

## LEARNING, I: THE ACQUISITION OF KNOWLEDGE

Most animals are small and don't live long; flies, fleas, protists, nematodes and similar small creatures comprise most of the fauna of the planet. A small animal with short life span has little reason to evolve much learning ability. Because it is small, it can have little of the complex neural apparatus needed; because it is short-lived it has little time to *exploit* what it learns. Life is a tradeoff between spending time and energy learning new things, and exploiting things already known. The longer an animal's life span, and the more varied its niche, the more worthwhile it is to spend time learning.<sup>1</sup>

It is no surprise; therefore, that learning plays a rather small part in the lives of most animals. Learning is not central to the study of behavior, as was at one time believed. Learning is interesting for other reasons: it is involved in most behavior we would call intelligent, and it is central to the behavior of people.

"Learning", like "memory", is a concept with no generally agreed meaning. Learning cannot be directly observed because it represents a change in an animal's *potential*. We say that an animal has learned something when it behaves differently now because of some earlier experience. Learning is also *situation-specific*: the animal behaves differently in some situations, but not in others. Hence, a conclusion about whether or not an animal has learned is valid only for the range of situations in which some test for learning has been applied. Some new situation, not previously tried, may show evidence of learning. And as we saw in Chapter 4, even this test is one that cannot be perfectly applied in practice because we never have exact replicas of the animals we study. Lacking replicas, we cannot really compare the same animal with different histories to see which differences in history make a difference to the animal's behavior.

In defining a particular change as due to "learning" some kinds of experience are usually excluded. A change in behavior is necessary but not sufficient: For example, a change brought about by physical injury or restraint, by fatigue or by illness doesn't count. Short-term changes, such as those termed *habituation*, *adaptation* or *sensitization* (cf. Chapters 2 and 4) are also excluded — the change wrought must be relatively permanent. *Forgetting* has an ambiguous status: The change is usually permanent and does not fall into any of the forbidden categories, yet it is paradoxical to call forgetting an example of learning. Evidently it is not just *any* quasi-permanent change that qualifies. *Learning* is a category defined largely by exclusion.

There is little point in attempting a neat definition. What we need is some understanding of the ways in which an animal's environment can produce long-lasting changes in its behavior. If we can understand how these effects come about and, especially, if we can find the general principles that underlie them, then learning will be one of the phenomena explained. An adequate theory of learning must be just one aspect of an adequate theory of behavior in general.

The first step in understanding how the environment produces persistent changes in behavior is obviously to classify, but to classify *what*? Not types of change, per se; we would not expect the particular behavior affected — pecking, lever-pressing, calling, or whatever — to be especially revealing. Rather it is the *way in which* the animal's past experience guides its future behavior that is important — the set of *equivalent histories*, in the terminology of Chapter 4. If we can also come up with a compact description of the *changes produced* (the problem of "what is learned?") we will have a solid basis from which to make generalizations.

For convenience, I will distinguish two main types of learning, *template learning* and *reinforced learning*. The first type tells us about the kinds of change produced by learning; the second tells us about how these changes occur. The discussion in this chapter and the next also shows that these two aspects of learning are not independent. Imprinting and the song learning

of birds are discussed under the first heading. Classical conditioning and related forms of learning about causal relations that do not involve the animal's own behavior are discussed under the second.

## TEMPLATE LEARNING

### *Imprinting*

The young of many precocial<sup>2</sup> birds (such as ducks and chickens) learn about their parents during the first day or so after hatching. This learning, known as *imprinting*, takes place spontaneously, without any explicit “training” by the parents. The function of filial imprinting (i.e., imprinting by the young on the parent) is to bring the young under the mother's control, so that they will attend to her warnings of danger and follow her when necessary. In a few species, filial imprinting leads also to sexual imprinting, which enables sexually mature animals to seek out for mating individuals of the correct sex and species.

Imprinting is not restricted either to birds or to visual stimuli: Goats and shrews imprint to olfactory signals; ducklings in the egg form preferences for the frequency of the maternal call; Pacific salmon imprint to the odor of the stream where they were hatched; touch, temperature and texture are important in the attachment of a young monkey to its mother.<sup>3</sup> But we know most about the imprinting of precocial birds.

Chicks imprint in two steps. First, the chick tends to follow moving objects. Since in nature the only object it sees is the mother chick, it will follow her. After a little experience of this sort (a few minutes can be sufficient), the young bird learns enough about the characteristics of the followed object to pick it out from others and to follow it on future occasions. With further experience, the young animal begins to treat the imprinted object as a member of its own species. In a few species, filial imprinting may persist into later life and lead to aberrant mate choice. In the laboratory the imprinting of chicks can appear strikingly unintelligent, in the sense that the young bird can become imprinted to a variety of strange objects — a human experimenter (see picture), a hand, a box with flashing lights. So described, the process seems quite mechanical and mindless.

Yet on closer examination the mechanisms that underlie imprinting are more selective than may at first appear. For example, if provided with a *choice* of moving objects during the first posthatch day, the young bird is much more likely to imprint on the object resembling its natural parent.<sup>4</sup> Even after imprinting on some artificial object, the initial imprinting may wear off when the animal is exposed to the natural object later — and this is more likely the more unnatural the original imprinting object.

Imprinting also involves more than just a stimulus-response connection. For example, in one experiment mallard ducklings were raised with adults of another species. When the mallards were released as adults onto a lake containing dozens of species of geese and ducks, the males attempted to mate almost exclusively with females of the species that raised them. Obviously the birds learned more than just to hang around with their foster parents. This kind of generalization is characteristic of most learning by birds and mammals.<sup>5</sup>



A famous photograph of German ethologist Konrad Lorenz swimming with human-imprinted ducklings.

What kinds of experience are necessary and sufficient for imprinting? And precisely what changes do they produce? There seem to be two features essential to imprinting: a moving or flickering<sup>6</sup> object of a certain type, and a *critical period* when the bird is sensitive to such objects. The young bird must be exposed to the object within the first one or two posthatch days. There has been much argument about the importance of the critical period. There are two problems understanding what it means. First, imprinting when it occurs is hard to reverse, so that the animal's *first* experience with an imprintable object blocks any effect of later objects. And second, after the very early period, young birds become *neophobic* — afraid of new things — which interferes with potential imprinting to something new. The growth of neophobia with age leaves a brief time “window” when the young bird can imprint to a novel, moving object. Afterwards, such objects elicit only fear. Of course, neophobia implies prior learning, since the unfamiliar is defined only in contrast to the familiar, and familiarity is learned. Until it has learned about the familiar, the animal cannot even recognize what is unfamiliar, much less be afraid of it. The critical feature of imprinting may therefore be its irreversibility, plus a built-in tendency for young birds to fear the unfamiliar.

