

sponsible for the chicks' following their mother and not another hen on the other side of the barnyard? Does it have anything in common with the way the great tit learns to recognize his neighbors' songs? What does either of these processes have in common with the process by which the great tit comes to sing his song in the first place? What makes the snail's ability to reduce its reaction to harmless shaking seem to represent one kind of learning and the chicks' ability to identify the hen another?

This chapter outlines a framework for answering these and other questions about learning. We begin with a brief review of the history of research on learning and then consider a general way of thinking about learning and some ideas about its evolution. The longest part of the chapter discusses associative learning, primarily Pavlovian conditioning. Armed with the general framework and some facts about associative learning, we will be better able to analyze other examples of learning in future chapters. We will also be equipped to assess claims in chapters 10 through 12 that some animals sometimes behave in ways that cannot be the products of associative learning but rather require self-consciousness, a theory of mind, or the like.

3.1 General Processes and Adaptive Specializations

3.1.1 "Constraints on Learning"

Experimental studies of learning were stimulated by Darwin's (1859) claim that animal minds should share properties with human minds. As studies of learning came to focus more and more on simple associative processes (Boakes, 1984; Jenkins, 1979), the possibility that there could be diverse kinds of learning within each individual or in different species tended to be overlooked. The result was an approach that has come to be known as general process learning theory (Seligman, 1970), an attempt to account for all learning with the same set of principles. These were principles of associative learning, as studied in instrumental (operant) and Pavlovian (classical) conditioning. General process learning theory had its heyday in the 1940s and '50s. Although there were a few dissenters (e.g., Tolman, 1949), even now most texts on learning in animals devote the majority of their pages to associative learning, with perhaps an introductory chapter on habituation and sensitization and a concluding mention of "special" learning processes like imprinting and imitation.

Relatively recently, however, there have been a number of stimuli to considering a wider range of kinds and aspects of learning. In the mid-1960s, psychologists discovered several puzzling phenomena that seemed to be evidence of "constraints on learning," cases in which the supposedly general principles of conditioning ought to apply but do not seem to (Rozin & Kalat, 1971; Seligman, 1970; Shettleworth, 1972). The most important examples of "constraints" were conditioned taste aversion, or poison avoidance learning (box 3.1), and autoshaping. In autoshaping (Brown & Jenkins, 1968), pigeons are placed in an operant chamber and a disk on the wall (the pecking key) is

Box 3.1. Flavor Aversion Learning

When rats (and many other vertebrates) sample a flavor and become ill later on, they learn to avoid consuming that flavor. As first described by John Garcia and his colleagues in 1966, flavor aversion learning has two remarkable properties. First, it takes place even with delays of hours between sampling the flavor (the CS in this Pavlovian paradigm) and becoming ill (the US) (Garcia, Ervin, & Koelling, 1966). Second, in rats, learning with illness as a US is specific to flavors. Garcia and Koelling (1966) had rats drink from a tube of flavored solution and also exposed them to a noise and a light each time they licked ("bright noisy tasty water"). Some of the rats were made ill after drinking, whereas some were shocked through the feet while they drank. When tested with the light plus the noise or the flavor alone after conditioning, the poisoned rats avoided drinking the "tasty water" while the rats that had been shocked avoided drinking the "bright noisy water." Figure B3.1 displays data from a later experiment with this basic design but with the mode of presentation of the various CSs and USs better controlled than in the original experiment (Domjan & Wilson, 1972).

When they were first reported, these findings attracted tremendous attention because long-delay learning and CS-US specificity seemed to contradict then-current assumptions about the generality of the laws of learning. Some investigators rushed to test hypotheses that various uninteresting "general process" factors might have been responsible, while others were equally quick to claim far-reaching implications for them (Domjan, 1983). As discussed later in this chapter, the idea that learning might

