Behavioral Neuroscience
1994, Vol. 108, No. 5, 835-847

0735-7044/94/$3.00

Context, Conditioning, and Hippocampal Rerepresentation in Animal Learning

Catherine E. Myers and Mark A. Gluck

The researchers argue that a previous computational account of hippocampal region function in associative learning (M. Gluck & C. Myers, 1993) has emergent implications that accurately describe the role of the hippocampal region in contextual processing. This article unifies 2 seemingly conflicting views of contextual processing: It accords contextual cues no special representational status (e.g., R. Rescorla & A. Wagner, 1972), yet it also allows context to stand in a superordinate relationship to the cues it contains (e.g., L. Nadel & J. Willner, 1980). As a result, the account correctly expects that context can develop occasion-setting properties and that context shifts can weaken learned responses or attenuate latent inhibition. The article also explains data suggesting that hippocampal lesions reduce contextual sensitivity. It may help unify several previous theoretical accounts of the hippocampal region's role in contextual processing.

The conditioned stimuli in a standard classical conditioning experiment generally consist of sensory cues presented for a carefully controlled temporal duration. These phasic stimuli are expected to accrue associative strength as a function of how reliably and uniquely they predict the arrival of a reinforcing (and usually response-evoking) event. At the same time, many other sensory cues are also present; these may include unrelated noise in the room, the smells and feel of the experimental apparatus, and so on. These unchanging or tonic stimuli are usually grouped together as *contextual* or *background* cues, and most conditioning methodologies try to eliminate or control for their effects. Nevertheless, experimenters have long recognized that these contextual cues can and do affect learned associations (e.g., Hull, 1943; Konorski, 1967; Pavlov, 1927).

More recently, it has become apparent that the hippocampal region plays a role in the processing of these contextual cues (e.g., Hirsh, 1974; Nadel & Willner, 1980; Penick & Solomon, 1991). In many cases, hippocampal damage appears to reduce the ability to integrate contextual information during learning (Good & Honey, 1991; Hirsh, 1974; Penick & Solomon, 1991; Winocur & Gilbert, 1984; and others). For example, in rats, hippocampal lesions interfere with the acquisition of a conditioned freezing response to the context, although there is no effect on conditioned freezing response to a tone stimulus (Phillips & LeDoux, 1992). In other experiments, however, hippocampal-lesioned rats had no difficulty in learning to respond differently in different contexts (e.g., Good & Honey, 1991), suggesting that lesioned rats can use context as a conditioned cue.

Many previous models of associative learning have assumed that contextual cues are fundamentally like explicitly conditioned cues (e.g., Mackintosh, 1975; Rescorla & Wagner, 1972; and others). Under normal circumstances, the argument goes, contextual cues are simply "tuned out" because they are not useful predictors of when a salient or reinforcing event will occur. Under circumstances where there are no better predictors, however, learning is expected to use contextual cues.

Other theorists, however, have taken a very different view of context in learning. Nadel and Willner (1980), for example, argued that context is qualitatively different from—and stands in a hierarchical relation to—the cues that compose it. Bouton and Swartzentruber (1986) have also noted that context does not always enter into direct associations, as phasic cues do, but instead may serve to modulate or "disambiguate" the meaning of phasic cues.

This article argues that these two seemingly opposing views can be reconciled within a recent computational theory of hippocampal region function in associative learning (Gluck & Myers, 1993). This theory assumes that stimuli in normal, intact animals are processed in two ways: First, stimuli enter into direct associations with any salient or reinforcing events they predict, and second, stimuli mediate the representations of any other stimuli with which they co-occur. The first, *associational* process is assumed to be hippocampal independent, whereas the second, *representational* process is assumed to depend on the hippocampal region. Therefore, the theory predicts that hippocampal lesion differentially disrupts representational—but not necessarily associational—processes. This two-process theory has been shown to account for a wide range of conditioned behaviors in both intact and hippocampal-lesioned rabbits and rats (Gluck & Myers, 1993).

In this article, we argue that this two-process theory of hippocampal region function has the emergent property that

Catherine E. Myers and Mark A. Gluck, Center for Molecular and Behavioral Neuroscience, Rutgers—The State University of New Jersey.

This research was supported by National Institute of Mental Health National Research Service Award 1-F32-MH10351-01 to Catherine E. Myers and by the Office of Naval Research through the Young Investigator Program and by Grant N00014-88-K-0112 to Mark A. Gluck.

For comments and advice, we thank Mark Bouton, James Goebel, Howard Eichenbaum, Lynn Nadel, Paul Solomon, Tim Curran, and the anonymous reviewers of this article.

Correspondence concerning this article should be addressed to Catherine E. Myers, Center for Molecular and Behavioral Neuroscience, Rutgers—The State University of New Jersey, 197 University Avenue, Newark, New Jersey 07102. Electronic mail may be sent to myers@pavlov.rutgers.edu.

contextual information can operate in the superordinate manner suggested by Nadel and Willner (1980) and Bouton and Swartzentruber (1986), without requiring any special representational status for contextual cues. In brief, the theory expects that new stimulus representations are affected by all co-occurring stimuli, tonic as well as phasic, and are therefore context sensitive. Thus, our theory accounts for the fact that contextual cues can develop occasion-setting properties, helping to "disambiguate" the meaning of phasic cues (Bouton & Swartzentruber, 1986). The theory also accounts for the disruption of learning that occurs after a change of context.

Our theory implies that, in contrast to these hippocampal-dependent representational processes, elementary associational processes are not hippocampal dependent. The theory therefore accounts for the behavioral data indicating that hippocampal-lesioned animals are less able to incorporate contextual information into learned associations (e.g., Penick & Solomon, 1991), but no less able to learn simple discriminations on the basis of context (e.g., Good & Honey, 1991).

Context and Classical Conditioning

A typical conditioning experiment involves learning responses on the basis of the presence of particular salient sensory events, such as tones, lights, and tactile cues. The contextual cues are usually defined to include all other stimuli present during the experiment, including background sights, sounds, and smells of the conditioning apparatus and experimental room. However, the division of stimuli into conditioned and contextual cues is not always simple. For example, contextual cues such as the color of a background wall may acquire conditioned strength if an animal obtains a food reward in a red, but not a blue, experimental chamber. Conversely, stimuli such as tones and lights that are explicitly conditioned stimuli in one experiment may be irrelevant in another experiment—and thus come to be regarded as part of the context or background. What is a contextual cue and what is a conditioned stimulus are neither well defined across experimental paradigms nor consistent across theoretical accounts of contextual processing.

A more objective division of stimuli is to identify them as either tonic or phasic cues (e.g., Nadel & Willner, 1980). Phasic cues, as the name suggests, are of limited duration. Tonic contextual cues are present and essentially unchanging for the duration of an experiment. Thus, one possibility for modeling contextual cues is to represent them identically to phasic conditioned cues but to present them at a different frequency.

Rescorla and Wagner's (1972) associative learning model is one example of a theoretical account of classical conditioning that treats contextual cues in this way. In their theory, cues are assumed to acquire associative strength on the basis of how reliably they predict reinforcement, and they compete with other co-occurring cues for this association. Training to respond to a phasic cue A is considered equivalent to mixing reinforced $AX+$ trials with unreinforced $X-$ trials, where X represents the set of contextual cues present during conditioning (e.g., Wagner & Rescorla, 1972). Because the contextual cues, X, are correlated with both reinforcement and nonreinforcement, they tend to accrue little or no net associative strength. Phasic cue A, however, is correlated only with reinforcement and is thus expected to acquire associative strength. Many other associative learning models make similar assumptions about the equivalent representation of phasic conditioned and tonic contextual cues (e.g., Mackintosh, 1975; Pearce & Hall, 1980; Winocur, Rawlins, & Gray, 1987; and others).

