7 Cortico-Hippocampal Interaction and Contextual Processing

Chapter 6 presented some examples of computational models of cortico-hippocampal interaction in classical conditioning. These models considered how conditioned stimuli were associated with responses and what role the hippocampal region might play in this association. But any conditioning experiment—indeed, any form of learning—takes place against a background, or **context,** including the sights, sounds, and smells of the environment. There are also internal contextual cues such as motivation and drives. Typically, researchers try to minimize contextual cues or control for them by making sure that all subjects experience similar context. Nevertheless, it has long been recognized that contextual cues can and do affect what is learned.[1]

From the early days of hippocampal research, it has been apparent that the hippocampal region plays an important role in contextual processing. Indeed, two early influential theories of hippocampal-region function suggested that the region's chief function is contextual processing in general[2] or processing spatial contexts in specific locations.[3] This chapter reviews several different approaches to interpreting the role of the hippocampus for learning about context.

7.1   OVERVIEW OF CONTEXTUAL PROCESSING

One of the difficulties in developing a theory of contextual processing is that context can be interpreted in several distinct manners. Context can be thought of as a set of unchanging, or **tonic,** cues. This contrasts with the more familiar, experimentally manipulated **phasic** cues, such as tones and lights, which have well-defined and limited duration and presentation rate. The CS is presented against a background of other tonic stimuli, such as the overhead lighting, the noises of the ventilation system, and the subject's knowledge of the time of day. Normally, the very nature of these background tonic cues means that they are not particularly useful as predictors of exactly when the US will appear. By contrast, a phasic cue, such as a light or tone, that always occurs just before the US will be a much better warning signal.

However, if no phasic cues are present, or if they are not good predictors of the US, then tonic contextual cues may be the only available predictors of reinforcement. For example, suppose food reinforcement is occasionally available in one experimental chamber X but never in a second chamber Y and there are no phasic cues, as schematized in figures 7.1A and 7.1B. In this case, animals will tend to approach the food cup in context X but not in context Y, as shown in figure 7.1C. Apparently, the animals have learned an association between context X and reinforcement, as if context were not fundamentally different from any other kind of CS. This kind of learning is not dependent on the hippocampal region: Animals with hippocampal damage can still learn to respond in context X but not in context Y.[4]

Context = "Just Another CS"

(A) Context X→US
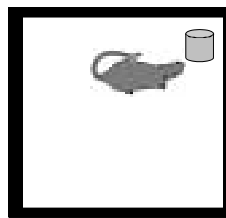
(B) Context Y→no US

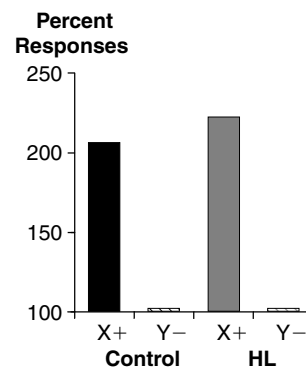(C) Behavioral Data

Percent Responses
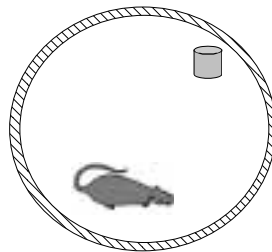
Figure 7.1   In an experimental paradigm, there may be both phasic cues, with explicit temporal duration, and tonic or contextual cues, which are present throughout the experiment. If no phasic cues reliably predict the US, then contextual cues may be the best available predictors. For example, if food is available in context X (A) but not in context Y (B), then animals will learn to make a response (e.g., approach the food cup) in context X but not Y. In this case, the context may enter into direct associations with the food US, just like any other CS. (C) Hippocampal-lesioned (HL) rats learn this distinction just as well as controls. Data are presented as percentage of baseline responding in unrewarded context (Y–). (Plotted from data presented in Good & Honey, 1991.)

However, context can also function as something more than just a collection of tonic cues. *Context can provide a framework against which other learning occurs.*[5] For example, suppose an animal is placed in a particular experimental chamber (context X) and trained that a cue, such as a light, predicts a US, as shown in figure 7.2A. The animal will learn to respond to the light. Once this light-US association is well learned, the animal is then tested in a new context Y that is dramatically different from X, as illustrated by figure 7.2B.

Under many conditions, if the light is presented in this new context Y, normal (control) animals will give a significantly reduced response, compared to the responding in the old context X, as shown in figure 7.2C.[6] This reduced responding in the new context implies that the animal learned not only about the light, but also about the context X in which the light originally occurred. In this case, the context appears to modulate responding to the light, helping to determine what kind of response is appropriate on the current occasion.
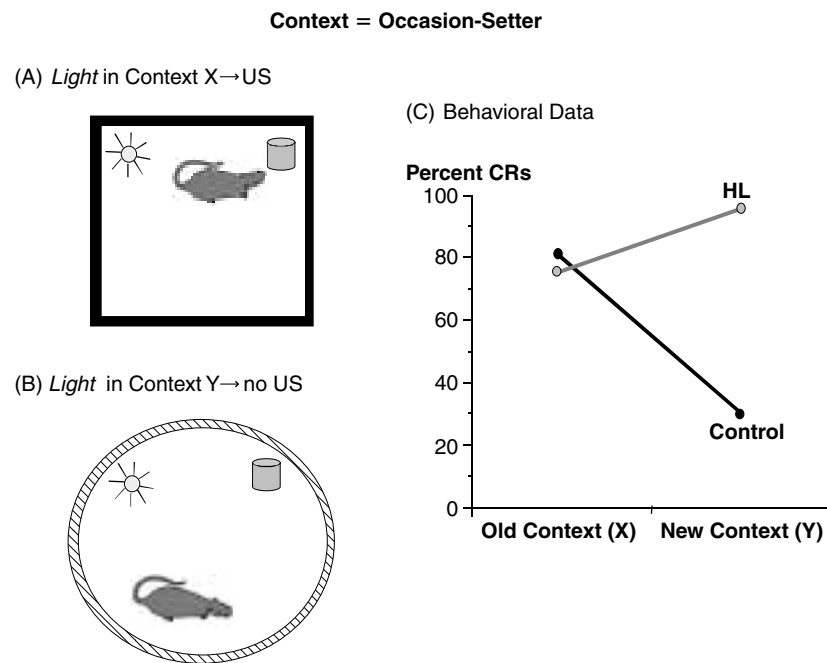
Context = Occasion-Setter

(A) *Light* in Context X→US

(C) Behavioral Data

Figure 7.2 Context can also provide a framework for interpreting the meaning of phasic CSs. (A) For example, an animal is trained that a light cue predicts a US (e.g., food) in one context X. (B) If the light cue is then presented in a novel context Y, there may be a decreased level of responding. (C) In rabbit eyeblink conditioning, hippocampal-lesion (HL) eliminates the context shift effect. (Plotted from results presented in Penick & Solomon, 1991.)

This kind of contextual processing—sometimes called **occasion setting**—does appear to be dependent on the hippocampus, since hippocampal-lesioned animals show responding at the same level when a familiar cue is presented in a novel context (figure 7.2C).[7] Note that this different behavior for hippocampal-lesioned animals is not necessarily more or less desirable than that seen in intact controls; what is clear is that the hippocampal-lesioned animals generalize differently than the intact controls to novel contexts.