In nature, the female duck interacts with her chicks in ways that are not possible for an artificial imprinted object. Maternal calling is an important part of these interactions. It is likely that they serve to confirm imprinting in nature and make it less easily altered than some imprinting in the laboratory.

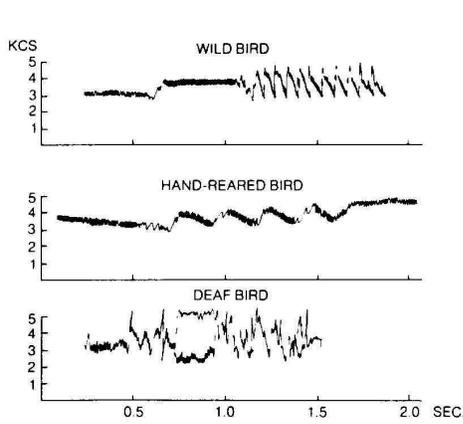
The essential change produced by imprinting seems to be one of *classification*. Young precocial birds seem to be innately equipped with an empty mental slot labeled “parent” (I define what I mean by “slot” more precisely later). The imprinting experience provides them with a set of natural objects to put into that slot. The animals have a ready-made repertoire of filial and, later, sexual behavior and just need to know where to direct it. Experiments have shown that the imprinting object can function as a reinforcer, in the sense that the young animal will learn to make some more or less arbitrary response to bring it closer. Chicks are clearly “comforted,” in some sense, by proximity to the imprinted object: they cease making “distress calls” and begin to make “comfort calls,” they move around more, and so on. Yet the concept of “reinforcement” has little to do with their learning: the imprinting object functions as a reinforcer because it has been classified in a certain way; it is not so classified because it is reinforcing. The real driving force seems to be the young animals' need to fill the empty “maternal” slot.

The slot may initially be empty, but it is not shapeless. Clearly some objects will fit it better than others, since the young birds imprint much more effectively to some objects than to others. This is the basis for placing imprinting under the category of *template learning* — a term originally coined to describe song learning in birds (about which more in a moment).

Although the effects of sexual imprinting are delayed, the major effects of filial imprinting are immediately apparent. Moreover, we do not know if the learning involves *practice* in any obvious sense. The young chick will follow the imprinting object, and imprinting is clearly better if it is allowed to do so; but following is not absolutely essential. In song learning by passerine birds, the effects are nearly always delayed and practice is essential.

### *Song Learning*

Songbirds differ in the extent to which song development depends upon experience. Song sparrows (*Melospiza melodia*), with one of the most intricate and beautiful songs, need relatively little experience. In one experiment<sup>7</sup> song sparrows were foster-reared by canaries in a soundproof room. The songs of the adults were indistinguishable from the songs of wild-reared birds: Song development was not impaired by foster rearing, and the sparrows showed no tendency to imitate the canary song. Other experiments have shown a small effect of depriving song sparrows of early experience, but clearly much of the information for the song-sparrow song is innate. At the other extreme, white-crowned sparrows (*Zonotrichia leucophrys*) develop

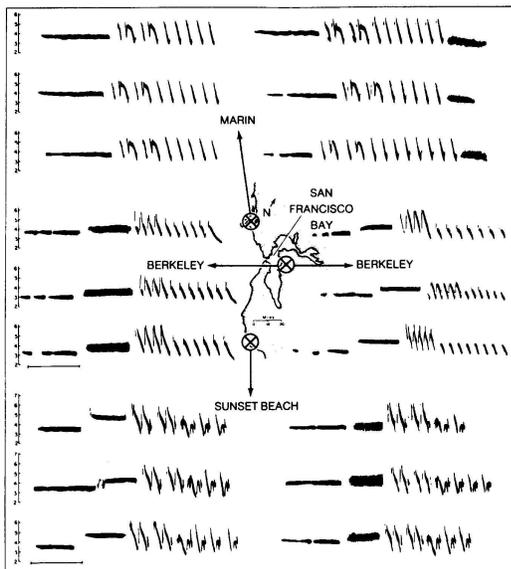


severely abnormal songs if deprived of the opportunity to hear species' song during the first six months of life. The upper sonogram in

**Figure 14.1.** Sonograms of songs of three white-crowned sparrows raised in different auditory environments: a wild bird, a hand-reared individual, and a bird deafened before the onset of singing. The sonograms show frequency as a function of time. Since most passerine song is close to being a pure tone (so-called whistle song) the sonograms mostly show just a single line of varying frequency. (From Konishi & Nottebohm, 1969.)

Figure 14.1 shows two seconds of the song of a wild sparrow, the middle sonogram shows the much simpler song of a deprived animal.

The Oregon junco (*Junco oreganus*) is an intermediate case. Individually isolated nestlings when adult produce wild-type songs but with a simpler syllabic structure. But juncos raised isolated in groups, or in general laboratory rooms where they could hear other species, produce more complex syllables and larger song repertoires, including sound patterns present neither in wild junco populations nor copied from other species. This study (Marler, Kreith, & Tamura, 1962), and work with song sparrows, shows that the ability to develop normal song in isolation does not preclude the ability to copy and to improvise new sounds.



**Figure 14.2.** Sonograms of the songs of 18 male white-crowned sparrows from three different localities within the San Francisco Bay area. Songs from a given area are similar to each other, and different from the songs of other areas. (After Marler & Tamura, 1964)

Song learning need not be delayed as it is in the white-crowns: Mocking birds, starlings, mynah birds and many parrot species imitate other sounds, and do so soon after hearing them. And the learning in this case is usually reversible: starlings seem to come up with a new repertoire every season, and most other imitative birds continue to add and delete items from their repertoire throughout life.

The song learning of male white-crowned sparrows and many other songbirds proceeds in two phases. For white-crowns the first phase occurs during a critical period from 10 to 50 days after hatching, when the immature males must have the opportunity to hear adult male song. Birds kept in isolation, or deafened, during this critical period never sing normally. But if the animals are prevented from hearing song either before or after

this period, the later development of adult song is unimpaired.

At about 8 months of posthatch age, the birds become sexually mature and begin to sing; this is the second phase. The onset of singing is not immediate, but is preceded by a period of *subsong*, a twittering unorganized pattern quite different from adult song — and similar in some

respects to the pre-speech *babbling* of human infants. Adult song emerges gradually from subsong over a period of several weeks.