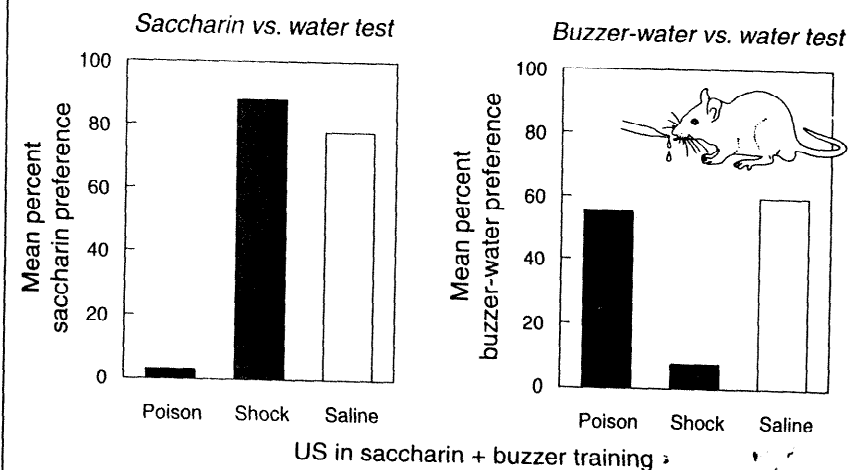


Figure B3.1. Test performance of three groups of rats that had experienced saccharin-flavored water infused into their mouths while hearing the sound of a buzzer and had then been shocked, injected with lithium chloride to produce illness, or given a control saline injection. In the test rats could drink plain water and either saccharin (left panel) or water in a bottle that produced a buzzing sound when licked. Conditioning is evident in reduced preference relative to controls. Redrawn from Domjan and Wilson (1972) with permission.

Box 3.1. Flavor Aversion Learning (*continued*)

be especially fast with certain functionally appropriate combinations of events is now generally accepted. However, it is still controversial whether the change in behavior that occurs when a flavor is followed many hours later by poison actually represents associative learning. Although the overall form of the function relating performance to the delay between CS and US is the same in flavor aversion learning as in other conditioning preparations (see figure 3.11), the aversions that develop with delays of hours may represent a nonassociative effect, namely reduced habituation of neophobia (i.e., avoidance of novel flavors) brought about by the poisoning (DeCola & Fanselow, 1995).

repeatedly lit for a few seconds and followed by presentations of food. There is no requirement for the pigeon to peck the key, but nevertheless pecking reliably develops and persists. This finding seemed related to the Brelands' (Breland & Breland, 1961) reports that animals being reinforced with food engaged in species-specific food-related behaviors that were counterproductive in the experimental situation. For instance, a racoon rewarded for depositing coins in a bank began to delay reinforcement by "washing" and rubbing the coins together in its paws. Attention was also drawn to the difficulty of training rats to perform anything other than defensive behaviors in avoidance experiments with shock (Bolles, 1970). To such observations of constraints on what animals could learn (or at least, do) in laboratory paradigms was added information about song learning, imprinting, and other "unusual" examples of learning observed by ethologists.

Around 1970 a small flood of articles appeared on the theme that general process theory had overlooked the biological aspects of learning (Bolles, 1970; Garcia, McGowan, & Green, 1972; Rozin & Kalat, 1971; Seligman, 1970; Shettleworth, 1972). Animals had learning mechanisms adaptively specialized for solving problems they face in nature. The narrow psychological approach to learning had overlooked these, but the newly discovered phenomena could be understood by thinking about the functions that learning had evolved to serve. However, despite proclamations that a revolution in the study of learning was on its way, in the ensuing years the original candidates for "biological constraints" and "adaptive specializations" were absorbed into a liberalized general theory of associative learning (Domjan, 1983). At the same time, by not formulating a clear research program with testable predictions, proponents of the "biological constraints" approach failed to stimulate research into other, related, phenomena that might have provided better evidence for adaptively specialized learning processes (Domjan & Galef, 1983). The term "constraints" in itself seems to assume a general process that is constrained in particular species and situations. However, consistent with the viewpoint presented in chapters 1 and 2, it is more appropriate to think in terms of evolved predispositions than in terms of constraints (Hinde & Stevenson-Hinde, 1973).

3.1.2 "Constraints," Modularity, and Other Areas of Psychology

At about the same time as these views about learning in animals were starting to be expressed, similar ideas were arising about cognition in humans. Some cognitive psychologists, most notably Neisser (1978), were discontented with a science of the human mind built almost exclusively on laboratory studies using such "unnatural" materials as nonsense syllables and lists of words. They advocated a more "ecologically valid" approach. Experimenters began to study peoples' memory for events in their own personal past, material learned at school, or events of public importance. Whether such research adds new theoretical insights into cognition is debatable (Banaji & Crowder, 1989; Bruce, 1985; Loftus, 1991).