The use of context to disambiguate the meaning of a stimulus is not limited to classical conditioning experiments. Most of us will react differently to a snake, depending on whether we encounter it in the context of a barefoot walk through the woods or in the context of a visit to the zoo. Context is also not limited to external cues such as physical setting. Context can include internal states such as motivation. For example, a rat can be placed in the base of a T-shaped maze and trained to run into the left or right arm of the maze to obtain a reward. If the left maze arm contains food and the right maze arm contains water, then the correct response (run left or run right) depends on whether the animal is hungry or thirsty. After a few trials, the rat will learn to run toward the water when it is thirsty and toward the food when it is hungry.[8] In this example, the maze itself is the conditioned stimulus that elicits responding, and the animal's internal state is the context that determines which of two competing responses is appropriate. Hippocampal-region damage greatly disrupts an animal's ability to learn the correct response, indicating that the hippocampal region is also involved in using internal contextual cues to guide responding.[9]

In summary, depending on the situation, context sometimes can act just like any other kind of conditioned stimulus cue, thereby entering into direct associations with the US. The first part of section 7.2 describes a computational model of this kind of contextual processing. On the other hand, context also stands in a special, superordinate relationship to the cues that it contains simply by virtue of its tonic nature throughout a conditioning experiment; as such, context appears to be able to mediate responding to cues, without itself entering into direct associations with the US. The second part of section 7.2 describes a computational model that includes this kind of contextual mediation. Animal data suggest that the former, associative properties of context are not dependent on the hippocampal region, while the latter, mediating properties of context do depend on the hippocampal region. Richard Hirsh extrapolated from these and similar results and proposed that contextual processing is a chief role of the hippocampus. Specifically, Hirsh argued that, *while the hippocampus is not needed to mediate the associative properties of context*

(the ability of context to enter into direct associations with the US), the hippocampus is critical for contextual occasion-setting (using context to mediate choice between competing responses to a stimulus).[10] The last part of section 7.2 shows how computational models of hippocampus can incorporate both these kinds of contextual processing without requiring any special representational status for contextual cues.

7.2   COMPUTATIONAL MODELS

Context as "Just Another CS"

In the original description of the Rescorla-Wagner model, the authors noted that contextual cues (such as the odor and texture of the experimental chamber) could acquire associative strength just like the experimentally manipulated cues, such as tones and lights.[11] Rescorla and Wagner proposed that if the set of contextual cues is collapsed into a single compound cue X, then X could be associated with the US just like any other CS. Training to respond to a light in context X could be thought of as mixing presentations of the light&X compound with presentations of X alone (figure 7.3A); light&X would be paired with the US, while X alone would not. In this case, the light is a perfect predictor of the US and will accrue strong associative strength. Context X, by contrast, is only sometimes paired with the US; it will accrue little or no associative strength. Thus, there will be little responding to X alone in the absence of the light.

Alternatively, the experimenter may arrange things so that the context is a better predictor of the US than any other available cue. For example, suppose that the US sometimes appears in context X but never appears in context Y. Under these conditions, context X will accrue some positive associative strength (figure 7.3B). If the US appears only occasionally, context X will accrue associative strength proportional to the rate of US presentation compared to the total time spent in context X. On the other hand, since the US never appears in context Y, Y will not accrue any associative strength. In fact, this appears to be what actually happens: Animals show a higher rate of CRs in context X than in context Y.[12] This behavior is also shown in rats with hippocampal-region damage,[13] a result suggesting that the hippocampal region is not necessary for learning to condition to tonic contextual cues.

The convention that context can be viewed as just a collection of tonic cues that compete with phasic CSs for associative strength has been adopted by several other computational models.[14] This seemingly simple idea does, however, lead to surprisingly subtle and important implications.

(A) Light predicts US in context X.

(B) US sometimes appears in context X
but never in context Y.

X− US trial

Light/X+ trial

X   0.3

Y   0.0

US

X

0.0

Light   1.0

US

X− noUS trial

X   0.3

Y   0.0

US

X− noUS trial

X

0.0

Y− noUS trial

Light   1.0

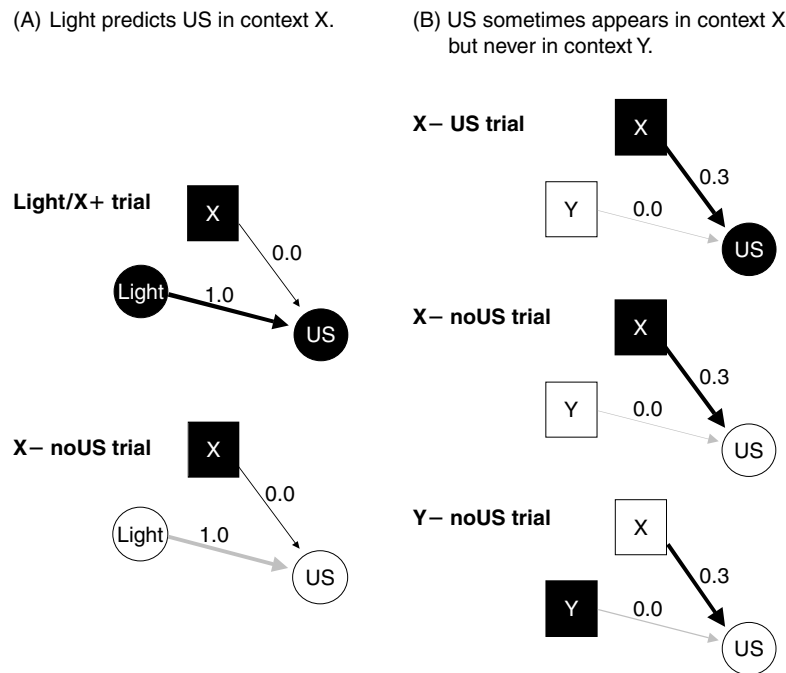US

X   0.3

Y   0.0

US

Figure 7.3   If context is treated as just another kind of cue, it can accrue direct association with the US if it is the best predictor of the US. (A) Learning that light predicts the US in a particular context X can be interpreted as mixing light-X+ trials with X− trials. In this case, light is a good predictor of the US and will accrue strong associative strength; X is only occasionally paired with the US and will acquire little or no associative strength. (B) Alternatively, the context may be the best available predictor of the US. If the US always occurs in context X but never in context Y, this is equivalent to intermixing X+, X−, and Y− trials. Context X will acquire associative strength, and the amount of associative strength will be related to how often the US appears in context X as a function of total time spent in context X. Context Y will acquire no associative strength, since it is never paired with the US.

For example, several models of spatial navigation adopt this same general framework.[15] Figure 7.4A shows an example of a canonical network that learns to map from a set of contextual cues (representing current location) to generate the next appropriate move. At each time step, the current location is provided as input to the network. Initially, the network generates a random output, representing a movement direction; some external process may then compute the resulting location and distance from the goal. If the goal is now closer than it was at the start of the trial, the network weights are updated to increase the likelihood that, next time the network is in the same initial position, it will make the same move. If the goal is farther away, then the weights
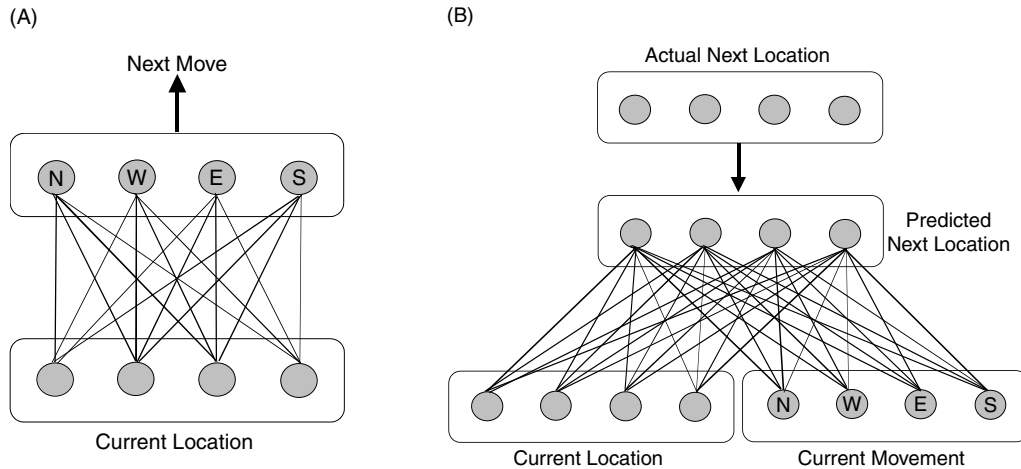
(A)

Next Move

(B)

Actual Next Location

Figure 7.4   (A) A simple network that learns to navigate. Inputs represent the contextual cues that are available at the current location. Outputs represent possible next moves (typically, a movement of one step in the north, west, south, or east direction). Only one output may be active at a time. If a reward is encountered, the weights are adapted to increase the probability of making the same movement the next time that location is encountered. This general format can be made more sophisticated in a number of ways: The inputs can be extended to include additional information about head direction and egocentric coordinates, the implementation can be mapped onto various brain regions, and so on. The basic principle still remains that responses are made by allowing contextual information about spatial location to accrue direct associative strength. (B) A slightly more sophisticated network that learns to predict the next location in a spatial navigation task. Inputs represent the current location and the current movement; outputs represent the predicted next location. This is compared to the actual next location, and weights are updated accordingly. If the network is allowed to "imagine" the expected results of each possible movement, another module can choose between possible movements on the basis of which generates the most desirable next expected location (e.g., closest to goal). This network can also act as a sequence generator if the predicted next location is fed back into the inputs to generate further predicted next locations. With enough iterations, the network may generate an entire path from the current location to the goal location.

are updated to increase the chances that the network will choose a different move next time. When the network is fully trained, it can accept a single location as input and generate a movement trajectory leading to the goal.