Experiments<sup>8</sup> have shown that the first, early-experience, phase is essential for the bird to store a model or *template* of adult song. As in imprinting, the animal is selective in what it will store. For example, white-crowns exposed during the critical period only to songs of the song sparrow incorporate none of this into their model. Song development is just as impaired as if the birds had been reared in complete acoustic isolation. On the other hand, in the wild, the birds' early receptivity permits them to learn the local *dialect* of white-crown song (indeed, this may be the evolutionary function of song learning in this species). Figure 14.2 shows sonograms of dialect variation in the San Francisco Bay area: while the dialects are different from one another, there is obviously a family resemblance. Evidently the birds' mechanism for template formation restricts the class of possible templates. It is as if the bird were a musician programmed to learn a single piece of baroque music: such an individual would be immune to the charms of Brahms or the Beatles, but would learn at once anything by Bach or Vivaldi.

Even after the template has formed, song learning is not complete. The period of *practice*, when the bird is showing subsong and incomplete versions of adult song is also essential. The final song is evidently built up by a circular process in which the bird sings, hears the results of its own singing, and slowly modifies what it sings until its production matches the stored template.

This loop can be interrupted in two places: either by disrupting the motor apparatus, to prevent practice, or by deafening the bird, to prevent it hearing its own production. Experimentally deafened birds show normal subsong, but it never evolves into normal adult song. An example of the grossly abnormal song of an early-deafened bird is shown in the bottom panel of Figure 14.1. Moreover, this is not owing to some disruptive effect of the deafening operation: birds reared in a noisy environment so that they cannot clearly hear their own song also develop abnormal songs. These results prove the importance of auditory feedback. It is obviously difficult to impair motor function in a reversible way so as to prevent practice but permit singing in adulthood. As far as I know, this has not been attempted. It is conceivable that a bird denied the opportunity to practice would nevertheless sing normally as an adult when motor function is restored, but it seems unlikely. Almost certainly, both halves of the loop — motor and perceptual — are essential for normal song development in white-crowns, chaffinches and many other songbirds.

The requirement for practice is not limited to birds whose template is formed by learning. Song sparrows require little experience with conspecific song to develop normal singing, but they do need to hear their own early singing efforts. Song sparrows deafened at an early age produce extremely abnormal song. On the other hand several nonpasserine<sup>9</sup> birds, such as domestic chickens and ringdoves (*Streptopelia risoria*), need no auditory experience and develop their normal rather rudimentary calls even if deafened soon after hatching.

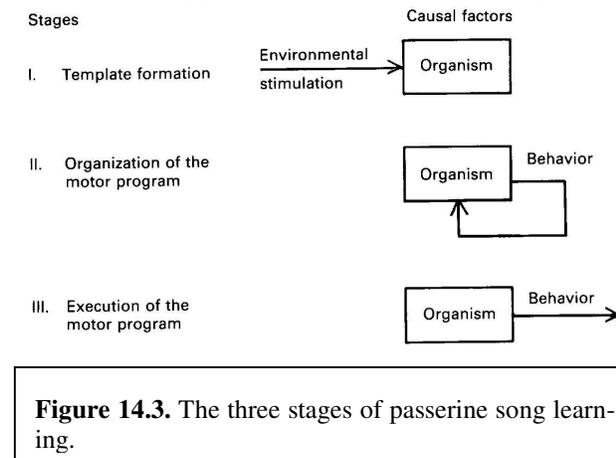
Once song sparrows, white-crowns and other songbirds have developed their adult song, it is largely immune to further interference. For example, white-crowns deafened as adults continue to sing essentially normal songs for many months.

These diverse processes of song development presumably have an ecological basis, although this is not always obvious. For example, the sensitivity of young white-crowns allows them to learn the dialect of their location, which raises an obvious question: What evolutionary advantage is conferred on a male that learns a dialect, rather than being innately provided with an adequate song? White-crowns occupy a wider range of habitats than some other similar species, suggesting that genetic sub-populations may be especially well adapted to particular habitats. Perhaps kin selection leads males to be more hostile to males singing a different dialect, thus favoring males that can learn dialects of the place they were born. There are probably other possibilities. The imitative abilities of mockingbirds may be related to the evolutionary benefits of a

large song repertoire, something which is itself not well understood.<sup>10</sup>

## Conclusion

For all their apparent diversity, imprinting, and the song learning of passerines, share two central features: In both cases, behavior is based upon a stored model or template; and the perceptual learning that leads to template formation is broadly *selective*: only objects, or songs, that satisfy certain general characteristics will be accepted as models. These properties are not universal, however. It is likely that the song learning of doves does not involve any kind of stored perceptual model: their simple song seems to be just a *motor program*, like walking or flying, that needs little practice and no specific external stimulation to develop normally.



**Figure 14.3.** The three stages of passerine song learning.

These characteristics fit conveniently into the three learning stages illustrated in Figure 14.3. In the song learning of white-crowns, the first stage is the laying down of the template; we know very little about the details of this process, because no overt behavior accompanies it. This stage is inferred from the results of deafening and deprivation experiments. The learning may be laborious or quite rapid. The mechanism involved has the property of selectivity I have just described. I am assuming that imprinting also involves the formation of a template, but in this case overt behavior (following, learning responses that bring the imprinted object closer, etc.) is tied to the template as soon as it is formed. But in neither case is the organism's own behavior directly involved: template formation is induced entirely by outside stimulation.

The second stage of song learning involves the circular interaction between the bird's own singing and the stored template. It is here that our ignorance of the nature of the template becomes most obvious. All we observe is a circular process that continues for several weeks and eventually stabilizes in adult ("crystallized") song. But whether the learning involves progressive approximation to a stored model (as the "template" idea implies), or whether the process converges by some completely different means we do not know. We can be sure only that the final form of the song is determined by early acoustic experience. In imprinting, there is nothing corresponding to the second or third stages.

The third stage of song learning is the final, "automatic" song, that is almost unaffected by disruption of the feedback loop that was essential to its original development.

The song learning of chaffinches and white-crowned sparrows involves all three of these stages, but the development of song-sparrow song skips the first stage. Evidently the template here is innate and requires no specific environmental input. Deafening experiments show that the second stage is necessary, however.

The development of song in non-passerines such as chickens and doves skips all but the last stage. In these species, the motor program itself is innate and requires no external input — either from the environment or the bird itself — to develop normally.

These three types of behavioral organization are based on experiments on imprinting and song learning in birds. Nevertheless, I believe that they are to be seen in other species learning quite different things in quite different situations. Indeed, I will argue that most learning of habits and skills follows just these three stages. I turn now to a consideration of the many types of learning that involve some kind of "reinforcement": the occurrence of a hedonic event in systematic relation to other aspects of the environment or to the animal's own behavior.