Others, however, have argued that it is insufficient to represent context in the same way as phasic cues (e.g., Nadel & Willner, 1980; see also Bouton & Swartzentruber, 1986). According to this line of reasoning, contextual cues are fundamentally different from phasic cues in that they do not enter directly into associations with the reinforcer. Rather, contextual cues are presumed to form a framework or map that can be used to predict the presence of phasic cues. This view of context as superordinate to phasic cues is related to Hirsh's (1974) proposal that context disambiguates phasic cues by indicating which of several stored associations should be retrieved. There are several examples (including the occasion setting discussed below) in which contextual cues seem to operate qualitatively differently from phasic conditioned cues.

Gluck and Myers (1993) recently presented a computational theory of hippocampal region function in which tonic contextual cues are not represented differently from phasic cues; contextual cues do, however, influence the representation of phasic cues in a manner analogous to that suggested by Nadel and Willner's argument. In the next section, we briefly review this computational theory and its account of contextual processing.

Context and Representations in the Hippocampal Region

The idea of a stimulus representation, defined as the pattern of activity evoked by a stimulus input, is central to the account of hippocampal region function presented by Gluck and Myers (1993). Learning that a particular stimulus predicts some salient or reinforcing outcome is equivalent to developing an association between the representation of that stimulus and the reinforcement.

What has been learned about one stimulus generalizes to another stimulus as a function of the similarity or overlap between their representations. If the representations of two stimuli are very similar, associations that accrue to one stimulus should generalize strongly to the other. Conversely, associations generalize only weakly to stimuli that are represented very differently. Learning can be speeded by a choice of representation that appropriately decreases the generalization between stimuli associated with dissimilar reinforcement. Learning can also be facilitated by noticing correlations in the input: If two stimuli tend to co-occur, then they will tend to make similar predictions, and the generalization between them should be increased. Gluck and Myers (1993) proposed that the hippocampal region is able to construct new stimulus representations that facilitate learning in these ways.

A simple connectionist network model can instantiate these processes (Gluck & Myers, 1993). An autoencoding network representing the associative learning capability of the hippocampal region is shown on the right-hand side of Figure 1A; given input representing any stimuli present (including contextual cues), it learns to reconstruct these inputs as well as to predict future reinforcement. An important feature of this

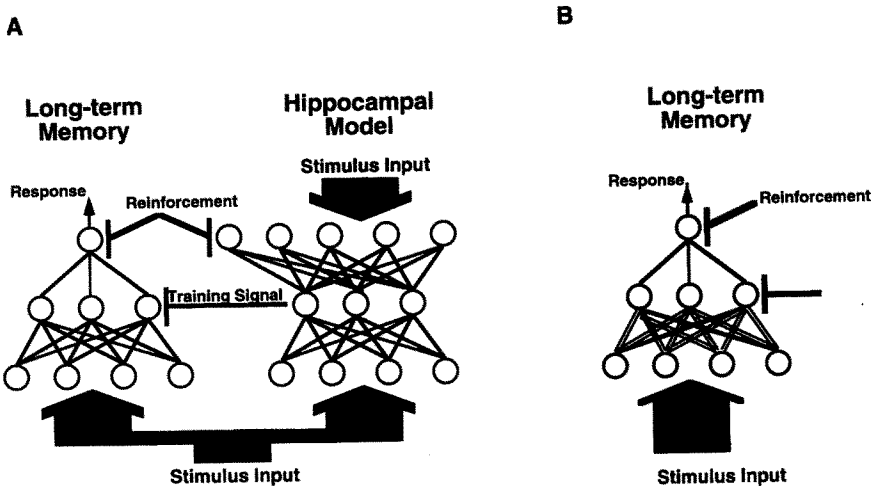

Figure 1. A: Intact cortico-hippocampal model (Gluck & Myers, 1993). The hippocampal network (see right) is a predictive autoencoder (Hinton, 1989) that learns to reproduce its stimulus inputs as well as predict future reinforcement. In the process, it develops an internal representation that incorporates information about all co-occurring stimuli including context. These representations may be adopted by other brain regions, such as the cerebellar and cerebral cortices (see left) which are assumed to be the sites of long-term memory. The output of the cortical network is interpreted as the behavioral response of the model. B: Hippocampal lesion is simulated by disabling the hippocampal network. In this case, the cortical network can no longer adopt new stimulus representations; it can, however, continue to learn behavioral responses that are based on its preexisting representations. Together, the intact and lesioned models have been shown to account for a wide range of trial-level classical conditioning data in intact and hippocampal-lesioned animals (Gluck & Myers, 1993).

network is a narrow internal layer, which forms new stimulus representations that compress redundancies while preserving predictive information. The network also performs pattern completion: Given a noisy or partial pattern, it tends to reconstruct the most similar stored pattern as its output (Hinton, 1989; see also McNaughton, 1991; Rolls, 1990). This pattern completion allows the presence of one cue to predict the presence of another that has reliably co-occurred in the past.

A second network, shown on the left-hand side of Figure 1A, represents a highly simplified model of some aspects of conditioned learning in cerebellar and cerebral cortices (e.g., see Gluck, Myers, & Thompson, in press). This cortical network can adapt its lower layer of weights to acquire the new representations constructed by the hippocampal model. The cortical network can also modify its upper layer of weights to learn associations between these representations and a prediction of future reinforcement. In modeling a conditioning task, this prediction is assumed to be related to a behavioral output, such as an anticipatory reflex response.

Hippocampal lesion is simulated in this model by disabling the hippocampal network, as shown in Figure 1B. In this case, the cortical network can still modify its upper layer of weights to form new associations on the basis of its existing internal representations. However, in the absence of hippocampal information, the lower layer of cortical network weights is fixed, and no new representations can be acquired.

Simulations with this model have been shown to exhibit accurately a wide range of conditioned behaviors in intact rats and rabbits as well as rats and rabbits with broad hippocampal region damage (Gluck & Myers, 1993). Among other behaviors, the intact but unlesioned model correctly exhibits successive reversal facilitation, latent inhibition, and sensory preconditioning. The model also makes the novel predictions that easy–hard transfer and compound preconditioning should be disrupted by hippocampal region damage. Full implementation details of the model, which is used in the simulations reported in this article, are given in the Appendix.

Context sensitivity emerges as a natural consequence of the basic assumptions that govern learning in the intact model. The complete input to the hippocampal net is assumed to include any phasic conditioned cues, as well as a series of tonic background cues. These tonic cues change slowly with time, but are static compared with the frequent onset and offset of phasic cues. Every block of training includes one presentation of each phasic input (or combination of phasic inputs) interspersed among large numbers of presentations of the background tonic stimuli alone. These background presentations represent intervals during which the animal remains in the conditioning apparatus, but no particular phasic cue is presented. Aside from this difference in the baseline presentation rate, tonic contextual cues are represented in exactly the same way as phasic cues.

The model also includes a larger sense of context that is superordinate to and predictive of the cues that occur within it. The stimulus representations constructed in the hippocampal region are assumed to be influenced by all other co-occurring stimuli. Thus, the presence of a particular set of stimuli may suffice to allow the pattern completion mechanisms to generate an expectation of other cues that have consistently cooccurred in the past. In this way, a set of contextual cues provides a framework with which to predict the arrival of phasic cues.

This use of context as a framework is qualitatively different from learning which individual cues or cue compounds predict reinforcement, as suggested by Nadel and Willner (1980). However, as Rescorla and Wagner (1972) argued, no special representation of context is assumed. Therefore, this theory provides a way to reconcile these two seemingly opposing views on how context should be represented.

In the remainder of this article, we concentrate on how this computational account of hippocampal region function can provide an interpretation of three broad classes of contextual processing phenomena: occasion setting, context shifts, and the decreased sensitivity to context generally observed in hippocampal-lesioned animals.

Occasion-Setting Properties of Context

One important contextual processing phenomenon involves the ability of context to signal or retrieve associations between phasic cues without itself entering into direct associations (e.g., Bouton & Swartzentruber, 1986). This property is often called occasion setting and is developed by contextual cues, but not usually phasic cues, during a conditional discrimination (e.g., Holland, 1989a, 1989b). The intact model can account for this phenomenon, as shown below. Because the model assumes contextual and phasic cues are not represented differently, the simple assumption of different presentation rates for tonic and phasic cues is sufficient to generate the occasion-setting phenomenon.