The ecological approach to perception was better developed. As we saw in chapter 2, perceptual systems have evolved for responding to ecologically relevant stimuli. Some psychologists have argued that laboratory studies with simplified, abstracted stimuli fail to capture important functions of the system as a whole (Gibson, 1966; Shepard, 1984, 1994). The structure of the social environment should also be reflected in the nature of cognitive processing, a theme emphasized by evolutionary psychologists (cf. Barkow, Cosmides, & Tooby, 1992). Just as general process learning theory has dominated research on the animal mind, they argue, a model of human cognition as a single, general-purpose mechanism, the Standard Social Science Model (Cosmides & Tooby, 1992), has unduly dominated research on the human mind.

The idea that there may be a number of different, adaptively specialized, kinds of learning or cognition is related to two other ideas in the study of human cognition: intelligence is modular (Fodor, 1983), and memory consists of more than one system (Sherry & Schacter, 1987; Tulving, 1985). A module, in Fodor's sense, is an informationally encapsulated perceptual system. This means that it acts on a restricted kind of input unconsciously but in an apparently intelligent way. For example, visual perception includes a module that makes inferences about relative sizes and distances of objects. It is encapsulated because its operation is impervious to all but a specific kind of information. For example, measuring the lines in the Muller-Lyer illusion (figure 3.1) and discovering that they are equal does not prevent one's perceiving them as different lengths. The module that unconsciously perceives relative size is impervious to the "higher level" information obtained from measurement. The "intelligent" inference implicit in operation of a particular module can be applied only to that module's own constrained kind of information.

Striking examples of cognitive modularity abound in animal behavior. For instance, Cheng (1986) trained rats to find food buried in a rectangular box (figure 3.2). At the start of each trial the rat was exposed to the location of the food for that trial by being allowed to eat some of it. The rat was then removed and placed in an identical box with food hidden in the same relative

1993; Shanks, 1994). Some knowledge of basic conditioning principles is therefore useful for understanding the mechanisms behind some important phenomena in behavioral ecology. For instance, the ability to learn from punishing experiences may play a role in organizing animal society (Clutton-Brock & Parker, 1995).

3.4.1 The Conditions for Learning

Associative learning can be described as the process by which animals learn about causal relationships between events and behave appropriately as a result (Dickinson, 1980; Mackintosh, 1983; Macphail, 1996; Rescorla, 1988b). Associations are hypothetical connections within the animal that represent causal connections between events in the animal's environment. This functional description makes almost perfect sense of the conditions for associative learning. It also reflects the philosophical basis of the study of conditioning in associationism, which suggested that effects should be associated with their causes (Hall, 1994; Young, 1995). Associations have traditionally been thought of as the building blocks of all cognition, but seeing them as allowing animals to represent distinctively causal relationships makes associative learning just as adaptively specialized as, for example, learning about spatial or temporal relationships (Gallistel, 1992). In chapter 11 we consider whether different mechanisms are needed for processing information about distinctively social causation, as opposed to the physical causation being discussed here.

Contingency As we will see, the presence of an association can be evidenced in a variety of ways. However, to be sure that one is studying behavior reflecting the animal's knowledge of a predictive relationship, it is necessary to be able to discriminate this behavior from similar behavior brought about for other reasons. In the terminology of section 3.2, animals that have experienced a predictive relationship between CS and US at T1 must behave differently at T2 from control animals that experienced some other relationship between CS and US. The best way to isolate the effects of a predictive CS-US relationship is to expose the control group to the CS and US occurring randomly with respect to each other. The behavior of these latter, random control animals is a baseline that takes into account possible effects of exposure to CS and US individually in the experimental context (Rescorla, 1967). Against this baseline, the effects of both positive contingency (CS predicts US) and negative contingency (CS predicts absence of US) can be assessed (figure 3.6).