## REINFORCED LEARNING

Reinforced learning has been a major concern of psychologists since Thorndike and Pavlov. Unfortunately, for many years research tended to focus on ever-more standardized experimental “paradigms” in which a simple, preferably unidimensional, response could be studied under highly controlled conditions. The field has become increasingly divided into two topics: classical (Pavlovian) conditioning, and operant conditioning. This division conforms to no characteristic of the processes involved; the distinction between classical and instrumental conditioning is much more a matter of procedure (open-loop vs. closed-loop, in the terminology of Chapter 5) than of process (cf. Chapter 4). The division goes along with different theoretical and practical preconceptions. These different interests have led the operant conditioners almost to abandon work on the *process* of learning (i.e., of behavior change) and concentrate on steady-state adjustments. The same tendency has overtaken many classical conditioners, who, while affirming a deep interest in the learning process, usually present data only on steady states. The implicit notion that the learning process is something that should be describable by smooth curves has led others to ever more refined situations and ever larger groups of disposable subjects.

Learning means change, and change has given psychologists the same kind of trouble that motion gave Zeno: How can one study motion, when a body must be in one place or another? How can something move at all? Learning involves the same kind of moving target: At any time the organism “knows” some things and not others; we can clearly identify the beginning and ending states (ignorance → knowledge), but what happens in between is often obscure. The Procrustean solution is to act as if there is a single learning “process” and then design our experiments so as to preclude any other. Some of the more tightly controlled conditioning procedures have this flavor. A safer tack is to look at a range of learning phenomena in different species and see what useful generalizations emerge. In this section I shall look at the learning of bees, taste-aversion learning and several experiments on classical conditioning. A fair conclusion is that while there are some general learning *principles*, there is no single learning *process*. Learning is almost certainly *discrete* rather than continuous, and seems to involve the assembly of many elements, both preformed and remembered, to form new programs for action.

The learning of insects is most easily analyzed. Careful experiments have often been able to break down apparently complex learning into simple elements. For example, honeybees (*Apis mellifera*) learn the color, shape, location, odor and payoff schedule (many flowers secrete nectar only at certain times of day) of their food sources.<sup>11</sup> In their searches for new food sources and their exploitation of old ones they do not seem to be significantly inferior to many “higher” species even though their learning is built up out of very simple elements.

For example, the color of a flower is clearly visible to the bee before landing, while standing on the flower and while circling the source after takeoff. Yet the color of the flower is learned only during the final three seconds before the bee lands. A naive bee carried to an artificial feeder from the hive will circle after taking on a load of sugar water, but will not be able to pick out the correct feeder color on a second trip a few minutes later. Yet if the bee is interrupted during its first feeding, so that it must take off and land again, it remembers the feeder color perfectly. In the same way, bees learn landmarks only when taking off — the circling flight that reliably follows takeoff presumably serves to scan the environment for this purpose. A bee feeding at a new location and removed before finishing has no recollection of the location, even though she had plenty of opportunity to scan landmarks when first arriving.

These peculiarities are necessary adaptations to the bee’s limited brain size: Like a small computer, the bee must do one thing at a time, and cannot afford to store useless information. It makes functional sense for the animal to note landmarks only after leaving, because she then knows whether the place is valuable enough to be worth recording (bees don’t circle when leav-

ing a potential food site that provided no food). Presumably color is learned when landing because that could be done equally well before or after landing, but landmark-learning belongs better to the takeoff phase and preempts the “processor” at that time. While on the flower, the bee is looking for and gathering nectar; evidently this activity is sufficiently demanding that no processing resources are left over for color learning. Since the bee must obviously look at the flower in order to land on it, color learning can occur most efficiently at that time.

The learning of a bird or a mammal is not so perfectly sequential as the learning of bees, although it too is subject to subtle constraints. (Human visual attention has its own rigidities, which are taken advantage of by sleight-of-hand artists and magicians, for example.) When a pigeon homes on its loft or a swallow returns to its nest after a foraging flight, it must process the same kind of information as the bee, even if not in such a rigid order. Moreover, the phenomenon of *attention* discussed in earlier chapters represents exactly the same kind of constraint as the compartmentalized learning of the bees. Mammals, birds — and man — have limited processing resources, and if the task at hand is difficult, dealing with it effectively limits the amount of attention that can be paid to other tasks. We may all be able to walk and chew gum at the same time, but few can do mental arithmetic at the same time as they answer questions or read a book. The larger brains of mammals and birds give them this advantage over the bees: they can *shift* attention in relatively flexible ways, rather than being pre-programmed to attend to particular things at particular times.

The learning of bees provides two further lessons. First, learning an apparently straightforward task, such as the location and identity of a new food source, involves the encoding and storage of a number of elements, the solution of a number of subsidiary problems. Second, the actual storage of information — remembering a color or the properties of landmarks, for example — does not require a behavioral explanation. Remembering the color of a flower seems a relatively simple process for a bee. It takes place rather rapidly, has no behavioral accompaniment, and does not seem susceptible of further behavioral analysis. Such events might be termed *atomic acts*. Further analysis of such acts must be at the level of neurophysiology. For complete understanding we need to know how the visual system encodes physical properties of the visual world, and what changes take place in the animal’s nervous system when it sees, and remembers, a particular colored object. These are fascinating problems for neuroscience, but take us to a nonbehavioral level of analysis.<sup>12</sup> The task for behavioral analysis is to discover the atomic acts, the program, involved in the learning of particular tasks.

### *Learning as program assembly.*

Analysis into elements reduces the process of learning to the building of a program, in which rules and elements are combined to produce a system that can solve a particular problem, such as getting to a particular food source. In bees the process is often relatively simple, since steps are added one at a time, as in a small digital computer. Moreover, their learning seems largely to involve defining *variables*, to be operated upon by built-in programs, rather than building the programs themselves. Thus, the worker honey bee is innately provided with a set of routines for reacting to landmarks and colored food sources; its learning consists largely in applying those routines to particular landmarks and flowers. The bee is like a hand-calculator that knows innately how to multiply and relies on the environment only for the input numbers.

In higher animals, the process is much harder to get at, because in addition to defining variables (the process termed *stimulus control*) program elements — rules — are combined in new ways during learning. Many things seem to be happening at once, or at least in rapid succession, often with no obvious behavioral accompaniment, which further complicates the task of analysis. These difficulties force us to deal with learning in birds and mammals at a more abstract level, in terms of poorly defined concepts such as “template,” “coordinate system,” “internal representation” and “response strength” even though our ultimate objective is to reduce this

learning also to the level of atomic acts.