A *conditional discrimination* involves learning a response to a phasic cue A that is conditional on the presence of one or more other cues. For example, in a *feature-positive* conditional discrimination, A predicts reinforcement only when paired with a second phasic cue, B ($BA+$ and $A-$). This task is usually solved by development of direct associations between B and the reinforcement, although A may acquire little or no association (Holland, 1989a). Alternatively, in a *feature-negative* conditional discrimination, A predicts reinforcement unless paired with a second phasic cue C ($CA-$ and $A+$). Here, C usually acquires strong negative associations with reinforcement (Holland, 1989b).

However, the situation may be very different if the feature cues B and C precede A in time ($B \rightarrow A+$ and $A-$ or $C \rightarrow A-$ and $A+$). In this case, the feature cues B and C may not acquire direct associations with the reinforcement. Instead, A may acquire associations with both reinforcement and nonreinforcement, and the presence of B or C signals which association should be retrieved (Holland, 1989a, 1989b). In this case, the feature cues B and C are termed *occasion setters.* The development of occasion-setting properties is assumed if there is no evidence that the feature cues have entered into any direct associations with reinforcement or with other cues.

Bouton and Swartzentruber (1986) noted the parallel between occasion setting and context; in particular, contextual cues tend to precede the arrival of phasic cues, in the same way that feature cues precede target cues in an occasion-setting task. They conducted a series of experiments demonstrating that contextual cues develop occasion-setting properties. First, an animal was trained to respond to a phasic cue A in one context, X, but not in a second context, Y ($AX+$, $X-$, $AY-$, and $Y-$); this task contains both feature-positive and feature-negative aspects. Bouton and Swartzentruber noted that there are at least four ways in which this task can be solved: (a) X can become a conditioned excitor, (b) a response can be learned to the configuration AX, (c) Y can become a conditioned inhibitor, and (d) X or Y or both can become occasion setters. These mechanisms are not all mutually exclusive; however, only the first three involve direct associations between context and reinforcement. Bouton and Swartzentruber found no evidence for such direct associations and so concluded that the contexts had instead become occasion setters.

The intact cortico-hippocampal model can be trained on an analogous conditional discrimination ($AX+$, $X-$, $AY-$, and $Y-$), and quickly learns to respond only to A in context X (see Figure 2A). Next, it is trained on a second conditional discrimination to test which of the putative mechanisms it used in solving the original task. As outlined below, there is no evidence of direct associations between context and reinforcement, implying that the intact model, like animals, uses contextual occasion setting.

The first mechanism suggested by Bouton and Swartzentruber (1986) is for context X to become a conditioned excitor. This is clearly not the case in the intact model, as during original training, there is little or no response to $X-$ in the absence of phasic cue A (see Figure 2A). Also, the model does not appear to learn to respond to the configuration AX. If it did, the configuration AX would accrue associations with reinforcement, whereas there would be little or no learning to the individual components A and X. This can be tested by a transfer task involving learning a new discrimination in the previously positive context X ($EX+$ and $X-$). If all learning involved the configuration AX, there should be little or no transfer, and this task should be no faster than learning in a completely novel context Z ($EZ+$ and $Z-$) (cf. Bouton & Swartzentruber, 1986). However, this is not the case; the model shows significant facilitation in the previously positive context (see Figure 2B). Thus, it does not appear that context X has acquired direct association to reinforcement, either alone or in configuration with A.

Another possible mechanism for solving the original task is for context Y to become a conditioned inhibitor: The model would learn a positive association to A, but stronger negative associations to Y, so that the net response to the compound AY would be low. One way to test whether conditioned inhibition has occurred is by pairing the putative inhibitor with a novel phasic cue, E, in a transfer task ($EY+$ and $Y-$). If Y has acquired inhibitory strength, then this task should be slowed, as E must acquire still stronger excitatory strength to allow responding to the compound EY (cf. Bouton & Swartzentruber, 1986). However, there is no such slowing in the model: Learning in the previously negative context Y is not retarded compared with learning a parallel task in a novel context Z ($EZ+$ and $Z-$; see Figure 2C). This indicates that Y has not acquired direct inhibitory associations with the reinforcement.

Together, these results indicate that the model does not learn the original discrimination ($AX+$, $X-$, $AY-$, and $Y-$) by forming direct associations involving the contexts X and Y. Therefore, by Bouton and Swartzentruber's (1986) argument, the results imply that contexts operate as occasion setters in the intact model.

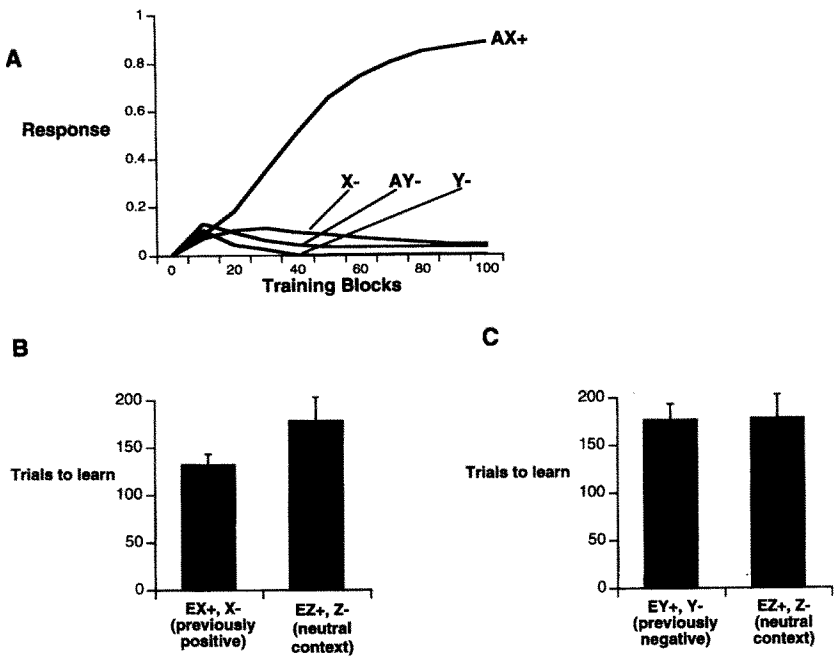

Figure 2. Occasion setting by contextual cues in the intact cortico-hippocampal model. A: A conditional discrimination in which the correct response to a phasic cue A is determined by the context ($AX+$, $X-$, $AY-$, and $Y-$). This task is not learned by excitatory associations to X, because there is little or no response to X in the absence of A. B: If the task were learned by acquiring a response to the configuration AX, this would imply little learning to the components A and X. However, learning a new discrimination in the previously positive context ($EX+$ and $X-$) is facilitated relative to learning in a novel context Z ($EZ+$ and $Z-$). This implies there was learning about context X and not merely to the configuration AX. C: Finally, there is no evidence that the original conditional discrimination was solved through the acquisition of negative associations to Y. This would imply slower learning of a new discrimination in the previously negative context Y ($EY+$ and $Y-$) compared with learning in a novel context Z ($EZ+$ and $Z-$). However, there is no such slowing, implying that Y did not accrue negative associations. Together, these results can be taken to imply that the contexts X and Y do not enter directly into associations but instead function as occasion setters. By similar arguments, contextual cues can be shown to function as occasion setters in normal animals (Bouton & Swartzentruber, 1986).

Contextual Shift Effects

Another emergent property of the cortico-hippocampal model is an effect of context shift on learned associations. The model includes contextual information in stimulus representations. This implies that there may be a drop in the learned response to a stimulus if that stimulus is presented in a novel context: The stimulus representation will be less fully evoked, and in turn the response will be weakened. Figure 3A shows this effect in the model. Normal intact rabbits and rats often show a similar drop in response to a trained stimulus when it is presented in a novel context (Antelman & Brown, 1972; Good & Honey, 1991; Honey, Willis, & Hall, 1990; Kim & Fanselow, 1992; Penick & Solomon, 1991).