Figure 7.4B shows a related spatial navigation network, which takes inputs coding both the current location and the current move and learns to predict the next location.[16] If the network is allowed to experiment with different moves, it can "imagine" the effects of each possible movement and determine which results in an expected next location that is closest to the goal. This kind of network is not limited to spatial learning, but can also perform

general sequence learning, mapping from one state to output that predicts the next state.*[17]

Networks such as those shown in figures 7.4A and 7.4B allow tonic contextual cues representing place to enter into direct associations with outputs that can be used to guide behavioral responses—either directly (figure 7.4A) or indirectly (figure 7.4B). Such networks largely ignore the occasion-setting properties of context, by which context can influence stimulus-response learning without itself accruing associative strength.

### Combining the Associative and Occasion-Setting Properties of Context

Section 7.2A reviewed some computational models that focus on the *associative* properties of context: namely, the ability of tonic contextual cues to enter into associations if they are reliable predictors of reinforcement. This is distinct from the *occasion-setting* properties of context, whereby context is used to disambiguate competing responses to a stimulus. The associational processes appear to survive hippocampal-region damage, while the occasion-setting properties of context may depend on the hippocampal region. Apparently, these two contextual processes are separable and may be instantiated in different locations within the brain.

Our cortico-hippocampal model (figure 7.5) is also a two-process model, which proposes that *associational* CS-US learning is mediated by extrahippocampal substrates, while some kinds of *representational* CS-CS learning may require the involvement of the hippocampal region. To the extent that tonic contextual cues enter into associative learning, they are no different from any other CS, and so this kind of learning is based in the cortico/cerebellar network within the model.

For example, the background contextual cues are present not only when the US arrives, but also during long intertrial intervals when the US is *not* present. In contrast, a phasic cue that is present only just before US arrival is a much better predictor of the US. Because all cues (phasic and tonic) compete for associative strength, phasic CSs usually accrue associations at the expense of tonic contextual cues.

---

*These spatial navigation networks are interesting as computational devices that can be used to guide robot movements, but there is also a neurobiological tie-in. There are neurons in the hippocampus that respond when an animal is in a particular region of space. These neurons, called place cells, may provide the kind of information needed as input to the networks shown in figure 7.4. Thus these networks provide one possible interpretation of how place cells might be exploited in the brain. For a fuller discussion of place cells in the brain, see O'Keefe, 1979, Kubie & Ranck, 1983; Taube, 1991; Redish, 1999.

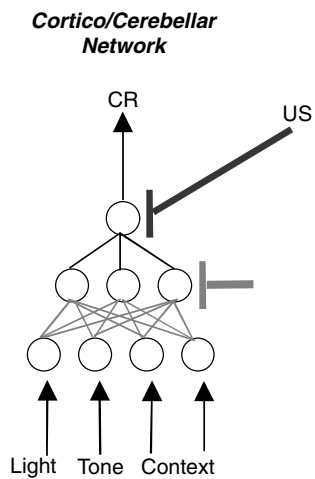(A) Intact Cortico-Hippocampal Model
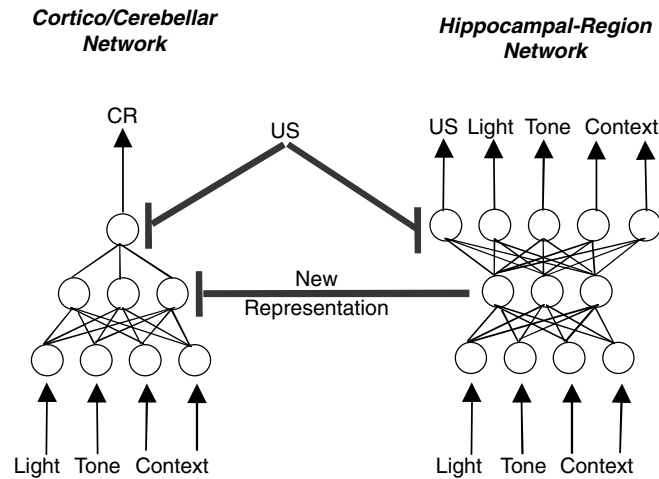
(B) HR-Lesion Model

Figure 7.5   The cortico-hippocampal model, as introduced in section 5.2D, instantiates two kinds of contextual processing. In the intact model (A), the hippocampal-region network forms new stimulus representations that include information about all co-occurring cues, including the context in which a phasic CS occurs. Thus, it can implement contextual occasion-setting. At the same time, contextual cues are free to enter into associations with the US if they are the best available predictors of that US. This ability of the context to enter into direct associations with the US can occur in both the hippocampal-region and cortico/cerebellar networks. In the HR-lesion model (B), the occasion-setting properties of context are lost, but context can still enter into direct associations with the US.

However, in the absence of phasic CSs that reliably predict the US, contextual cues will compete for associative strength (refer to figure 7.1). If the US occasionally occurs in context X but never in context Y, both the intact and HR-lesion models will show stronger responding in X than in Y (figure 7.6A).[18] Similarly, both intact and hippocampal-lesioned rats can learn to make a conditioned response more often in context X than context Y (figure 7.6B).[19]

On the other hand, the hippocampal-region network contains representational processes in which the representation of any cue is affected by all the other cues—phasic and tonic—with which it co-occurs. Because of this, tonic contextual cues can develop occasion-setting properties and be used to disambiguate the meaning of phasic CSs. *Thus, our cortico-hippocampal model provides a way to reconcile the associational and occasion-setting properties of context without requiring any special treatment of context.*[20] Because the contextual occasion-setting processes are localized within the hippocampal-region
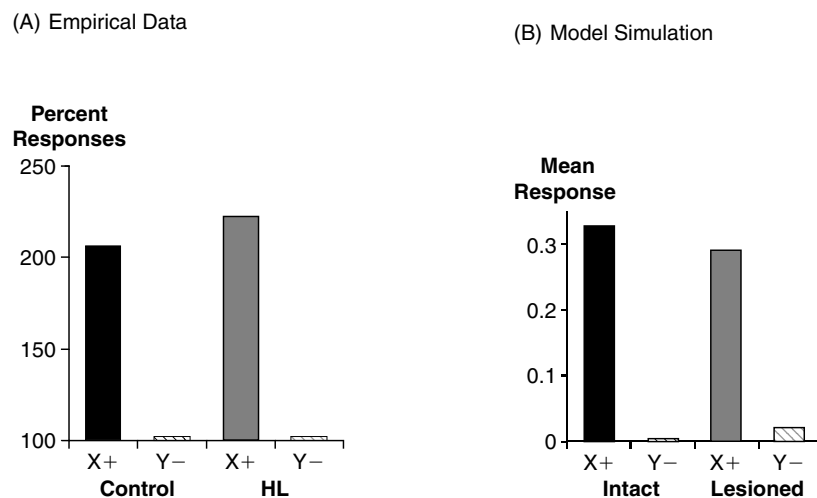

Figure 7.6   (A) In the absence of any phasic cue that reliably predicts the US, associations can be formed directly between context and US. One example is a situation in which a food US often arrives in context X but never in context Y. Normal rats will learn to approach the food cup more often in context X than in context Y. Hippocampal-lesioned (HL) rats show the same discrimination. Responses are plotted as a percentage of mean responding in context Y: typically about 2.5 to 3.5 responses per minute. (Plotted from mean response rate data reported in Honey & Good, 1993.) (B) The cortico-hippocampal model shows similar behavior, with both the intact and HR-lesion models giving a stronger response in context X than in context Y. After extended training, the strength of the response in X will be related to the frequency of US arrival in X compared to the total time spent in X. (Taken from Myers & Gluck, 1994, figure 5C.)

network, they do not occur in the HR-lesion model. This is consistent with animal data demonstrating that hippocampal-lesioned animals are generally less sensitive to the effects of context on learned CS-US associations. Two examples of this effect occur in the behavioral paradigms of context shift and latent inhibition, which will be discussed in the next section.