### *Reinforcement and learning*

The bee learning provides some insight into an old issue in learning theory: the role of reinforcement in learning. The question was, Does reinforcement affect *learning* or just *performance*? We know that reinforcement affects performance, that is, action, because we can train animals to do things by rewarding them for it — and often they will not do certain things *unless* they are rewarded for it. But is this *all* that is affected by reinforcement?

This question was difficult to answer because learning is a vague concept, and reinforcement seems to be retroactive. And how is *performance* different from learning? Obviously we cannot assess learning other than by watching the animal's behavior, so how can we ever detect learning that produces no change in performance? Is the question even a meaningful one? The second problem is the indirect relation between reinforcement and its effects: a reward comes *after* the behavior being rewarded, so the immediate effect of reward must be on the animal's *internal state* (in the sense defined in Chapter 4), rather than directly on the behavior being rewarded. Presumably the animal's state is changed in such a way that the rewarded behavior becomes more likely. So perhaps the question should be rephrased as follows: Are changes in the animal's state (of the sort we might like to term *learning*) produced only by reward or punishment, or can they take place independently of reward and punishment?

The bee learning provides an answer: Bees store the color of a flower *before* they have landed on it, hence before they have any opportunity to be rewarded. Clearly color, an essentially neutral stimulus, can cause a change in the bee's state independently of reward. On the other hand, the bees store information about landmarks only after, and if, they have been rewarded. And they only demonstrate by their behavior that they have learned if they have eaten: It was necessary to reward the bees so that failure to choose correctly could be attributed to something other than lack of reward. Moreover, it is quite likely that information about flower color is *erased* from the bee's modest memory if the flower fails to yield nectar, although this might be hard to demonstrate behaviorally. Consequently it might well be impossible to demonstrate what was once termed *latent learning* in bees, even though other experiments have proved that the animals can remember things without being rewarded for doing so.<sup>13</sup>

The question as originally posed is an example of one of those undecidable propositions about black-box systems discussed in Chapter 4. Since we must infer learning from performance, and we know that reward is often necessary for performance, we simply cannot give a general answer to the question of whether reward is necessary for learning. As the bee example illustrates, it is quite easy to come up with learning process in which reward is essential to the demonstration of learning, yet deeper analysis shows memory effects that are independent of reward.

On the other hand, based upon both logic and the bee experiments, we can say with certainty that reward and punishment are not the only events that can cause a change in an animal's internal state. And we can decide experimentally whether a given thing is learned before or after the animal is rewarded. The bees encoded flower color before being rewarded by the flower, but it is conceivable that they could have learned it afterwards, as they learn landmarks. It is logically possible that an animal notice nothing until it has obtained a reward, at which time it scans its environment and learns something about it: All learning could be like bee landmark learning. The converse is also a theoretical possibility: Animals could learn everything in advance of reward, using reward just as a guide to action, as in bee color learning. Neither of these extreme arrangements seems uniquely efficient. Real learning is surely a balance of both.

The Pavlovian or classical conditioning “paradigm” provides a convenient method for the study of problems such as the order in which things are learned, what is learned, innate biases in learning, and the ways in which causal relations are represented by animals.

### Classical Conditioning

The evolutionary function of “reinforced learning” is to detect regularities in the environment related to things of value to the animal. For example, the function of the bee learning just described is to identify and locate sources of food. Learning of this sort is a process of *inference*, in the sense that certain environmental features, such as landmarks (a particular location) and colors are taken as signals of food

Any inference such can be mistaken: if the environment changes in between observations, or if too few observations are made, for example. If the experimenter changes the color of the feeder after the bee’s first visit, the animal will be unable to locate the food source: its initial inference that “red” predicts food will have been proved wrong. If the feeder is only full some of the time, a single observation will give unreliable information. In other words, some cues are more predictive than others. One can imagine an experiment with artificial feeders having two properties: a color, and a pattern. The experimenter can arrange a predictive relation between food and either color or pattern or both. Let there be two patterns, X and Y, and two colors, A and B. Suppose that pattern X predicts food (all X-type feeders contain sugar water), but pattern Y does not (all Y-type feeders are empty). Color is made irrelevant (A-and B-color feeders both contain food half the time). These contingencies (cf. Chapter 5) are illustrated in Table 14.1, which shows the four possible feeder types (AX, AY, BX, BY) and the probability of finding food at each. If the bee can learn about only one aspect at a time (as seems likely), and if it learns on first exposure, then its choice of whether to attend to color or pattern is critical. If it attends first to color, it will always be wrong. Obviously this is a case where the animal must make *several observations* before it can detect the real invariance in the relation between visual appearance and food potential.

This need for more observations is almost independent of the animal’s computational capacity. Even if the bee could attend to *both* color and pattern on each trial, it still would not know which one was more important until it had sampled several flowers, that is, until it could fill in from its own experience the probabilities in Table 14.1.

This example shows why (in a functional sense) animals learn some things rapidly and others more slowly. Speed of learning is determined both by the processing limitations of the beast, and by the necessity to sample the environment a sufficient number of times to obtain a reliable estimate of invariant features.

The problem of precisely *how much* sampling an animal should do is an exceedingly difficult one theoretically, even in a simple environment with known properties.<sup>14</sup> Nevertheless, the main factors are straightforward: (a) How much does the animal already know? That is, what are the *prior probabilities*? to use the Bayesian term,<sup>15</sup> and how many of them are there? (b) How large is its computational capacity? That is, how many factors can it assess on each trial? (c) How important are the consequences of the learning? This will be determined by the relative costs and benefits of “hits”, “misses” (failing to respond to a valid signal) and “false alarms” (responding to an invalid signal), together with the costs of sampling.<sup>16</sup>

**Innateness of priors.** Factor (a), the prior probabilities, is likely to be related to the variability of the environment: Properties that are variable should be learned more slowly than properties that are fixed because the animal can be less certain of the meaning of a given conjunction: Should the child bitten by a dog be afraid of all dogs, just German shepherds, or just *this* German

|       |   | Pattern |   |    |
|-------|---|---------|---|----|
|       |   | X       | Y |    |
| Color | A | 1       | 0 | .5 |
|       | B | 1       | 0 | .5 |
|       |   | 1       | 0 |    |

**Table 14.1** Probability of food for feeders with different color-pattern combinations

shepherd? Obviously more data are needed: providing the risk is not too high, he should sample a few more dogs — unless the genetically coded experience of his ancestors (have dog-like animals been reliably dangerous in the past?) can bias him in one direction or another. The appearance of flowers, even of the same species, is quite variable: blooms vary in size and tint even from day to day. On the other hand, the *odor* of a given flower doesn't vary at all. Odor is a much more *valid cue* to the identity of a flower than is color. Moreover, the number of possible flower colors is very large, whereas the number of possible odors is much smaller. Correspondingly, bees should take some time to learn the color of a flower, because of the need to experience a representative sample, whereas odor learning should be rapid. Bees show precisely this difference: odor is learned on one visit, while color takes several and is never learned perfectly. Time of day is learned more slowly still, presumably reflecting typical variation in the time at which flowers secrete nectar.