However, this response decrement is not always obtained (Bouton & Brooks, 1993; Bouton & Peck, 1989; Hall & Honey, 1990; Winocur & Olds, 1978). There have been several suggestions as to why the response decrement should fail to appear under certain experimental conditions. One suggestion is simply that different paradigms are differentially dependent on the context (Bouton, 1993; Good & Honey, 1991); different paradigms may therefore show more or less effect of a context shift. Another suggestion is that contextual dependence is increased when phasic cues are low in salience, infrequently shown, or incompletely reliable as predictors of reinforcement arrival (Winocur & Gilbert, 1984).

Our theory suggests a third and possibly additional reason why response decrements might sometimes not be observed: Context dependence is expected to be a function of training time. Contextual information is included in a stimulus representation because of the model's constraint to compress the representations of co-occurring cues. Because the phasic stimulus always occurs in the same context, the representations of cue and context become compressed together. However, compression is not the only bias constraining developing representations; there is also a constraint to differentiate representations of stimuli that are useful predictors of reinforcement. In this task, the presence of the phasic cue is the best predictor of reinforcement, and so its representation will gradually become differentiated from all other representations. As a result, after extensive training, the representation of that phasic stimulus may become less dependent on the context. Our intact model predicts that the response decre-

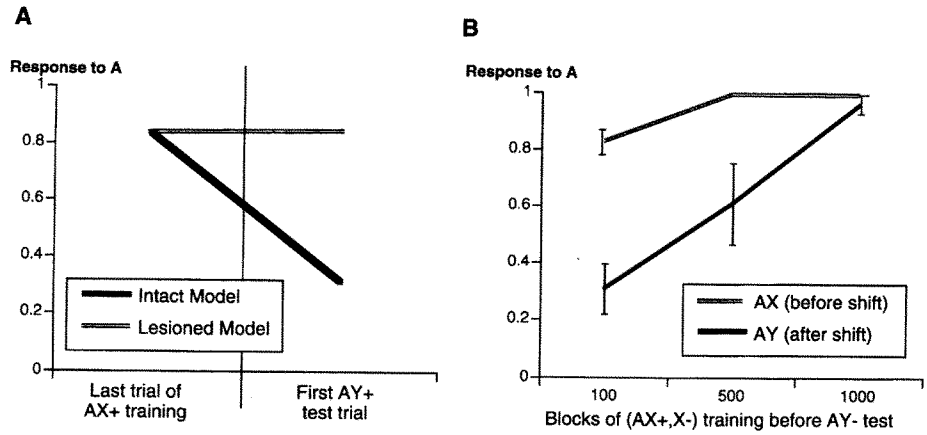

Figure 3. A: In the intact model, a strong response to cue A in context X is learned within 100 blocks of ($AX+$ and $X-$) training; that response is weakened if A is presented in a novel Context Y. This ocucrs because the context of learning is included in the representation of A; the new context Y does not activate this representation so strongly, and so the response is in turn weakly activated. In the lesioned model, context is simply tuned out of the association, and so there is no effect of a context shift. Likewise, hippocampal-lesioned animals do not show a response decrement after context shift (Antelman & Brown, 1972; Penick & Solomon, 1991). B: With further training, representations in the intact model become more context independent. This predicts that the response decrement after context shift will be eliminated after overtraining of the response. There is some evidence that this occurs in intact rats as well (Hall & Honey, 1990).

ment with context shift should be lessened or eliminated after extensive training in the original context (see Figure 3B).

In animal paradigms, different kinds of discrimination are learned at widely varying rates (Eichenbaum, 1992). Because we predict that the response decrement effect occurs only transiently, the exact number of trials during which it occurs is dependent on the paradigm. For example, rat conditioned suppression is a fast learning system, in which a response is acquired within a few trials (e.g., Bouton & King, 1983). In this paradigm, the response decrement effect appears if context is shifted after a single training trial (Hall & Honey, 1990), whereas with 15–23 training trials, the effect is not found (Bouton & King, 1983). In a slow learning paradigm, such as conditioning of the rabbit nictitating membrane response, the response decrement effect is seen if the context is shifted after more than 600 trials (Penick & Solomon, 1991); our account predicts that after even more extensive overtraining, the effect should disappear. Table 1 summarizes these and other representative data and predictions; although context shift has not been tested after overtraining in a variety of paradigms, we are aware of no data that explicitly contradict our prediction that the response decrement will occur if context is shifted early in training but not late in training.

If this interpretation of context shift effects is correct, then it should apply to other paradigms in which performance changes after contextual shift. One such paradigm is latent inhibition (Lubow, 1973). In latent inhibition, a phase of unreinforced preexposure to a phasic stimulus A ($AX-$ and $X-$) retards later acquisition of a response to the same A stimulus ($AX-$ and $X-$). The intact cortico-hippocampal model shows this effect (see Figure 4). In the model, preexposure results in compression of the representations of A and X, because they often cooccur; neither is predictive of any future reinforcing event, so there is no opponent process serving to differentiate their representations. In the subsequent acquisition phase, the representations of A and X must be differentiated to allow learning to associate a response with A but not with X alone. This will make acquisition slower than in a control condition where there was no preexposure and hence no compression of the representations of A and X.

Because latent inhibition in the model is caused by compression of the representations of A and a particular set of contextual cues, X, the effect will be context sensitive; no impairment on training is expected in a different context, Y. Therefore, our model expects a release from latent inhibition after a context shift (see Figure 4). Likewise, a context shift between preexposure and acquisition phases abolishes latent inhibition in normal rats and rabbits (Bouton & Brooks, 1993; Hall & Honey, 1989; Lubow, Rifkin, & Alek, 1976).

In a similar fashion, our model expects that latent inhibition will be abolished by a rest period of context-alone presentations between the preexposure and training phases. During the rest period, stimulus–stimulus learning continues, as the hippocampal network develops new representations of the contextual cues that do not predict A. The effect is to separate the representations of the context and A, and so training to respond to A will proceed at a normal rate (see Figure 4). Likewise, animals show release from latent inhibition given context-alone trials intervening between preexposure and training phases (Hall & Minor, 1984; Kraemer & Roberts, 1984; McIntosh & Tarpy, 1977).

As a final note, in the same way that overtraining eliminates the response decrement effect in our intact model, overexposure also eliminates the latent inhibition effect. During the $AX-$ preexposure phase, neither cue A nor the context X predicts any reinforcement, and so there is a tendency to compress their representations together. However, the hippocampal network is trying not only to predict reinforcement but

Table 1
Summary of Representative Behavioral Experiments on the Phenomenon of Response Decrement With Context Shift

Training trials	Effect of context shift	Reference
	Fast learning: Rat conditioned suppression	
1	Decrement	Hall & Honey (1990), Experiments 1 & 2
15	No effect	Bouton & King (1983), Experiment 3
23	No effect	Bouton & King (1983), Experiments 1 & 4
24	No effect	Hall & Honey (1990), Experiment 3
96	No effect	Hall & Honey (1989), Experiment 3
	Moderate learning: Rat conditioned approach to food	
48	Decrement	Honey & Good (1993), Experiment 1
96	Decrement	Hall & Honey (1989), Experiments 2 & 4
Overtraining	Prediction: no effect	No empirical data
	Moderate learning: Pigeon instrumental conditioning	
320	Decrement	Honey, Willis, Hall (1990), Experiments 1 & 2
Overtraining	Prediction: no effect	No empirical data
	Slow learning: Rabbit nictitating membrane response	
≥600	Decrement	Penick & Solomon (1991)
Overtraining	Prediction: no effect	No empirical data

Note. One explanation, suggested by the cortico-hippocampal model and consistent with these experiments, is that context sensitivity decreases with extended training and hence overtraining should eliminate the effect of response decrement after context shift. In rat conditioned suppression, the context-shift decrement is found only with limited training. In other, slower paradigms, a decrement is usually found after context shift with limited training; the effects of context shift after overtraining remain to be tested, but this explanation predicts that overtraining should eventually eliminate the context-shift decrement effect. The experiments noted here are representative rather than exhaustive, but we are aware of no data that explicitly contradict this explanation.

also to reproduce its inputs. This means that on $AX-$ trials the hippocampal network should reproduce A in its output, whereas on $X-$ trials A should not be reproduced. Therefore, although the presence of A among the inputs may not signal any particular reinforcement, it is a very important predictor of desired hippocampal network outputs. For this reason, there is a competing bias to differentiate the representations of A and the context. With extensive preexposure, the compression may be undone, and therefore latent inhibition will be abolished. The expectation that extensive preexposure eliminates latent inhibition is, to the best of our knowledge, a novel prediction that remains to be tested in intact animals. However, it has long been known that prolonged exposure to stimuli facilitates later discrimination rather than retarding it (e.g., Gibson & Walk, 1956).