Context Shift Effects.   As we noted in section 7.1, an intact animal that is trained to give a strong response to a cue in one experimental context X will often show a decrement in responding when that cue is presented in a different experimental context Y, as is illustrated in figure 7.7A.[21] Our intact cortico-hippocampal model shows a similar response decrement after context shift (figure 7.7B).[22] Initially, the presentation of a tone cue in context X (tone&X+) and X-alone (X−) may evoke a very similar internal-layer representation in the hippocampal-region network, since the inputs overlap a great deal.

However, since tone&X+ and X− are to be associated with very different responses, the hippocampal-region network generates representations that are very distinct and overlap little (figure 7.8A,B). Once this is done, it is fairly trivial for the upper layer of weights to be set so that the representation
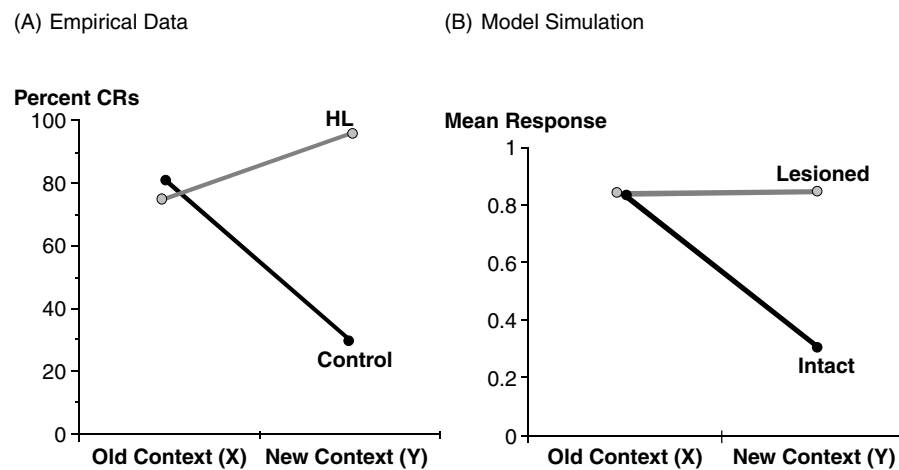
(A) Empirical Data

(B) Model Simulation

Figure 7.7   (A) In the eyeblink conditioning preparation, control rabbits show a decrement in responding when the trained cue is presented in a novel context X. Hippocampal lesion (HL) eliminates this response decrement after context shift. (Data plotted from results in Penick & Solomon, 1991.) (B) The cortico-hippocampal model shows similar behavior: The intact model but not the HR-lesion model shows decreased responding in the novel context. (Adapted from Gluck & Myers, 1994, figure 3A.)
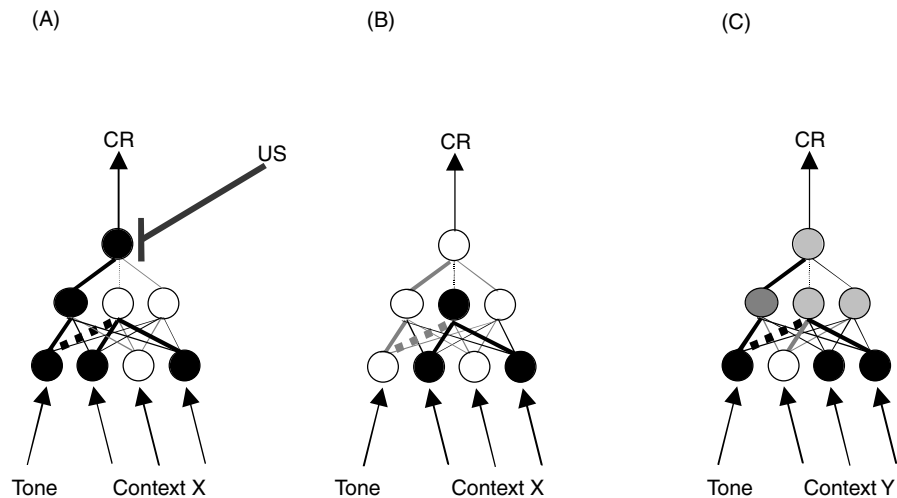
Figure 7.8   Contextual learning in the intact cortico-hippocampal model. (For simplicity, only the cortico/cerebellar network is shown.) Dark circles represent active nodes, lighter circles are more weakly activated nodes, and white circles represent inactive nodes. The network has been trained that a phasic cue, a tone, in context X predicts the US but context X alone does not. Internal-layer representations evolve that differentiate the tone&X inputs and X-alone inputs. (A) Tone&X activates a particular representation (strongly activating the leftmost internal-layer node but inhibiting the center one), and this representation is in turn mapped to a conditioned response. (B) When context X is presented alone, a different representation is activated. (C) The tone is presented in a new context Y, which shares some features with X. Elements of the representation of tone&X are weakly activated, and other nodes may be partially activated too, depending on their preexisting (fixed) connections with the input nodes encoding Y. The resulting representation may cause some output activity, but this is typically weaker than the robust conditioned response produced by tone&X.

of tone&X+, when activated, causes a strong response, while the representation of X− does not.

Now suppose that the tone is presented in a different context Y (figure 7.8C). The new stimulus tone&Y will activate a different subset of internal-layer nodes than tone&X; this in turn will mean that tone&Y will evoke a weaker response than tone&X. Thus, the intact model shows decreased responding when a trained cue is presented in a novel context. In this case, context modulates the response to a cue without itself accruing associative strength: Context X-alone does not evoke any kind of behavioral response in the absence of the tone cue.

The situation is quite different for the HR-lesion model. Recall that the HR-lesion model has only a cortico/cerebellar network and the internal-layer representations are not adapted. Thus, initially, tone&X and X-alone generate

some random activity patterns in the internal layer nodes, and these representations are not adapted. Since there is high overlap between the inputs encoding tone&X and X-alone, there will generally be high overlap between their internal-layer representations too (figure 7.9A,B). The network must learn to ignore those common elements and strongly weight associations to the output node from the internal-layer nodes that are activated by tone&X but not by X-alone. In this way, tone&X generates a CR, but X-alone does not. In effect, the network learns to completely ignore context and respond on the basis of whether the tone is present. When the tone is presented in a new context Y (figure 7.9C), context is still ignored, and the usual CR is generated. Thus, the HR-lesion model shows little decrement in responding after a context shift (figure 7.7B).[23] Consistent with this account, there is no decrement in responding after a context shift in hippocampal-lesioned animals (figure 7.7A).[24]

