. (1994). Lesions of the dorsal hippocampal formation interfere with background but not foreground contextual fear conditioning. *Learning & Memory*, *1*, 34–44.

Port, R., Mikhail, A., & Patterson, M. (1985). Differential effects of hippocampectomy on classically conditioned rabbit nictitating membrane response related to interstimulus interval. *Behavioral Neuroscience*, *99*(2), 200–208.

Port, R., Romano, A., & Patterson, M. (1986). Stimulus duration discrimination in the rabbit: Effects of hippocampectomy on discrimination and reversal learning. *Physiological Psychology*, *4*, 124–129.

Port, R., & Patterson, M. (1984). Fimbrial lesions and sensory preconditioning. *Behavioral Neuroscience*, *98*, 584–589.

Port, R., Romano, A., & Patterson, M. (1986). Stimulus duration discrimination in the rabbit: Effects of hippocampectomy on discrimination and reversal learning. *Physiological Psychology*, *4*(3&4), 124–129.

Rawlins, J. (1985). Associations across time: The hippocampus as a temporary memory store. *Behavioral and Brain Sciences*, *8*, 479–496.

Reilly, S., Harley, C., & Revusky, S. (1993). Ibotenate lesions of the hippocampus enhance latent inhibition in conditioned taste aversion and increase resistance to extinction in conditioned taste preference. *Behavioural Neuroscience*, *107*(6), 996–1004.

Rescorla, R., & Wagner, A. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and non-reinforcement. In A. Black & W. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp.64–99). New York: Appleton-Century-Crofts.

Ribot, T. (1882). *The diseases of memory*. New York: Appleton.

Riley, D. (1968). *Discrimination learning*. Boston: Allyn & Bacon.

Ross, R., Orr, W., Holland, P., & Berger, T. (1984). Hippocampectomy disrupts acquisition and retention of learned conditional responding. *Behavioral Neuroscience*, *98*(2), 211–225.

Rumelhart, D., Hinton, G., & Williams, R. (1986). Learning internal representations by error propagation. In D. Rumelhart & J. McClelland (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition* (pp.318–362). Cambridge, MA: MIT Press.

Schmajuk, N. (1994). Stimulus configuration, classical conditioning, and spatial learning: Role of the hippocampus. In *World Congress on Neural Networks*, 2 (pp.II723–II728). San Diego: INNS Press.

Schmajuk, N., & DiCarlo, J. (1990). *Backpropagation, classical conditioning and hippocampal function* (Technical Report, Northwestern University).

Schmajuk, N., & Moore, J. (1985). Real-time attentional models for classical conditioning. *Physiological Psychology*, *13*(4), 278–290.

Schmalz, L., & Theios, J. (1972). Acquisition and extinction of a classically conditioned response in hippocampectomized rabbits (*Oryctolagus cuniculus*). *Journal of Comparative and Physiological Psychology*, *79*, 328–333.

Shanks, D.R. (1991). Categorization by a connectionist network. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *17*(3), 1–11.

Shanks, D.R., & Gluck, M.A. (1994). Tests of an adaptive network model for the identification, categorization and recognition of continuous-dimension stimuli. *Connection Science*, *6*(1), 59–89.

Shepard, R. (1958). Stimulus and response generalization: Deduction of the generalization gradient from a trace model. *Psychological Review*, *65*, 242–256.

Silveira, J., & Kimble, D. (1968). Brightness discrimination and reversal in hippocampally-lesioned rats. *Physiology and Behavior*, *3*, 625–630.

Solomon, P. (1977). Role of the hippocampus in blocking and conditioned inhibition of the rabbit's nictitating membrane. *Journal of Comparative and Physiological Psychology*, *91*(2), 407–417.

Solomon, P., & Moore, J. (1975). Latent inhibition and stimulus generalization of the classically conditioned nictitating membrane response in rabbits (*Oryctolagus cuniculus*) following dorsal hippocampal ablation. *Journal of Comparative and Physiological Psychology*, *89*, 1192–1203.