Contextual Processing and Hippocampal Lesion

The model assumes that the hippocampal region is required to form new stimulus representations that include contextual information. Hippocampal lesion is assumed to eliminate this ability. However, hippocampal lesion is not expected to impair simple associations between stimuli and reinforcement. For example, Figure 5A shows that the lesioned model exhibits no particular impairment in learning to respond to a phasic cue A in a particular context. Hippocampal-lesioned animals similarly show no particular deficit on simple discrimination learning (e.g., Eichenbaum, Fagan, & Cohen, 1986; Port, Romano, & Patterson, 1986). On the other hand, effects such as latent inhibition, which the model attributes to representational changes, are absent in animals with hippocampal-region damage (Kaye & Pearce, 1987; Solomon & Moore, 1975). In the lesioned model, there is no stimulus–stimulus learning; all learning is driven by the error between the actual and predicted reinforcement. During the preexposure phase, there is no reinforcement (and none is particularly predicted), and so there is no learning. Thus, unreinforced preexposure does not have any effect on later learning, and there is no latent inhibition (see Figure 5B).

Because contextual cues are represented in the same way as phasic cues, hippocampal lesion is not expected to impair contextual learning selectively. Rather, learning about contextual cues is expected to be disrupted in the same way as learning about phasic cues. Therefore, because hippocampal lesion is not assumed to disrupt learning to discriminate two phasic cues, neither is there an expected impairment of learning to respond in one context, X, but not in a second context, Y. Given training with occasional reinforcement in context X but never in context Y ($X+$, $X-$, and $Y-$ trials intermixed), both intact and lesioned systems learn not to respond in Y but to respond in context X (see Figure 5C). Because reinforcement is sporadic in context X, the response strength never reaches the level it does, for example, in Figure 5A, where reinforcement can be perfectly predicted. Likewise,

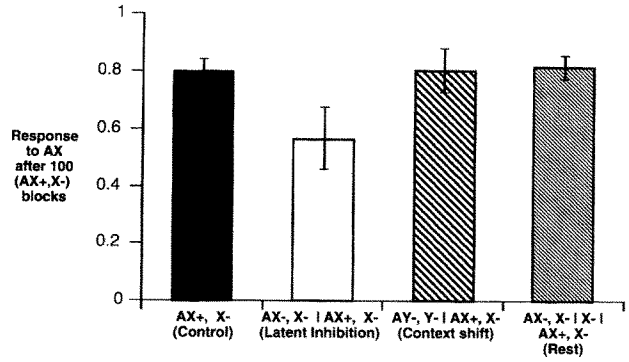

Figure 4. Latent inhibition effects in the intact model. In the control condition, the model shows a strong response to cue A after 100 ($AX+$ and $X-$) training blocks. Latent inhibition is seen if the training is preceded by nonreinforced preexposure to A; so, given 100 ($AX-$ and $X-$) blocks followed by 100 ($AX+$ and $X-$) blocks, the response to A is weaker than in the control condition. Because the model attributes latent inhibition to redundancy compression between A and the context, a context switch between preexposure and training phases should eliminate it. In fact, if 100 ($AY-$ and $Y-$) blocks are followed by 100 blocks in a new context ($AX+$ and $X-$), the response to A is as strong as in the control condition, and there is no latent inhibition. Finally, the model predicts that latent inhibition will also be eliminated by a rest interval between preexposure and training phases. Thus, if 200 context-alone ($X-$) blocks intervene between preexposure and training phases, the response to A after training is as strong as in the control condition. Normal animals show a similar latent inhibition effect (Lubow, 1973), which can be eliminated by a context shift (Hall & Honey, 1989; Lubow, Rifkin, & Alek, 1976). The elimination of latent inhibition by an intervening rest period is a novel prediction of the model that remains to be tested in animals.

hippocampal-lesioned animals can also learn to discriminate contexts (Good & Honey, 1991).

However, just because a task can still be learned after hippocampal lesion does not necessarily mean that it is learned in the same way. During learning of a simple response to a phasic cue A, the intact model includes contextual information in a new representation of A. Therefore, a response decrement is expected if A is presented in a new context. The lesioned model, by contrast, simply develops an association between A and the reinforcement. There is little learning about the contextual cues that are equally present during reinforced and nonreinforced trials; as a result, contextual cues do not influence responding enough to affect the response to A in a new context (see Figure 5D). Behavioral data are consistent with this expectation: Hippocampal-lesioned animals do continue responding at previous levels after a context shift (e.g., Antelman & Brown, 1972; Penick & Solomon, 1991). Similarly, normal animals receiving tone–shock pairings show conditioned fear to the context as well as the tone; hippocampal-lesioned animals appear to become conditioned to the tone but not the context (Phillips & LeDoux, 1992). These results suggest that hippocampal-lesioned animals are less sensitive than normal animals to contextual information.

However, it should be noted that other experiments have suggested exactly the oppositive conclusion. One such result has been obtained in the case of relearning: A response to a phasic cue A is trained, and then after a few days' interval, the response to A is tested again. Both normal intact animals and hippocampal-lesioned animals show good retention of the response; if, however, A is tested instead in a new context, the hippocampal-lesioned animals show much worse retention than the normals (Winocur & Olds, 1978). In a similar experiment, the usual impairment on reversal learning in hippocampal-lesioned animals was greatly alleviated by a context switch between tasks; the normal animals showed only a small effect of a context shift (Winocur & Olds, 1978). These results are in stark contrast to the usual empirical finding that hippocampal-lesioned animals are less sensitive to contextual changes than normal animals are.

One factor that may help explain these results is the magnitude of the contextual change. In most context-shift experiments, a contextual change is effected by alterations in the location and visual, tactile, and olfactory characteristics of the experimental apparatus (e.g., Good & Honey, 1991; Penick & Solomon, 1991; and others). Winocur and Olds (1978) went further: Their context shifts involved both a change of experimental apparatus (alternating a Yerkes–Thompson chamber and a rectangular box) and of the task itself (alternating the goals of obtaining food and water)—although the conditioned stimuli were identical in each case. The paradigm may therefore have involved switching to a novel task, rather than merely shifting context, especially because their normal subjects did not show the expected context-shift effects either. Nevertheless, these results do stand in contrast to the studies discussed above, and it remains to be seen if similar results can be obtained with lesioned animals using a more conventional context shift.

Summary and Discussion

We previously presented a theory of hippocampal involvement in constructing new stimulus representations during learning (Gluck & Myers, 1993). The basis of this idea is that the hippocampal region is required to form new stimulus representations during learning. In this article, we have argued that this same proposed hippocampal region function can account for a broad range of data regarding contextual learning and can resolve a seeming contradiction between the view that contextual cues are represented no differently than explicitly conditioned cues and the observation that context does sometimes function differently from the cues that compose it.