Latent Inhibition.   Latent inhibition is a behavioral paradigm in which unreinforced exposure to a stimulus, such as a tone, can retard later learning
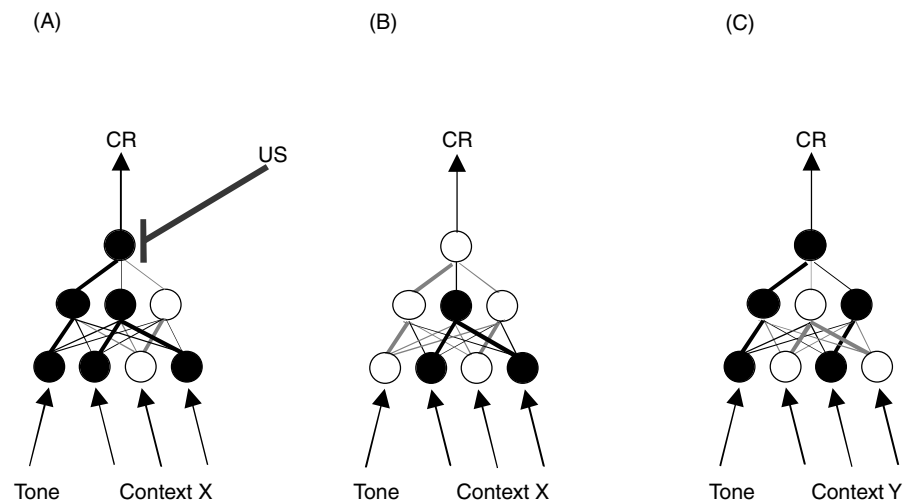

Figure 7.9   The HR-lesion model cannot construct new stimulus representations to differentiate tone&X from X alone. Whatever (fixed) representation is evoked by tone&X (A) will be partially activated by X alone too (B). To solve the task of responding only when A is present, the network must learn to strengthen weights from exactly those elements of the representation that are activated by tone&X but not by X alone. In effect, the system learns to ignore the context and focus only on whether the CS is present. Then, if the tone is presented in a new context Y (C), the network is relatively oblivious to this change, since the context has already been largely ignored. The result is that the lesioned system shows little or no decrement in responding when a trained cue is presented in a novel context.

about that stimulus. One way of interpreting latent inhibition is in terms of context exposure whereby exposure to a tone in context X (tone&X−, X−) retards subsequent acquisition of a response to that stimulus (tone&X+, X−).

Latent inhibition has been extensively studied in animals and humans with various forms of brain lesion and under various pharmacological agents. There are probably several substrates in the brain, all of which may contribute to the effect. One important contribution to latent inhibition appears to be attention: During the exposure phase, subjects learn to "tune out" the cue, which makes it harder to later acquire associations involving that cue. This attentional effect is beyond the scope of the cortico-hippocampal model. However, it is an important issue, given that latent inhibition is indeed disrupted in subjects who have attentional disorders (such as schizophrenia) or have been given attention-altering drugs (such as amphetamine).[25]

The cortico-hippocampal model provides an alternative, representational account of latent inhibition. During the first exposure phase (tone&X−, X−), neither the tone nor the context X predicts any salient future events. As the hippocampal-region network forms new representations, it tends to compress the tone and X, since they frequently co-occur and do not predict different outcomes. In effect, the network treats the tone as an occasionally occurring feature of the context X. In the second acquisition phase (tone&X+, X−), the network must differentiate the representations of tone&X and X if it is to learn to map one to a conditioned response and the other to no response.

This differentiation takes time, retarding learning relative to a condition in which there was no prior exposure to the tone and hence no compression of the tone and X to undo. Thus, in the intact model, after 100 trials of tone&X+ and X−, simulations that received exposure only to the context are already giving strong (0.8) responses to tone&X; in contrast, simulations that received exposure to tone&X− and X− are giving much weaker responses to tone&X (figure 7.10B).[26]

Because these representational processes depend on the hippocampal-region network, no latent inhibition is produced by the lesioned model (figure 7.10B). This is consistent with data showing attenuation or elimination of latent inhibition in animals with hippocampal-region damage (figure 7.10A).[27] More recent data have shown that various substructures of the hippocampal region may contribute differentially to latent inhibition; we will return to this issue in chapter 9.

Our cortico-hippocampal model also has several implications for latent inhibition in intact systems. For example, since latent inhibition is assumed to reflect representational compression of the tone and context X during the exposure phase, our model expects that the effect should be context
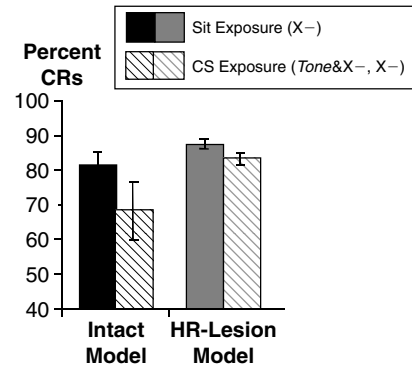
Latent Inhibition
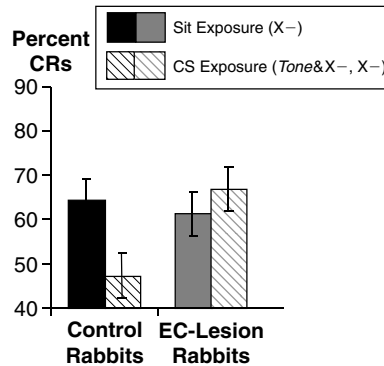
Figure 7.10   (A) In the rabbit eyeblink preparation, control rabbits show latent inhibition: Animals are slower to learn a response to a cue A following unreinforced exposure to a tone, relative to a Sit Exposure group that received equivalent exposure to the context X alone. Hippocampal-region damage can attenuate or eliminate latent inhibition, so exposure does not retard subsequent learning. (Plotted from data presented in Shohamy, Allen, & Gluck, 1999.) (B) The model shows a similar latent inhibition effect: In the intact model, unreinforced exposure to the tone slows subsequent learning compared to simulations that received equivalent amounts of exposure to the context alone; the lesioned model does not show the latent inhibition effect (Myers & Gluck, 1994). Note that the hippocampal region is only one of several substrates implicated in latent inhibition, which may be one reason why latent inhibition is more robust in the intact animal than in the intact model.

sensitive.[28] Thus, if exposure takes place in context X but acquisition takes place in context Y, there should be a release from latent inhibition, as is seen in figure 7.11. There is now good evidence that latent inhibition is similarly reduced by a context shift in rats and in humans.[29]

In a similar fashion, the cortico-hippocampal model shows disrupted latent inhibition if there is an intervening exposure to the context alone between exposure and acquisition phases (figure 7.11).[30] During the context-exposure period, the hippocampal-region network continues to update its representations. Since the context continues to appear without the tone, the tone gradually disappears from the representation of context. Given enough time, the representations of the tone and context will be redifferentiated, eliminating the compression that gives rise to latent inhibition. Empirical
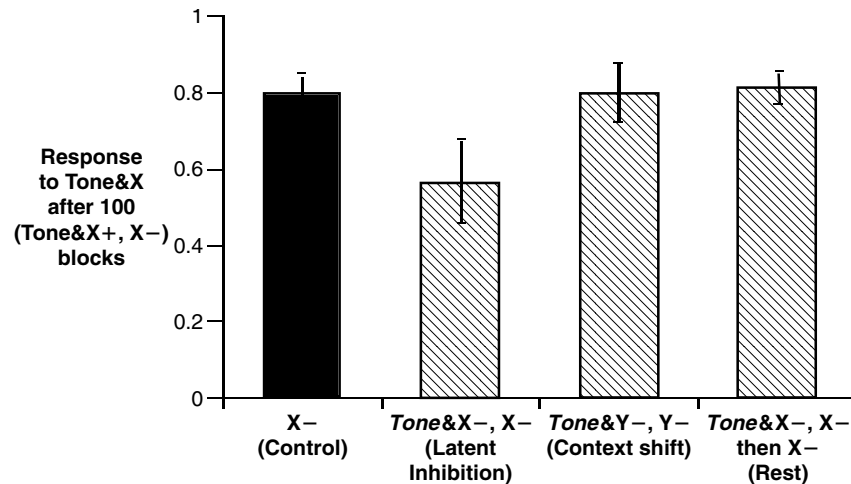
Figure 7.11   The cortico-hippocampal model predicts that latent inhibition should be context sensitive. Prior exposure to the stimulus (tone&X−, X−) results in slower learning to respond to tone&X+ than a control condition that received exposure to the context alone: the latent inhibition effect. However, if there is a context shift between phases, the latent inhibition effect disappears: tone&Y−, Y− exposure does not slow tone&X+, X− learning relative to the control condition. Normal animals do show such a release from latent inhibition following context shift (Lubow, Rifkin, & Alek, 1976; Hall & Honey, 1989). Similarly, exposure to the context alone between exposure and training phases eliminates the latent inhibition effect. A similar effect may also occur in animal conditioning, though further studies are indicated to test this prediction. (Adapted from Myers & Gluck, 1994, figure 4.)