In simple animals the priors are innate. Bees do not learn colors more rapidly if raised in an environment in which colors never vary and are always reliable predictors of food. Nor can they learn to be less impulsive in learning odors.<sup>17</sup> Even in mammals and birds, examples of innate bias of this sort are common. For example, a rat made sick by a harmless injection of lithium chloride half an hour after tasting a new food in a new cage will develop a strong aversion to the novel taste, but not to the new cage. *Taste-aversion learning*, of which this is an example, follows the classical-conditioning paradigm (an unconditioned stimulus [US], sickness, follows a conditioned stimulus [CS], the taste of food) — but has several unusual properties.<sup>18</sup> It is relatively insensitive to long delays (conventional Pavlovian conditioning would be almost impossible with CS-US delay of 30 minutes or more), it is highly selective (a taste CS is much more effective than CS of some other kind), and it occurs on a single trial.

The difficult experiment of raising rats in an environment in which something other than taste is correlated with gastrointestinal effects has not been done, so we do not know if the priority of taste as a CS for sickness is completely immune to alteration by experience. However, taste-aversion experiments have been done with very young animals, and here also taste is dominant over other stimuli. The dominance of taste is evidently innate.

Previous experience does play a key role in taste-aversion learning, however. If a rat is made sick some time after having tasted both a familiar and a novel food, an aversion develops only to the novel food. Here is a perfect example of Bayesian inference. The prior probability that the familiar food is poisonous is obviously low, since it has been eaten safely on previous occasions. If the animal must make a decision on the basis of a single exposure, then obviously the novel food is the preferred candidate.

The other characteristics of taste-aversion learning are also highly adaptive. The primacy of taste in general follows from the invariable relation (in the history of any species not subject to X-rays or intrusive experimenters) between food and sickness: animals may feel sick for reasons other than poison, such as disease, but nonfood stimuli never cause sickness. The rapid learning follows from the severe risk associated with eating poisonous substances (factor (c), above: “false-alarms” can be very costly). The cost of omitting a safe item from the diet will generally be small; the cost of eating poison may be very high. This skewed payoff matrix obviously biases selection in favor of conservatism: rapid taste-aversion learning and *neophobia*, avoidance of novel objects, particularly foods. And indeed, rats (especially feral rats) are very cautious when confronted with new foods; they eat only a little at first and wait some time, usually without eating any other food, before incorporating the new food into their diet.

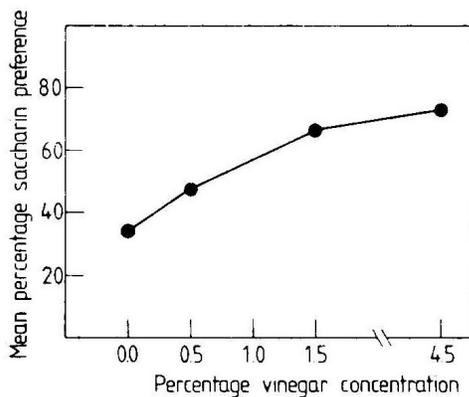
The long delay sustainable by taste-aversion learning is probably not (or at least, not entirely) a specially programmed feature. Because taste is the effective stimulus, and tastes and other stimuli are not easily confused, a taste experienced half an hour before sickness will not be confused with all the non-taste stimuli experienced in the interim. The main reasons that long-delay Pavlovian conditioning with non-taste stimuli fails are (a) because there is nothing to dis-

tinguish the CS from concurrent non-taste stimuli and stimuli that occurred before and after the CS, so that the animal has no way of knowing which stimulus to associate with the US; and (b) even with repeated trials in which the only invariant relation is between the CS and the (delayed) US, these other, irrelevant stimuli impair recall for the CS and thus prevent the animal making any connection between it and the US. If the animal cannot remember a particular stimulus that occurred 30 minutes ago, it is in a poor position to show a conditioned reaction to that stimulus as a consequence of a US that just occurred. (Cf. the discussion in the previous chapter of the sources of proactive and retroactive interference in temporal control).

Two kinds of experiment are necessary to back up these conclusions. One is to show that when interfering stimuli are eliminated, conventional Pavlovian CSs like lights and tones can become conditioned even with long delays. The second is to show that taste-aversion learning can be impaired if interfering tastes are allowed to occur before the initial taste or between it and the US.

Both types of experiment have been done. For example, the experiments by Lett, discussed in the previous chapter, have shown that conditioning (her experiments were operant rather than classical, but this is not a critical point) with conventional visual signal stimuli can be successful with long delays if the stimuli are highly salient and if interfering stimuli are minimized.

The effects of interference on taste-aversion learning have been shown in several studies by Revusky (1971). In one study, four groups of rats were allowed to drink 2 ml of saccharin solution; then, 15 minutes later, different groups were allowed to drink water containing differ-



**Figure 14.4.** Preference of rats for drinking saccharin solution vs. a novel coffee solution, as a function of the concentration of vinegar drink in between original experience with saccharin and lithium-chloride-induced sickness. The groups that tasted the stronger vinegar solutions showed greater preferences for (i.e., less aversion to) saccharin. (After Revusky, 1971.)

ing concentrations of vinegar (a novel taste). One hour after the saccharin drink, all the rats were made sick by an injection of lithium chloride. Three days later, each group received a series of preference tests in which the animals were offered a choice between saccharin water and dilute coffee. Figure 14.4 summarizes the results: The weaker the intervening vinegar solution, the less saccharin water drunk (i.e., the stronger the aversion to saccharin). In a second experiment Revusky showed that the vinegar interferes with the development of saccharin aversion even when it precedes, rather than follows, the saccharin. He also showed that the interfering effect of the vinegar is much reduced if it is a familiar rather than a novel taste.

There are two possible reasons for the interfering effect of the interpolated vinegar taste in Revusky's first experiment: (a) The vinegar coming after the saccharin interfered with the animal's recall of the saccharin (retroactive interference); consequently the animal could not associate the saccharin with the sickness one hour later. (b) The vinegar was closer in time to sickness than the saccharin, so that the animal associated the vinegar rather than the saccharin with sickness. This would be an example of an *innate prior*, since recent events are much more likely to be the cause of something than more remote events. This dominance of one potential CS over another is termed *overshadowing*. In either case, we might expect the rats to form an aversion for vinegar, and this presumably occurred, although Revusky did not test for it.