In our cortico-hippocampal model, contextual cues are represented in the same way as all other cues. However, hippocampal-mediated stimulus representations are assumed to include information about the context in which an association is learned. Thus, contextual information can influence the association retrieved to a stimulus, without itself entering into direct associations with reinforcement. This account is sufficient to account for the development of occasion-setting properties by contextual but not phasic cues, as well as a range of data regarding the effects of contextual shifts on learned discriminations and on latent inhibition.

The cortico-hippocampal model also assumes that hippocampal lesion impairs the ability to form new stimulus representations that include contextual information. This correctly implies that there should be no specific deficit for contextual

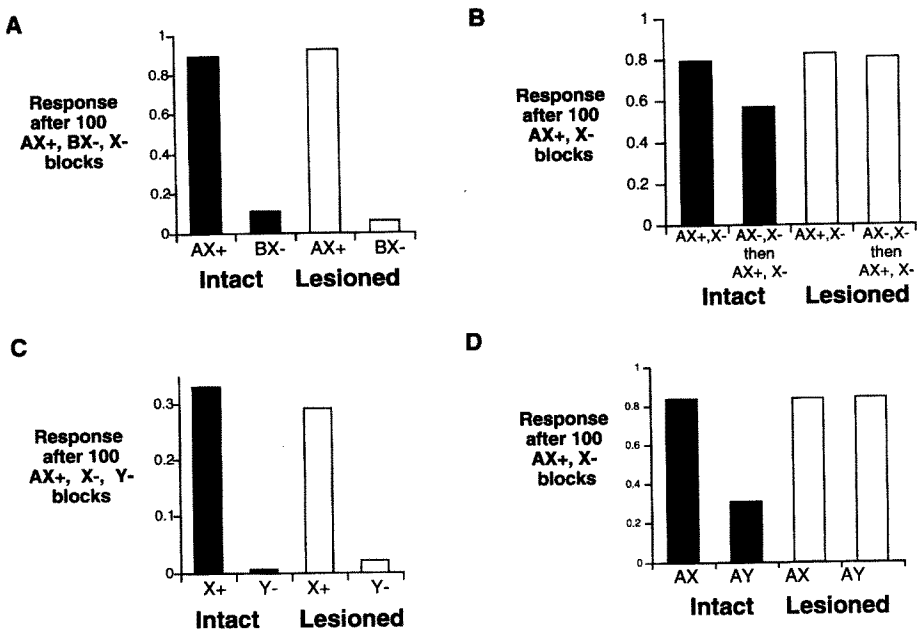

Figure 5. A: Both the intact and lesioned models can learn a stimulus discrimination with phasic cues ($AX+$, $BX-$, and $X-$) within 100 training blocks. Similarly, hippocampal-lesioned animals do not show impairment at discriminations involving phasic cues (Eichenbaum, Fagan, & Cohen, 1986; Port, Romano, & Patterson, 1986). B: However, latent inhibition, which arises because of redundancy compression in the intact model, is eliminated in the lesioned model: 100 unreinforced ($AX-$ and $X-$) blocks do not slow subsequent ($AX+$ and $X-$) learning. Hippocampal-lesioned animals likewise do not show latent inhibition (Solomon & Moore, 1975). C: Because the model postulates no specific hippocampal-lesion deficit for discrimination learning, learning to respond in one context but not another ($X+$ and $Y-$) is not disrupted in the lesioned model. Similarly, hippocampal-lesioned animals do not show impairment at discriminations involving phasic cues or tonic contextual cues (Good & Honey, 1991). D: Because the lesioned model cannot construct stimulus representations that include contextual information, a learned response is not expected to weaken when the stimulus is presented in a new context. Thus, after 100 blocks of training to respond to cue A in context X, the lesioned model does not show any response decrement when A is presented in a novel context Y. Similarly, hippocampal-lesioned animals do not show a response decrement after context shift (Penick & Solomon, 1991).

learning after hippocampal lesion where new representations are not required; however, learned associations may be less sensitive to contextual influences, as many experimental data also suggest.

Comparison With Other Theories

Our account of contextual processing can be compared with several previous theories regarding the role of the hippocampal region in contextual learning. In this section, we review the relationship between our account and previous characterizations of a hippocampal role in contextual retrieval (e.g., Hirsh, 1974), cognitive mapping (e.g., O'Keefe & Nadel, 1978), and configural association (e.g., Sutherland & Rudy, 1989).

Contextual retrieval. The central feature of our intact model is that, although the hippocampus is required for constructing new stimulus representations that include contextual information, it is not required for learning simple stimulus–response associations involving either tonic or phasic cues. This implication of our theory is consistent with the ideas of Hirsh (1974) that there are two distinct systems presumed to underlie memory in the intact animal. One, the performance line system, is assumed to store physiologically observable sequences initiated by a particular stimulus, and result in a specific response (or outcome). A second, the contextual retrieval system, allows information to be retrieved by cues that refer to, but are not necessarily contained in, that information. In other words, a stimulus can be used to retrieve the context in which it occurred, or vice versa. Hirsh identified the hippocampus with the contextual retrieval system and noted that hippocampal-lesioned animals show only behavior consistent with the performance line system.

Hirsh's performance line system is analogous to our lesioned model, which is capable only of stimulus-reinforcement learning. The contextual retrieval system Hirsh proposed is analogous to our intact hippocampal-region model, which can use context to reconstruct other stimuli that are predicted to cooccur. In this way, our model can be viewed as an instantiation and an extension of Hirsh's ideas.

Cognitive mapping. Nadel and colleagues have argued for the existence of multiple memory systems within which the hippocampal-dependent memory system is responsible for cognitive mapping functions (Nadel, 1992; Nadel & Willner,

1980; O'Keefe & Nadel, 1978), which associate objects and events with the framework (or context) in which they occurred. They have also argued that spatial memories are the most important part of this cognitive mapping system, drawing on a large body of work documenting impairment in spatial and navigational tasks after hippocampal damage (e.g., Morris, Garrud, Rawlins, & O'Keefe, 1982; Nadel, 1991; O'Keefe, 1983), as well as neurophysiological evidence for hippocampal involvement in spatial learning (McNaughton, Chen, & Markus, 1991; O'Keefe, 1979).

We (Gluck & Myers, 1993) and others (Eichenbaum, Stewart, & Morris, 1990; Taube, 1991) have instead suggested that a place is simply one specialized form of a context, consisting of a configuration of local views of space. Thus, we argue that spatial impairment after hippocampal lesion may be only a task-specific effect of disrupted contextual processes—even though in some predominantly spatial species, such as rats, it may be an especially salient effect. This view is consistent with the fact that some of the same hippocampal cells that show spatially determined responses during a spatial task can have other behavioral correlates during a nonspatial task (Kubie & Ranck, 1983). Thus, we expect that hippocampal representations should come to encode significant aspects of the context—whether spatial or nonspatial.

Configurality and context. Other theories have focused on a presumed hippocampal role in learning configurations (e.g., Sutherland & Rudy, 1989). Such configurations are necessary for tasks such as $A+$, $B+$, and $AB-$, the negative patterning task, in which the response to a cue compound is different from the response to its components. In some cases, the hippocampus has been suggested to be critically necessary for forming cue configurations during learning (Schmajuk & DiCarlo, 1992; Sutherland & Rudy, 1989). These same processes can be used to explain the hippocampal role in contextual learning, if contextual learning is assumed to require the construction of a configural element that contains information about both phasic and tonic contextual cues. Because this configural association is assumed to be hippocampal dependent, it is disrupted by hippocampal lesion. In this respect, these theories correctly expect that hippocampal-lesioned animals will show difficulty in incorporating contextual information during learning.

However, there are at least two difficulties with these attempts to derive contextual effects as a consequence of a configural deficit in hippocampal-lesioned animals. The first problem is that there are several recent experiments that suggest that configural learning is not always disrupted after localized hippocampal lesion (Gallagher & Holland, 1992; Jarrard, 1993; Saunders & Weiskrantz, 1989; Whishaw & Tomie, 1991). The second problem is that some contextual processing tasks, such as the conditional discrimination described in a previous section, seem to be solved not by configural approaches but by contextual occasion setting.