data on this question are mixed: Many animal studies found that adding context-alone trials between exposure and acquisition can reduce latent inhibition, but another careful study found no such effect.[31] One study in humans found that a delay between exposure and training phases did not affect latent inhibition, but here the delay between training phases lasted only 120 seconds; possibly, a longer duration delay would be needed to show the effect.[32]

Finally, the cortico-hippocampal model predicts that overexposure to a cue may eliminate latent inhibition. For example, during ordinary exposure, the tone and the context X are equally unpredictive of any future salient event, and so their representations are compressed together. However, the hippocampal-region network is not only learning to predict the US; it is also learning to reconstruct its inputs including cue and contextual information. This means that on tone&X− trials, the hippocampal-region network should reconstruct the tone and X, while on X− trials, it should reconstruct X alone. Therefore, although the presence of the tone may not signal the US, it does signal that tone should appear in the hippocampal-region network's out-

puts. For this reason, although there is a bias to compress the tone and X on the basis of their nonprediction of the US, there is a competing bias to differentiate the representations of the tone and X. With enough exposure, the hippocampal-region network will develop representations that do indeed differentiate the tone and X somewhat. At this point, latent inhibition will be abolished. We are aware of one study directly addressing this issue in rabbit eyeblink conditioning; the results showed that overexposure to the CS could reduce the latent inhibition effect, just as predicted.[33]
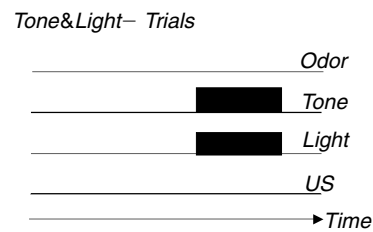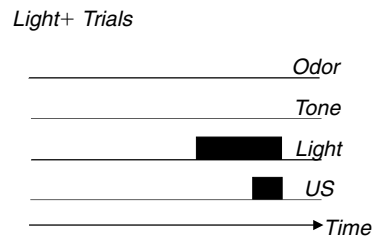
Thus, the representational processes in Gluck and Myers's cortico-hippocampal model appear sufficient to address a broad range of latent inhibition data. Like the other contextual effects, this is accomplished without giving context any special status: All cues, including context and phasic CSs, color the representation of all co-occurring cues. The model does not address the attentional explanation of latent inhibition. However, the representational and attentional explanations of latent inhibition are in no way mutually exclusive. Both may operate in parallel in the normal brain, producing similar latent inhibition effects for different reasons. More likely, the two mechanisms may cooperate.

The brain substrates that are believed to underlie the attentional component of latent inhibition—particularly the nucleus accumbens[34]—have reciprocal connections with the hippocampal region. In fact, it makes a certain amount of intuitive sense that the two systems would cooperate: Attention may be important in deciding which cues are emphasized or deemphasized within the hippocampal representation, while the hippocampus can help to detect novel cue configurations that are worthy of engaging attention.[35]
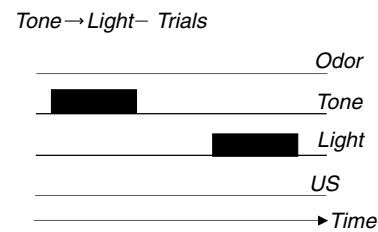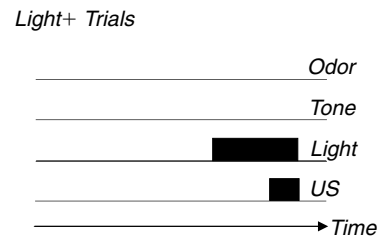
Occasion-Setting Properties of Phasic Cues

Occasion-setting is the ability to modulate responding without directly entering into associations with a US. Context is not the only kind of stimulus that can function as an occasion-setter. Peter Holland and others have documented the circumstances under which phasic cues can become occasion-setters.[36] For example, in ordinary **conditioned inhibition,** a cue (light+) predicts the US, while the compound (tone&light−) predicts no US (figure 7.12A). Normally, on light&tone trials, light and tone co-occur. However, if tone precedes light (figure 7.12B), then the presence or absence of tone determines the correct response to light. This kind of task is sometimes called a **feature-negative discrimination.** Similarly, in a **feature-positive discrimination,** light preceded by a cue (odor) predicts the US, while light alone predicts no US (figure 7.12C). In these cases, the tone and odor cues are
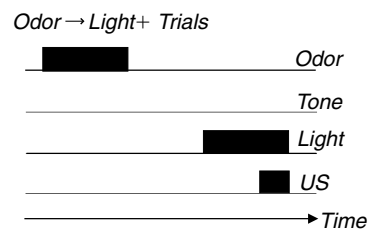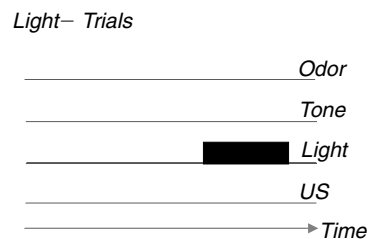
(A) Conditioned Inhibition

Light+ Trials

Odor
Tone
Light
US
Time

Tone&Light− Trials

Odor
Tone
Light
US
Time

(B) Feature Negative (FN)

Light+ Trials

Odor
Tone
Light
US
Time

Tone→Light− Trials

Odor
Tone
Light
US
Time

(C) Feature Positive (FP)

Light− Trials

Odor
Tone
Light
US
Time

Odor→Light+ Trials

Odor
Tone
Light
US
Time

(D) Simultaneous FN/FP

Tone→Light− Trials

Odor
Tone
Light
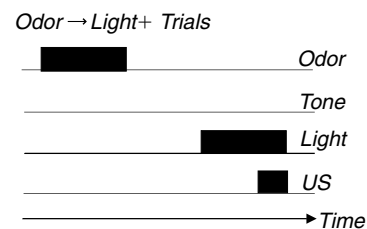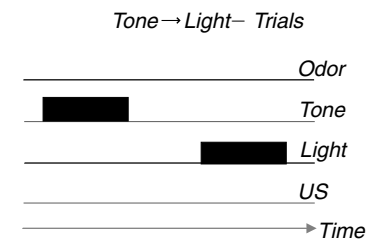US
Time

Odor→Light+ Trials

Odor
Tone
Light
US
Time

Figure 7.12   Occasion-setting with phasic cues. (A) In a conditioned inhibition paradigm, a CS (light) is reinforced when it occurs alone but not when it co-occurs with a second CS (tone). (B) In the feature-negative (FN) paradigm, light is reinforced when alone but not when preceded by tone. (C) In the feature-positive (FP) paradigm, light is reinforced when preceded by a cue (odor) but not when presented alone. (D) The FN and FP discriminations can be presented simultaneously so that light is reinforced when preceded by one cue (odor) but not by a second cue (tone). Unreinforced presentations of the individual cues (light, tone, and odor) may also be interspersed to ensure that the subject is not learning simple associations between tone or odor and reinforcement.

occasion-setters, which determine the correct response to light, although they themselves do not enter into direct association with the US.[37]

In at least one study, animals with hippocampal lesions could learn simultaneous feature-positive and feature-negative discriminations as well as controls; in fact, under some circumstances, the lesioned animals even showed facilitated learning.[38] This particular study used a lesion technique that destroys hippocampus and dentate gyrus but not nearby structures such as entorhinal cortex; it may be the case that broader hippocampal-region damage would indeed disrupt learning. Another study found that, while a selective lesion of the hippocampus did not disrupt feature-positive learning, a broader lesion that included nearby brain structures did indeed disrupt learning.[39] Clearly, further study is needed to resolve this issue. However, on the basis of the limited existing data, it appears that the hippocampus is not always needed for occasion-setting using phasic cues.