The proximity explanation, (b), is ruled out by Revusky's second experiment, in which the vinegar interfered even if given before the saccharin. We have already seen a comparable

result, however, in the discussion of temporal control in Chapter 14. Figure 14.1, for example, shows that temporal control by a brief, neutral stimulus is impaired by a preceding food delivery. I argued that this reflects the greater importance of food, a hedonic stimulus, which preempts the animal's attention after a few seconds and abolishes temporal control by the neutral stimulus. The animals in Revusky's experiment had only one opportunity to learn, so that the proximity of the saccharin to the sickness (*postremity* is the technical term) was outweighed by the higher prior probability that vinegar, a salient, novel taste, was the cause of later sickness. The animal had just two things to go on in the first experiment: saccharin, though familiar, was closer to sickness than the novel vinegar, but the vinegar was both salient and novel. In the event, given one trial, the vinegar won out.

A couple of thought experiments are instructive here. We can ask: what would have happened had both vinegar and saccharin been familiar? And, what would happen if the experience, vinegar→saccharin→sickness, were repeated several times?

Common sense, and data from other experiments, can answer the first question at once: when the two events are equal in every other respect, the rats will avoid the taste closer in time to the sickness. The innate properties of memory, of which retroactive interference is one, reflect facts about the causal properties of the world: in particular, the fact that a more recent event is more likely than a more remote event to be the cause of a present event (sickness).<sup>19</sup> The temporal resolution of memory *should* reflect the probable usefulness of the information remembered. As in an office filing system, the older an event, and the more similar events that have occurred subsequently, the less useful information about it is likely to be — and the less need to retain a record of it.

The effect of repeated experience is also easy to predict: even if the vinegar is novel at first, with repeated experience it necessarily becomes less so, and then the advantage of postremity is likely to favor the more recent stimulus, saccharin. Eventually (and if it has no other basis for preference) the animal is likely to avoid saccharin, the stimulus closest to sickness.

The discussion of the temporal resolution of memory in the previous chapter suggests one final thought experiment. In Revusky's second experiment, the vinegar occurred 15 min before the saccharin, and the animal was made sick 60 min after that. Suppose the rat had been made sick only 10 min after the saccharin? There is little doubt that *some* time, shorter than 60 min, but probably more than a few seconds, could be found such that the familiar saccharin would be avoided, not the preceding, novel, vinegar.

The general conclusion is that the animal's "choice" of which stimulus to avoid is the outcome of a delicate balance in which all the potential factors that could determine cause — stimulus type, novelty (i.e., past experience with each stimulus) and postremity — are combined in an equation whose weighting factors are determined by heredity.

**The order of learning.** Learning that occurs on one trial must require at least some acquisition of information ("change of state" in the sense discussed earlier) before the occurrence of the reinforcer. The animal must record the novel taste if he is to acquire an aversion to it based upon delayed sickness (this is like color learning in the bee example). On the other hand, this information need not dominate; and additional processing may go on after a reward has occurred. For example, in one experiment<sup>22</sup> rats were given a single electric shock while taking a pellet from a small hole in the center of a novel stimulus display. After a brief delay (different for different groups) a spiky toy rubber hedgehog was suddenly introduced into the chamber. The next day the rats were given choice tests in which the hedgehog was compared with the conditioned stimulus and the location in which the shock had been delivered. By a number of measures, the animals tended to avoid the hedgehog most, even though it had occurred *after* the shock; moreover, the degree of aversion was inversely related to the delay between shock and appearance of the hedgehog — indicating that temporal proximity to the shock was a key factor. This is an example of *backward conditioning*, something which is not supposed to occur (and

usually doesn't) and makes little causal sense — if B occurs after A it cannot be the cause of A.

From an ecological point of view, of course, the animals were acting perfectly sensibly. All conditioning is a process of inference. It is impossible to be certain of the validity of a causal inference on the basis of a single instance. Animals make mistakes and predators are often concealed before they attack. The prior probability (from the rat's vantage point) that a striped pattern (the CS in this experiment) will be dangerous is small; the prior probability that a lifelike dummy will be dangerous is much higher. Repeated experiences (denied the animal in this experiment) would serve to correct mistaken inferences drawn on the basis of a single experience. From this point of view, the animal's "guess" that the spiky hedgehog was the source of his pain is perfectly sensible.

This result also shows that considerable processing goes on after the US has occurred. Here the rats evidently acquired some fear of the hedgehog; it is likely that a subsequent trial in which presentation of the hedgehog was soon followed by shock would rapidly confirm this fear and produce long-lasting avoidance. Thus, the first conditioning trial not only produces some "conditioning," it also creates a set of CS *candidates* to which the animal can attend on later trials. I return to this point in later discussion of theories of classical conditioning (see also note 22).

Other experiments have demonstrated the occurrence of post-US processing in conditioning. For example, Wagner<sup>20</sup> and his associates studied the effects of post-US events on the conditioning of the rabbit nictitating membrane. In the first phase, the animals were trained with a standard CS-US combination until they developed a reliable conditioned membrane response. In the second phase, the animals were divided into two groups. Both groups continued to receive the standard CS-US trials, but after each trial they also received one of two types of "interference" trials. Both types involved a second CS: either a previously established, different CS+ (i.e., a stimulus that had been paired with shock), or a previously established CS- (i.e., a previously established safety signal). No US occurred on these interference trials. Thus, for the added-CS+ group they involved a violation of expectations; for the added-CS- group they did not. The animals that always received the incongruous post-trial event (CS+ → no shock) learned (i.e., developed a membrane response to the standard CS) more slowly than the animals that received the unsurprising sequence (CS- → no shock). Evidently the post-US "surprising" event interfered with essential retrospective processing (termed "rehearsal" by Wagner) of the preceding trial.

**What is learned?** The effect of any event on an animal usually depends upon its current state of knowledge. For example, some years ago, a colleague (Davis, 1975) was puzzled by the observation that a flock of feeding pigeons would sometimes be startled into general flight by the flight of one individual, whereas at other times, the takeoff of one bird had no effect. Perhaps pigeons have a special "alarm flight" that alerts others to the presence of danger and causes them to take flight? we thought. Reasonable as this hypothesis sounds, painstaking experimental work failed to find any distinguishing characteristic of "alarming" vs. "non-alarming" takeoffs. The answer to the puzzle turned out to be not the property of the takeoff, as such — there was nothing to differentiate alarming from non-alarming flights — but in the "intention movements" birds make *before* they take off. A takeoff preceded by the proper intention movements (head-raising, small wing movements) had no effect; but an *unsigned* takeoff — a takeoff for which the flock was not prepared — caused alarm and the flight of other birds.<sup>24</sup>

Birds in a flock closely monitor one another's behavior. Clearly the intention movements of a bird about to take off change the state of his fellows in such a way that his flight causes no alarm. In colloquial terms, the intention movements lead the other birds to *expect* a takeoff and only unexpected takeoffs cause alarm.