Our approach, like that of several others (e.g., Eichenbaum, 1994), proposes a slightly different interpretation of configural learning. We argue that configural associations are especially likely to require the kind of representational processing we propose as a hippocampal function, and therefore they are expected to be particularly sensitive to hippocampal lesion. However, our lesioned model does have existing stimulus representations, even though it can acquire no new ones, and therefore there is always some probability that these representations will suffice for learning a particular configural association. This is expected to happen rarely, but the possibility implies that there will be cases in which a hippocampal-lesioned animal can solve a configural association. Some empirical support for this prediction exists: For example, although Whishaw and Tomie (1991) found hippocampal-lesioned rats to be generally impaired at a configural discrimination, a few of their lesioned rats solved the task as well as, or better than, controls.

Future Issues

We have argued in this article that the hippocampal role in contextual processing can be derived as an emergent implication of our more general theory of hippocampal function in conditioned learning (Gluck & Myers, 1993). Although we have shown that this approach unifies a broad body of empirical data, our model is limited to trial-level classical conditioning. It cannot simulate many temporal aspects of a training trial, such as the order in which stimuli are presented. A related issue is the consolidation period for hippocampal-dependent memories, during which hippocampal lesion causes retrograde amnesia (e.g., Squire, 1987). In our modeling to date, we have not specified biological mechanisms by which hippocampal-mediated representations could be transferred to long-term storage in the cerebral and cerebellar cortices. Although a consolidation period for this transfer could certainly be assumed, the theory does not yet incorporate assumptions to account for this phenomenon.

Because our model cannot yet simulate one-trial learning, it cannot directly address the declarative or episodic memory formation that is disrupted in humans after hippocampal damage (e.g., Squire, 1987). However, the characterization of human hippocampal-damaged amnesics as unable to form new declarative memories seems qualitatively consistent with the account of hippocampal contextual processing presented here. In particular, declarative memories include information about the spatial and temporal context in which an event occurred; hippocampal-damaged amnesics are strongly impaired at retrieval of this contextual information, even though they may have acquired the information itself (Haist, Musen, & Squire, 1991; Weiskrantz & Warrington, 1979). Conversely, the procedural or implicit memories that can still be acquired after hippocampal lesion tend to be acquired over many exposures and are less strongly associated with a particular learning context. Thus, the general idea that the hippocampal region is required to include contextual information in stimulus representations may also be relevant to understanding contextual processing deficits in human hippocampal-damaged amnesics. However, further theoretical and experimental research is required to understand the relation between hippocampal-region processing and human contextual processing.

References

Antelman, S., & Brown, T. (1972). Hippocampal lesions and shuttlebox avoidance behavior: A fear hypothesis. *Physiology and Behavior, 9,* 15–20.

Bouton, M. (1993). Context, time and memory retrieval in the interference paradigms of Pavlovian learning. *Psychological Bulletin, 114, 80–99.*

Bouton, M., & Brooks, D. (1993). Time and context effects on performance in a Pavlovian discrimination reversal. *Journal of Experimental Psychology: Animal Behavior Processes, 19,* 165–179.

Bouton, M., & King, D. (1983). Contextual control of the extinction of conditioned fear: Tests for the associative value of the context. *Journal of Experimental Psychology: Animal Behavior Processes, 9,* 248–265.

Bouton, M., & Peck, C. (1989). Context effects on conditioning, extinction and reinstatement in an appetitive conditioning paradigm. *Animal Learning and Behavior, 17,* 188–198.

Bouton, M., & Swartzentruber, D. (1986). Analysis of the associative and occasion-setting properties of contexts participating in a Pavlovian discrimination. *Journal of Experimental Psychology: Animal Behavior Processes, 12,* 333–350.

Eichenbaum, H. (1992). The hippocampal system and declarative memory in animals. *Journal of Cognitive Neuroscience, 4,* 217–231.

Eichenbaum, H. (1994). The hippocampal system and declarative memory in humans and animals: Experimental analysis and historical origins. In D. L. Schacter & E. Tulving (Eds.), *Memory systems 1994* (pp. 147–201). Cambridge, MA: MIT 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,* 1876–1884.

Eichenbaum, H., Stewart, C., & Morris, R. G. M. (1990). Hippocampal representation in place learning. *Journal of Neuroscience, 10,* 3531–3542.

Gallagher, M., & Holland, P. C. (1992). Preserved configural learning and spatial learning impairment in rats with hippocampal damage. *Hippocampus, 2,* 81–88.

Gibson, E. J., & Walk, R. D. (1956). The effect of prolonged exposure to visually presented patterns on learning to discriminate them. *Journal of Comparative and Physiological Psychology, 49,* 239–242.

Gluck, M., & Myers, C. (1993). Hippocampal mediation of stimulus representation: A computational theory. *Hippocampus, 3,* 491–516.

Gluck, M., Myers, C., & Thompson, R. (in press). A computational model of the cerebellum and motor-reflex conditioning. In S. Zournetzer, J. Davis, T. McKenna, & C. Lau (Eds.), *An introduction to neural and electronic networks* (2nd ed.) San Diego, CA: Academic Press.

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

Haist, F., Musen, G., & Squire, L. R. (1991). Intact priming of words and non-words in amnesia. *Psychobiology, 19,* 275–285.

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,* 232–241.

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

Hall, G., & Minor, H. (1984). A search for context-stimulus associations in latent inhibition. *Quarterly Journal of Experimental Psychology, 36B,* 145–169.

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.

Holland, P. (1989a). Occasion setting with simultaneous compounds in rats. *Journal of Experimental Psychology: Animal Behavior Processes, 15,* 183–193.

Holland, P. (1989b). Transfer of negative occasion-setting and conditioned inhibition across conditioned and unconditioned stimuli. *Journal of Experimental Psychology: Animal Behavior Processes, 15,* 311–328.

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

Honey, R., Willis, A., & Hall, G. (1990). Context specificity in pigeon autoshaping. *Learning and Motivation, 21,* 125–136.

Hull, C. (1943). *Principles of behavior.* New York: Appleton-Century-Crofts.

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

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.

Kim, J. J., & Fanselow, M. S. (1992). Modality-specific retrograde amnesia of fear. *Science, 256,* 675–677.

Konorski, J. (1967). *Integrative activity of the brain: An interdisciplinary approach.* Chicago: University of Chicago Press.

Kraemer, P., & Roberts, W. (1984). The influence of flavor preexposure and test interval on conditioned taste aversions in the rat. *Learning and Motivation, 15,* 259–278.

Kubie, J. L., & Ranck, J. B., Jr. (1983). Sensory-behavioral correlates in individual hippocampus neurons in three situations: Space and context. In W. Seifert (Ed.), *Neurobiology of the hippocampus* (pp. 433–447). San Diego, CA: Academic Press.

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 preexposure and environmental preexposure determines subsequent learning. *Journal of Experimental Psychology: Animal Behavior Processes, 2,* 38–47.

Mackintosh, N. J. (1975). A theory of attention: Variations in the associability of stimuli with reinforcement. *Psychological Review, 82,* 276–298.

McIntosh, S., & Tarpy, R. (1977). Retention of latent inhibition in a taste-aversion paradigm. *Bulletin of the Psychonomic Society, 9,* 411–412.

McNaughton, B. (1991). Associative pattern completion in hippocampal circuits: New evidence and new questions. *Brain Research Reviews, 16,* 202–204.

McNaughton, B., Chen, L., & Markus, E. (1991). "Dead reckoning," landmark learning, and the sense of direction: A neurophysiological and computational hypothesis. *Journal of Cognitive Neuroscience, 3,* 190–202.

Morris, R., Garrud, P., Rawlins, J., & O'Keefe, J. (1982). Place navigation impaired in rats with hippocampal lesions. *Nature, 297,* 681–683.

Nadel, L. (1991). The hippocampus and space revisited. *Hippocampus, 1,* 221–229.

Nadel, L. (1992). Multiple memory systems: What and why. *Journal of Cognitive Neuroscience, 4,* 179–188.