This in turn suggests that what is disrupted following hippocampal damage is specifically the ability to use context as an occasion-setter; this again points to a critical dependence of contextual processing on hippocampal-region mediation.

7.3   RELATIONSHIP OF COMPUTATIONAL MODELS TO QUALITATIVE THEORIES

Several prominent theories of hippocampal-region function focus on contextual processing. Hirsh and others have suggested that contextual processing is the hippocampus's chief role.[40] Our cortico-hippocampal model suggests a way in which the hippocampal region could contribute to contextual occasion-setting without having to define this as the region's special purpose. Context acts to mediate between competing responses to a cue by virtue of its inclusion in stimulus representations. This is no different from the way in which all (tonic and phasic) cues mediate the representations of all cues with which they co-occur.

Thus, there is no special treatment of context in our model; nevertheless, the model can account for a wide range of contextual phenomena. A similar argument can be made for spatial mapping: Place is simply one specialized form of context, consisting of a configuration of local views of space.[41] Thus, the devastating spatial impairment seen after hippocampal-region damage is really only a task-specific effect of disrupted contextual processes. Of course, in an animal such as the rat that depends heavily on spatial navigation, it is an especially visible effect.

A somewhat different view of contextual occasion-setting has been set forth by Mark Bouton and colleagues after years of elegant empirical studies examining the nature of contextual processing.[42] One of the most telling results to emerge from Bouton's lab is the contextual sensitivity of extinction. In one example, a CS is paired with shock in one context (context X) and then extinguished in another context (context Y); when the CS is again presented in context X, the response is renewed.[43] Various manipulations have shown that this effect cannot be ascribed simply to direct excitatory associations between context X and the US (or direct inhibitory associations between context Y and the US). The most parsimonious explanation seems to be that there are two associations: CS-US and CS-noUS; the context determines which response is retrieved.[44] During initial training, a CS-US association is created in context X (figure 7.13A); subsequently, during extinction, a CS-noUS association is created in context Y (figure 7.13B). At this point, the
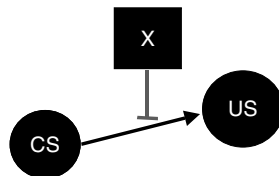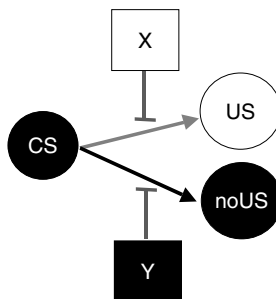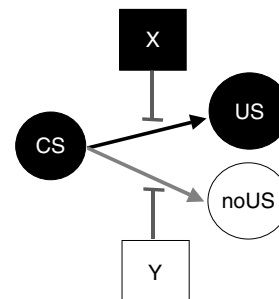

Figure 7.13   An account of extinction and renewal that involves contextual occasion-setting (Bouton, 1991). (A) Original learning that a CS predicts the US occurs in one context X. That context modulates the CS-US association but does not itself become associated with the US. (B) Subsequent extinction in context Y involves the development of a new association between CS and noUS. That association is retrieved during extinction training in context Y, making it appear as if the earlier CS-US learning has been extinguished. (C) However, if the CS is presented back in context X, the original association is retrieved, and the responding to CS is renewed. Thus, the context modulates which of the competing associations to CS is retrieved.

CS-US association learned in X has not disappeared, even though there is no responding to the CS in context Y. When the animal is returned to context X, the original CS-US association is retrieved, and the response "spontaneously" reappears (figure 7.13C).

We noted in chapter 6 that the cortico-hippocampal model does not adequately address extinction. When trained to give a response to a CS (tone+), this response can be eliminated by extinction trials (tone−). However, what the model does in this case is simply unlearn the response to the tone, setting the weights from the internal representation to the output nodes to zero. The model does not adequately learn a second, CS-noUS association, as is shown in figure 7.13.

It is possible that simple addition of multiple output nodes in the cortico/cerebellar and hippocampal-region networks would alleviate this problem, allowing internal representations to activate either a US or a noUS output. The particular representation that is activated on a current trial would depend not only on the CS, but also on the context, and thus context would affect the response retrieved in much the way Bouton and colleagues have suggested. If this is the case, there may be little theoretical difference between Bouton's theory of contextual occasion-setting and that embodied in our cortico-hippocampal model. Importantly, in Bouton's theory, as in the model, there is no qualitative difference in the ability of contextual cues and phasic CSs to operate as occasion-setters.[45]

There is some evidence that this kind of contextual modulation of extinction is dependent on the hippocampus and that various conditions that lead to renewal of an extinguished response in normal animals do not cause renewal in animals with hippocampal-region damage.[46] However, the data are mixed, and further study is needed to tease out the conditions under which hippocampal-lesioned animals may or may not show renewal of extinguished responding.

Another interesting aspect of the body of empirical studies amassed by Bouton and colleagues is their finding that, generally, there is no response decrement after context shift.[47] This is in contrast to the findings reported in the above section (refer to figure 7.7), but occasionally, other researchers have similarly failed to find a context shift effect.[48]

One factor that contributes to these conflicting results may simply be differential sensitivity to different kinds of contextual cues among different animals;[49] thus, some animal paradigms might be more likely than others to show context effects. Another possibility is that contextual dependence is increased when phasic CSs are poor predictors of the US.[50] Thus, if a CS only sometimes predicts US arrival, animals are more likely to respond to the

context alone;[51] interestingly, this effect is disrupted by hippocampal lesion.[52]

Our cortico-hippocampal model suggests a third factor that may contribute to contextual sensitivity of learned associations; as usual, this possibility may operate in parallel with the other factors suggested above. In the model, contextual sensitivity is a function of training time. It was mentioned above that latent inhibition in the model can be reduced by overexposure; it turns out that contextual sensitivity can be similarly affected by overtraining. Early in training, the hippocampal-region network is constrained to compress co-occurring cues; if a CS is always presented in a particular context X, then the features of X become part of the representation of the CS. However, *compression* is not the only bias constraining representations; there is also a constraint to *differentiate* the representations of stimuli that are useful predictors of future reinforcement. If the CS is the best predictor of US arrival, its representation should be differentiated from that of all other stimuli— including contextual stimuli. As a result, with extensive training, the representation of the CS may actually include less information about context. As a result, overtraining of a response to a CS will gradually decrease the amount of response decrement after context shift (figure 7.14).

In animal paradigms, different kinds of discrimination are learned at widely varying rates; an animal may acquire a conditioned fear response within a single trial, an odor discrimination within a few dozen trials, and a conditioned eyeblink response within a few hundred trials. If, as the model suggests, context sensitivity is reduced with overtraining, the exact number of trials required to eliminate contextual sensitivity may depend on the exact paradigm. However, there should still be a general trend to less context sensitivity with more training.

In 1994, we conducted a literature review and found this principle to hold across studies.[53] For example, rat conditioned suppression involves a hungry rat that is engaged in pressing a lever to obtain food; when a shock US appears, the animal briefly freezes, reducing (or suppressing) its normal fast rate of lever pressing. If a tone or light CS is paired with the US, the animal will come to suppress responding in the presence of the CS, anticipating US arrival. The degree of suppression to the CS is one measure of how well the CS-US association is learned. Conditioned suppression tends to be learned fairly quickly; most rats learn to suppress responding in the presence of the CS after only a few trials.[54] In this preparation, the response decrement effect occurs if the context is shifted after a single training trial[55] but not after 15–23 training trials.[56] In a slower preparation, rabbit eyeblink conditioning, the response decrement effect is seen even if the context is shifted after more
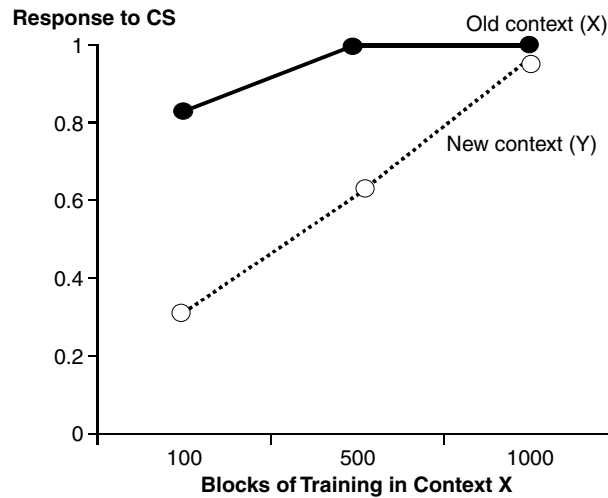
Figure 7.14   The cortico-hippocampal model predicts that contextual sensitivity may be reduced or eliminated by overtraining. After training to respond to a CS in one context X for 100 trials, the intact model develops a strong CR; if that CS is then presented in a novel context Y, the response drops dramatically. However, after 500 training trials in context X, there is less of a decrement in responding when the CS is presented in Y. After overtraining (1000 trials), there is no decrement in responding when the CS is presented in Y: The learned response is no longer context-sensitive. The model thus predicts that overtraining may similarly eliminate the response decrement with context shift in animal preparations. However, the precise amount of training that constitutes overtraining may vary, especially given that different kinds of learning are acquired at different speeds (e.g., it may take 40–70 trials to acquire a conditioned response but only one trial to acquire a fear response). (Adapted from Myers & Gluck, 1994, figure 3B.)