The importance of the random control group has been appreciated only relatively recently. Traditionally, temporal contiguity, or pairing, between two events was thought to be the necessary and sufficient condition for associative learning. The most popular control conditions eliminated contiguity by presenting just the CS or just the US or by presenting them systematically separated in time. But this condition has associative effects of its own. For in-

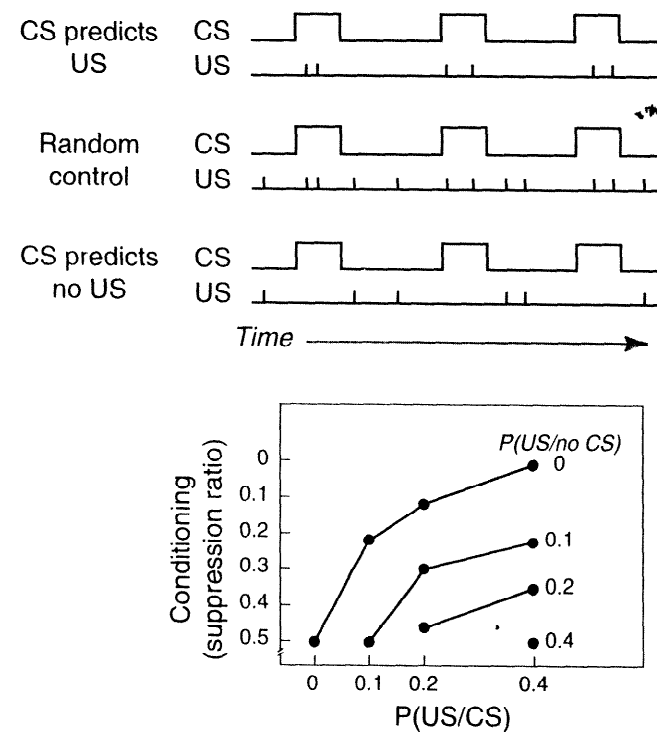


Figure 3.6. Effects of contingency on conditioning. In this example, illustrating the methods and results of Rescorla (1967), shock USs occurred with a constant probability per unit of time in the random control condition. Positive and negative contingencies were created by removing the USs between or during CSs, respectively. The effects of US rates on fear conditioning are plotted as a function of the probability of shock during the CS, with a separate function for each likelihood of shock in the absence of the CS. Rats were bar-pressing for food and fear conditioning was indexed by comparing bar-pressing rate during the CS to ongoing response rate. Zero indicates maximal conditioning; a suppression ratio of 0.5 indicates no conditioning. Redrawn from Rescorla (1988a).

stance, it can teach the animal that the US never follows the CS, thereby establishing a *conditioned inhibitor*. The truly random control condition is not without critics, however (Papini & Bitterman, 1990). An alternative approach often used with invertebrates to establish that they can learn at all is differential conditioning with two CSs. When a US is paired with one CS and concurrently not paired with another, associative learning should lead to differential responding to the two CSs. However, as a test of CS-US associations, this design is subject to a subtle confounding from possible differential habituation, since habituation may be selectively prevented to a CS that is always quickly followed by a US (Colwill, 1996). Colwill argues that the most convincing tests of associative learning make use of the fact that, as reviewed in section 3.4.2, a genuine CR reflects the quality and value of its associated US.

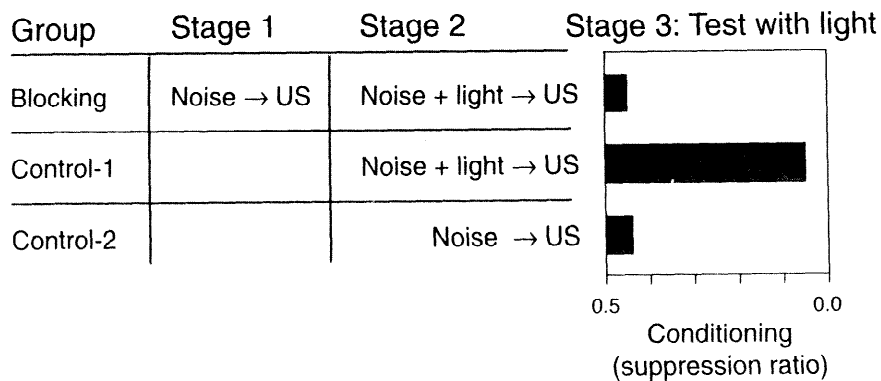


Figure 3.7. Design and results of Kamin's (1969) original demonstration of blocking of fear conditioning in rats. As in figure 3.6, conditioning was measured by suppression of bar pressing: lower suppression ratios correspond to greater conditioned fear.