Solomon, P.R., Stowe, G.T., & Pendlebury, W.W. (1989). Disrupted eyelid conditioning in a patient with damage to cerebellar afferents. *Behavioral Neuroscience*, *103*(4), 898–902.

Squire, L. (1987). *Memory and brain*. New York: Oxford University Press.

Squire, L., & Frambach, M. (1990). Cognitive skill learning in amnesia. *Psychobiology*, *18*(1), 109–117.

Stanton, P., & Sejnowski, T. (1989). Associative long-term depression in the hippocampus induced by Hebbian covariance. *Nature*, *339*, 215–218.

Sutherland, N., & Mackintosh, N. (1971). *Mechanisms of animal discrimination learning*. New York: Academic Press.

Sutherland, R., & Rudy, J. (1989). Configural association theory: The role of the hippocampal formation in learning, memory and amnesia. *Psychobiology*, *17*(2), 129–144.

Sutton, R., & Barto, A. (1981). Toward a modern theory of adaptive networks: Expectation and prediction. *Psychological Review*, *88*, 135–170.

Terrace, H. (1963). Discrimination learning with and without ''errors''. *Journal of Experimental Analysis of Behavior*, *6*(1), 1–27.

Thompson, R. (1971). Sensory preconditioning. In R. Thompson & J. Voss (Eds.), *Topics in learning and performance* (pp.105–129). New York: Academic Press.

Thompson, R. (1986). The neurobiology of learning and memory. *Science*, *233*, 941–947.

Thompson, R. (1990). Neural mechanisms of classical conditioning. *Philosophical Transactions of the Royal Society, London [Biology]*, *329* 161–170.

Topka, H., Valls-Sole, J., Massaquoi, S.G., & Hallett, M. (1993). Deficit in classical conditioning in patients with cerebellar degeneration. *Brain*, *116*, 961–969.

Weiskrantz, L., & Warrington, E.K. (1979). Conditioning in amnesic patients. *Neuropsychologia*, *17*, 187–194.

Werbos, P. (1974). *Beyond regression: New tools for prediction and analysis in the behavioral sciences.* PhD dissertation, Harvard University.

Widrow, B., & Hoff, M. (1960). Adaptive switching circuits. *Institute of Radio Engineers, Western Electronic Show and Convention Record*, *4*, 96–104.

Winocur, G. (1990). Anterograde and retrograde amnesia in rats with dorsal hippocampal or dorsomedial thalamic lesions. *Behavioral Brain Research*, *38*, 145–154.

Winocur, G., Rawlins, J., & Gray, J. (1987). The hippocampus and conditioning to contextual cues. *Behavioral Neuroscience*, *101*(5), 617–625.

Woodruff-Pak, D. (1993). Eyeblink classical conditioning in HM: Delay and trace paradigms. *Behavioral Neuroscience*, *107*(6), 911–925.

Zola-Morgan, S., & Squire, L. (1986). Memory impairment in monkeys following lesions limited to the hippocampus. *Behavioral Neuroscience*, *100*(2), 155–160.

Zola-Morgan, S., & Squire, L. (1992). Components of the medial temporal lobe memory system. In L. Squire & N. Butters (Eds.), *Neuropsychology of memory* (2nd edn.) (pp.325–335). New York: Guilford Press.

Zola-Morgan, S., Squire, L., & Amaral, D. (1989). Lesions of the amygdala that spare adjacent cortical regions do not impair memory or exacerbate the impairment following lesions of the hippocampal formation. *Journal of Neuroscience*, *9*(6), 1922–1936.

Zola-Morgan, S., Squire, L., Rempel, N., Clower, R., & Amaral, D. (1992). Enduring memory impairment in monkeys after ischemic damage to the hippocampus. *Journal of Neuroscience*, *12*(7), 2582–2596.

Copyright of Memory is the property of Psychology Press (T&F) and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.