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

O'Keefe, J. (1979). A review of the hippocampal place cells. *Progress in Neurobiology, 13,* 419–439.

O'Keefe, J. (1983). Spatial memory within and without the hippocampal system. In W. Seifert (Ed.), *Neurobiology of the hippocampus* (pp. 375–403). San Diego, CA: Academic Press.

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

Pavlov, I. (1927). *Conditioned reflexes.* London: Oxford University Press.

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,* 611–617.

Phillips, R., & LeDoux, J. (1992). Differential contribution of amygdala and hippocampus to cued and contextual fear conditioning. *Behavioral Neuroscience, 106,* 274–285.

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.

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

Rolls, E. (1990). Theoretical and neurophysiological analysis of the functions of the primate hippocampus in memory. *Cold Spring Harbor Symposia on Quantitative Biology, 55,* 995–1006.

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.

Saunders, R., & Weiskrantz, L. (1989). The effects of fornix transection and combined fornix transection, mammillary body lesions and hippocampal ablations on object-pair association memory in the rhesus monkey. *Behavioral and Brain Research, 35,* 85–94.

Schmajuk, N., & DiCarlo, J. (1992). Stimulus configuration, classical conditioning and hippocampal function. *Psychological Review, 99,* 268–305.

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.

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

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

Taube, J. (1991). Space, the final hippocampal frontier? *Hippocampus, 1,* 247–249.

Wagner, A., & Rescorla, R. (1972). Inhibition in Pavlovian conditioning: Application of a theory. In R. Boake & M. Halliday (Eds.), *Inhibition and learning* (pp. 301–336). San Diego, CA: Academic Press.

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

Whishaw, I., & Tomie, J.-A. (1991). Acquisition and retention by hippocampal rats of simple, conditional and configural tasks using tactile and olfactory cues: Implications for hippocampal function. *Behavioral Neuroscience, 105,* 787–797.

Winocur, G., & Gilbert, M. (1984). The hippocampus, context and information processing. *Behavioral and Neural Biology, 40,* 27–43.

Winocur, G., & Olds, J. (1978). Effects of context manipulation on memory and reversal learning in rats with hippocampal lesions. *Journal of Comparative and Physiological Psychology, 92,* 312–321.

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

Appendix

Implementation of the Model

The model generating the simulation results described in this article is functionally the same as that originally described in Gluck and Myers (1993). Implementation details are reviewed below.

Hippocampal Network

The hippocampal network is a three-layer network with full connectivity between 18 input nodes and 10 internal nodes and between those internal nodes and 19 output nodes. Input consists of an 18-bit pattern $I = (I_1, I_2, \ldots, I_{18})$ representing the current values of five phasic cues and 13 tonic contextual cues. Desired output $T = (T_1, T_2, \ldots, T_{19})$ is the same 18-bit input pattern, as well as a 1-bit prediction of reinforcement. The network is trained by a standard back propagation algorithm (Rumelhart, Hinton, & Williams, 1986). Node activations y are calculated as

$$y_j = f\left(\sum_i w_{ij} I_i + \theta_j\right) \quad \text{for each hidden node } j,$$

$$y_k = f\left(\sum_j w_{jk} y_j + \theta_k\right) \quad \text{for each output node } k, \text{ and}$$

$$f(x) = 1/1 + e^{-x}.$$

where w_{ij} is the weight between nodes i and j, and θ_j is the bias of node j. Weights and biases are initialized according to a uniform distribution $U(-0.3 \text{ to } 0.3)$.

Errors (err) are calculated as

$$\text{err}_k = (T_k - y_k) y_k (1 - y_k) \quad \text{for each output node } k \text{ and}$$

$$\text{err}_j = \left(\sum_k w_{jk}\, \text{err}_k\right) y_j (1 - y_j) \quad \text{for each hidden node } j.$$

Then weights and biases are updated as

$$w_{jk} = w_{jk} + \beta \text{err}_k y_j + \alpha w_{jk}^{\otimes} \text{ and}$$

$$w_{ij} = w_{ij} + \beta \text{err}_j I_i + \alpha w_{ij}^{\otimes},$$

where $w_{jk}^{\otimes}$ and $w_{ij\otimes}$ represent the previous changes to w_{jk} and w_{ij}, respectively; the learning rate $\beta = 0.05$ if $T_{19} = 1$ and $\beta = 0.005$ if $T_{19} = 0$; and the momentum $\alpha = .9$. Biases θ_k and θ_j are trained as if they were weights w_{bk} and w_{bj} from a node b that constantly outputs $y_b = 1.0$.

Cortical Network

The cortical network is a three-layer network with full connectivity between 18 input nodes and 60 internal nodes and between the internal nodes and a single output node. The input is the same as in the hippocampal network, whereas the desired output is the single bit T_{19} predicting reinforcement. Activation of nodes is computed as in the hippocampal network. The upper layer of weights, from hidden nodes j to output nodes k, is trained as in the hippocampal network, with learning rate $\beta = .5$ if $T_{19} = 1$ and $\beta = .05$ if $T_{19} = 0$. Weights and biases are initialized according to a uniform distribution $U(-0.3$ to $0.3)$. There is no momentum used in this network; otherwise, the

calculation of error and the updates of weights and biases are the same in this layer as in the hippocampal network.

However, the lower layer of cortical network weights is trained as

$$\text{err}_j = \sum_h v_{hj} y_h - y_j \quad \text{for each cortical hidden node } j,$$

where the v_{hj} are connection strengths from each hidden node h in the hippocampal network to hidden node j in the cortical network. These v_{hj} are nonadaptive and initialized according to a uniform distribution $U(-0.3, 0.3)$. $\sum_h v_{hj} y_h$ represents a composite training signal to cortical node j. Weight update then occurs as

$$w_{ij} = w_{ij} + \beta \text{err}_j I_j \quad \text{for each cortical hidden node } j.$$

In this layer, the learning rate $\beta = .1$ if $T_{19} = 1$ and $\beta = .01$ if $T_{19} = 0$. Weights and biases in this layer are initialized according to a uniform distribution $U(-3.0$ to $3.0)$. This initialization and the large number of internal layer nodes allows the lesioned model (cortical network only) to be able to solve random discriminations.

Stimuli and Training Schedule

Stimulus patterns are constructed by setting the first five bits to 0 or 1, depending on the presence or absence of five phasic cues. The next three bits code for a unique context: 101 for context X and 010 for context Y. The final 10 bits are a random string of 0s and 1s, constant across all stimulus patterns, but that evolve slowly with time, so that on any trial there is some probability, P = .01, that one of the 10 bits will be inverted; this inversion is permanent unless randomly inverted back.

The occasion-setting experiments are an exception because they require three distinct contexts. In this case, a five-bit code is used: 11101 for context X, 11010 for context Y, and 10110 for context Z. As a result, only eight bits are used for the random string of bits that evolve with time.

One block of training consists of a number of training trials, containing one presentation of each stimulus pattern being trained. These are intermixed with context-only presentations in a ratio of 1:20. For example, in the contextual conditional task ($AX+$, $X-$, $AY-$, and $Y-$), one block of training might consist of 10 presentations of context X, one presentation of phasic stimulus A in context X, 10 more presentations of X, 10 of context Y, one of A in Y, and 10 more presentations of Y. This ratio of context-only to training trials is about the minimum needed to ensure that background response to context alone remains low throughout training. Learning is assumed to have reached criterion when, for an entire block of training, the cortical network output deviates from T_{19} by less than 0.2. In the simulations reported in this article, each point in the figure graphs represents the average of 10 independent model simulations, except for data in Figure 2A, which represent the average of 30 simulations, which were then assigned among the three transfer conditions shown in Figures 2, B and C.

At the start of a simulation run, the network is initialized by training with 500 trials in which the input vector and output are both set to 0. This initialization ensures that the network has a low baseline output rate in the absence of input.

Received July 8, 1993
Revision received April 20, 1994
Accepted May 13, 1994 ■