than 600 training trials;[57] the cortico-hippocampal model suggests that with enough overtraining, even this effect should disappear. Currently, our prediction of parametric sensitivity remains to be explicitly tested; however, we are aware of no data that contradict the prediction that a context shift will affect responding early in training but not after overtraining.

7.4   IMPLICATIONS FOR HUMAN MEMORY AND MEMORY DISORDERS

Just as Eichenbaum and colleagues describe their hippocampal-lesioned rats as "inflexible,"[58] Daniel Schacter has noted that humans with MT amnesia often seem inflexible or "hyperspecific" in their learning.[59] Although amnesic subjects may be able to acquire new information, they are often unable to express that learning if test conditions differ significantly from learning conditions. For example, in one paper, Schacter and colleagues recount a

study of an amnesic individual who was first read a list of little-known facts.[60] A few minutes later, he retained a few of them but generally forgot where he had heard them. He claimed that he guessed the answers or read about them in a magazine. By contrast, control subjects recall most of the facts—and also that they learned them in the course of the experiment. Thus, one important aspect of learning in amnesic individuals is an inability to re-member the context in which information was acquired, a condition that is sometimes called **source amnesia.**[61]

The idea of contextual occasion-setting schematized in figure 7.13 also has implications for how to improve learning in amnesic subjects. Figure 7.13B suggests that a single CS can be associated with multiple outcomes (US or noUS) and multiple responses (CR or noCR) and that context helps to dis-ambiguate which of these is appropriate at the present time. If an amnesic in-dividual cannot use context in this way, then a single CS will be associated with multiple outcomes and responses, and there is no easy way to deter-mine which is currently appropriate. The default may simply be to choose the most recent or most often-repeated response rather than the one that is appropriate in the current context.

This may account for the finding that amnesic subjects are often dispro-portionately impaired at reversing a trained discrimination.[62] Typically, am-nesic subjects (and hippocampal-lesioned animals) can learn to respond to one stimulus A+ but not a second stimulus B− as well as controls; the prob-lem comes when it is time to acquire responses to A and B that compete with the old ones: An old association between A and the US conflicts with a newer one in which A predicts noUS. Control subjects may be able to interpret the two phases of the experiment as different "contexts": The response that was appropriate in the old context is no longer appropriate now. Without this contextual occasion-setting, amnesic subjects have a hard time favoring the newly developing A-noUS association in preference to the older, well-established A-US one. The result is that reversal may be very hard indeed.

This kind of interference may be a very general phenomenon. During most discrimination learning, there is a period of trial and error in which there are some correct responses and some incorrect responses. In normal subjects, the correct responses eventually win out, and the incorrect responses are sup-pressed; but in amnesic subjects, the incorrect responses may be much more able to compete, hindering learning. An amnesic subject's impairment in learning may therefore not only reflect absolute task difficulty, but also how many errors were made during learning—and thus how much interference was generated.[63]

One way to ameliorate the interference may be to reduce the likelihood that competing responses develop in the first place. One way to do this is

through a technique, developed by Herb Terrace, called **errorless learning.**[64] In errorless learning, the subject is prevented from making incorrect responses (or is guided toward making correct ones). For example, Terrace used errorless learning techniques to train pigeons to peck at a red light but not a green light to obtain food reward (for pigeons, this is a very difficult task to learn).[65] First, the experimenter turned on the red light and waited for the pigeons to randomly peck in the vicinity of the light; when they did, they obtained a food reward. The procedure was repeated many times, until the pigeons came to peck at the red light fairly reliably. Next, the experimenter would wait until the pigeons were not in a good position to peck and then briefly darken the red key. Slowly, the experimenter lengthened the amount of time the key was darkened and also gradually changed the color from red to green. The pigeons came to associate the red light with food and the green light with no food. Eventually, after many such trials, the experimenter could flick the light from red to green no matter where in the cage the birds were; the pigeons would peck only while the key was red. In this way, the pigeons were trained to discriminate red from green—all without ever having made a single erroneous response to the green key.

The same basic errorless learning strategy can be applied to humans.[66] One application is in the stem completion task. In the learning phase, subjects are shown a list of stems (e.g., "TH___") that can each be completed in a variety of ways to form common five-letter English words. Subjects are asked to guess the correct completion for each stem (e.g., "THUMB"). Often, of course, the subject will guess incorrectly (e.g., "THINK," "THING," "THANK"). If the subject does not guess correctly within four tries, the experimenter gives the correct answer. The process is repeated several times with the same list of word stems until the subject reliably generates the correct completion for each word stem. Next comes a testing phase, containing questions of the form: "One of the words you saw before began with TH. Can you remember what the word was?" Under these conditions, normal control subjects can correctly recall 60–80% of the studied words. Amnesic subjects achieve only some 30% correct, even if the testing phase comes immediately after the end of the learning phase.[67]

However, given an errorless version of the same task, amnesic performance improves considerably. In this version, subjects do not have to guess the word stem completions but are told that the correct completion for "TH___" is "THUMB." This prevents subjects from making any mistakes (incorrect guesses) during the learning phase. This errorless learning improves control subjects' performance on the testing phase: They now get 80–100% correct. Performance of amnesic subjects reaches as high as 70% correct—almost as good as control subjects.[68] Apparently, preventing the

amnesic subjects from making mistakes during learning somehow improves their ability to retain that information.

The same technique of errorless learning may be fruitful in more realistic situations. For example, a similar method was used to help a man with profound anterograde amnesia learn to program an electronic memory aid, even though he had completely failed to master this skill using normal (trial-and-error) learning techniques.[69]

SUMMARY

• Context—the tonic environmental stimuli, both external and internal—can and does affect what is learned.

• Some contextual learning involves direct associations between contextual cues and reinforcement (US). This kind of contextual processing can be described by the Rescorla-Wagner model and appears to be hippocampal-independent.

• Some contextual learning appears to involve contextual occasion-setting, in which context determines which of several competing responses to a stimulus should be retrieved. This kind of contextual processing appears to depend on the hippocampal region. It can be modeled by allowing context to influence CS-US associations but not itself enter into associations with the US.

• The cortico-hippocampal model assumes that all cues (including context) are part of the representation of all other cues with which they co-occur. In this way, context can affect the representations of CSs without itself requiring special representational status. The cortico-hippocampal model can show both context-US learning and contextual occasion-setting and assumes that the latter depends on hippocampal-region representational changes.

• Phasic cues (CSs) can also operate as occasion-setters; it appears that this kind of occasion-setting may survive hippocampal-region damage, meaning that contextual occasion-setting is selectively dependent on the hippocampal region.

• Hippocampal-region damage may eliminate the ability to use context to discriminate between competing responses to a stimulus. This in turn may imply that amnesic individuals are especially susceptible to interference from trials in which they made erroneous responses, as these errors will compete with memories of the correct response. Errorless learning may be a useful technique for reducing this interference.