Even the relatively simple stimuli used in most laboratory experiments on Pavlovian conditioning have many features. For instance, a tone comes from a particular location and has a particular loudness and duration. A visual stimulus has brightness, size, shape, and perhaps other features. How are multifeatured events processed—as a unit or as a sum of features? What about features that are added after initial learning? These questions are not settled yet even for simple associative learning (Fetterman, 1996; Hall, 1994). To see how they have been addressed, we start with a very influential model of Pavlovian conditioning based on the notion that animals learn separately about the features of events. The total *associative strength* of (or amount learned about) a compound stimulus, such as a light plus a tone, is the sum of the strengths of its separate elements. However, if an animal has learned that, say, a light causes food, and a new stimulus, say a tone, now accompanies the light so that the compound light + tone predicts food, learning about the tone, the new element, is blocked (Kamin, 1969; figure 3.7). Like the contingency effects in figure 3.6, *blocking* means that mere temporal contiguity between two events is not sufficient for associative learning. The CS must also convey new information about the US. In the case of blocking, the added CS conveys no new information about the occurrence of the US. When it does, as when the US is now larger or smaller than it was when predicted by the first CS, animals do learn (Mackintosh, 1978). This *unblocking* shows that blocking is not merely due to a failure of attention to the added element. It suggests that animals associate two events only when the second one, the US, is somehow surprising or unexpected.

The notion that surprisingness is essential for conditioning is captured in the influential formal model proposed by Rescorla and Wagner (1972). It generates the properties of associative learning reviewed so far, and many others besides. In that model, the amount of associative strength that a given CS accrues on a trial with a given US is proportional to the difference between the

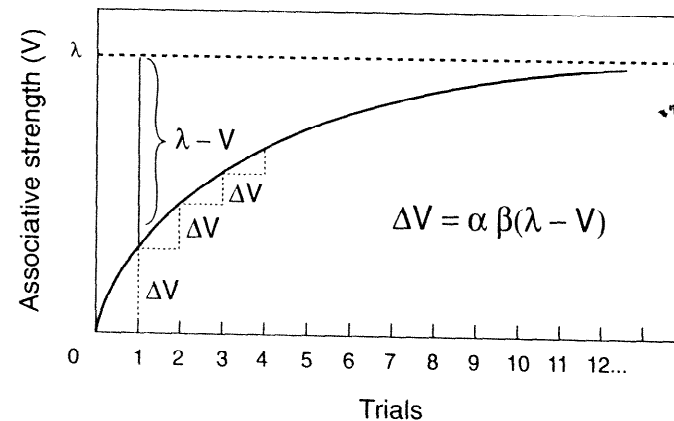


Figure 3.8. How the Rescorla-Wagner model generates a negatively accelerated learning curve.

maximum associative strength that the US can support (λ) and the current associative strength of all CSs currently present (V). The current associative strength of all CSs represents the degree to which the animal expects the US in the presence of those CSs, or the discrepancy between what the animal needs to learn and what it already knows (see figure 3.8). The parameters α and β in the equation are constants related to the particular CS and US being used. They reflect the fact that performance changes faster with salient or strong stimuli than with inconspicuous or weak ones.

The Rescorla-Wagner model readily accounts for blocking. When a novel stimulus, B, is added to an already-conditioned stimulus, A, the total associative strength of the compound is close to the maximum, thanks to the contribution of element A. If we assume that the environment in which an explicit CS occurs itself functions as a CS, the model also accounts for the effects of contingency as due to contiguity between CS and US. An explicit CS can be viewed as a compound of CS and context. When the predictive value of the CS is degraded by extra USs, as in figure 3.6, what really happens is that the context becomes associated with the extra USs and blocks conditioning to the CS. In this view, an animal exposed to random occurrences of CS and US is not an animal that has learned nothing; it may have associated the US with the environment or learned that the CS and US are unrelated.

The Rescorla-Wagner model has had a tremendous influence on the study of associative learning (Miller, Barnet, & Grahame, 1995; Siegel & Allan, 1996). Within the more than 25 years since it first appeared, most other proposals have been elaborations or variants on it. Two recent proposals have some essential differences from the Rescorla-Wagner model. In Pearce's (1994a) configural model animals process compound stimuli as unitary configurations rather than as the sum of elements. This model makes different predictions from the Rescorla-Wagner model for the results of some experiments on discrimination learning reviewed in sections 5.4 and 5.5. It is dis-