This same principle, that the effect of something depends upon the animal's state of knowledge, should obviously extend to learning. Thus an answer to the question, What are the

necessary and sufficient conditions for learning? also requires answer to the question, What is learned? Applying the rule suggests the general principle: *animals learn only when something violates their expectations*, or when they have no expectations (as in a novel situation). All the examples of reinforced learning I have discussed fit this generalization: When two tastes have an identical temporal relation to poisoning, conditioning develops only to the novel taste. When a *compound stimulus*, such as a tone-light combination, is used in one of the standard classical-conditioning arrangements (cf. Chapter 5) the amount of conditioning that takes place to each element depends upon what the animal knows about each. For example if, in a conditioned suppression (CER) paradigm, the tone had previously been paired with electric shock and the light was novel, no conditioning will take place to the light: The shock is already predicted by the tone, so no expectation is violated, and no new learning takes place. This is termed *blocking*. The Wagner et al. experiment is an explicit test of the role of surprise in conditioning, and I will discuss some others in a moment.<sup>21</sup>

In general, greater the violation of expectation, the larger the learned change in behavior. For example, if two stimuli are presented in alternation with shocks occurring only in the presence of one of them (CS+), then the other (CS-) becomes a “safety stimulus”, i.e., a stimulus that signals the absence of shock. If such a CS- is now presented with a novel stimulus, *X*, as a compound and paired with shock, then *X* acquires a larger increment in its power to suppress lever pressing than when presented alone (so-called “superconditioning”). The expectation here was presumably for a reduction in the probability of shock, so that the occurrence of shock represents a larger violation of expectation than if the CS- stimulus had no prior significance. It is interesting that the increment in conditioning goes to *X* rather than to the safety signal, but this presumably reflects the Bayesian process we have already encountered, since CS- has already been established as a safety signal.

This same process accounts for the differential conditioning to taste, when the US is sickness, or to audio-visual stimuli, when the US is electric shock. The (innate) priors (in rats) favor taste as the predictor of sickness, and an audiovisual stimulus as a predictor of pain.

The first occurrence of a surprising event is critical. In a blocking experiment by Mackintosh and his collaborators (Mackintosh, Bygrave, & Picton, 1977), five groups of rats were pre-trained with a light-shock combination in a standard CER arrangement. In the second phase of the experiment, a tone was added to the light, and different groups received either one or two additional trials, and a surprising additional shock just after the tone-light compound on the first or second of the additional trials. The six possibilities, five of which were tried, are summarized in Table 14.2. The entries in the table give the rank ordering of suppressive effect of the tone alone, as shown in a subsequent test. The group that received two tone-light trials with an extra shock on the first trial showed most suppression by the tone (rank 1). Next came the two groups that received two additional trials, but no additional shock on the first tone-light trial — evidently the added shock on the second tone-light trial had no immediate effect. Last were the two groups that received only one tone-light trial.

The added shock on the second tone-light trial evidently had *no direct effect*: The suppression after one trial was the same whether a shock was added or not (groups 1-neither and 1-1 in Table 14.2 showed equal suppression of rank order 3); and suppression after two trials was the same whether a shock was added or not, providing it was added only to the second trial (groups 2-neither and 2-2 were equal, rank 2). But the added shock did serve to create a candidate, the

|                             |   | Trial when added shock occurred |   |   |
|-----------------------------|---|---------------------------------|---|---|
|                             |   | Neither                         | 1 | 2 |
| Number of tone-light trials | 1 | 3                               | 3 | — |
|                             | 2 | 2                               | 1 | 2 |

**Table 14.2** Rank order of conditioning for five groups in an experiment by Mackintosh, Bygrave, and Picton (1977)

tone that could be confirmed on a subsequent trial. Thus, the group that received *two* tone-light trials with the added shock occurring on the *first* trial (group 2-1, rank 1) showed the greatest suppression by the tone.

There are some puzzles left by this rather elaborate experiment, however. Does effectiveness of the added shock on the first trial depend upon a general increase in attentiveness that carries over to the second trial and causes the animal to attend to the presence of the tone on that trial, or is it really a retrospective or “backward-scanning” effect that causes the animal to recall the presence of the tone on the first trial (as suggested by the hedgehog experiment)? In the latter case, the second trial only serves to confirm a prior “hypothesis.” A sixth group, in which the tone is omitted on the first added-shock trial but not the second is necessary to answer this question.

Learning is initiated by violation of expectation — *surprise*; but the change that subsequently occurs depends on Bayesian factors. For example, blocking can be abolished by either increasing or decreasing the strength of the US. Recall, blocking refers to the outcome of a two-part experiment: In the first phase, a conditioned stimulus, *A*, is paired with an unconditioned stimulus (usually electric shock, in the CER paradigm) until it has acquired maximal power to suppress lever pressing. In the second phase, a neutral stimulus, *X*, is presented along with *A*, and the compound, *AX*, is paired with the US as before. *X* acquires little power to suppress responding, presumably because the occurrence of shock is perfectly predicted by *A*. If the strength of the US (the size of the shock or number of shocks) is either increased *or decreased* on the first *AX* trial, however, then *X* acquires some suppressive power. Presumably either of these changes violates the animals’ expectations, based upon previous *A*→US pairings, so that some learning can now occur.

Consider the case where the US is incremented. The new learning could take any of three forms: an increment in the suppressive power of *A*; an increment in the suppressive power of *X*; an increment in both. There are comparable alternatives for US decrement. Yet in both cases, all the change (which is always an increment in suppression) takes place in the *novel* stimulus, *X*. This is perfectly Bayesian: the significance of *A* has already been established; a US change coincident with the introduction of a novel stimulus, *X*, is most plausibly attributed to *X*. Since shock continues to occur, and *X* does not suppress at first, the change is always an increase (Dickinson, Hall, & Mackintosh, 1976).

This separation between the surprise that initiates learning, and the Bayesian processes that determine its direction, can have unexpected implications. In one experiment (Hall & Pearce, 1979; see also Pearce & Hall, 1980) three groups of rats received standard CER training. One group (*T*→*s*) received pairings of a tone with weak electric shock; a second group (*L*→*s*) received pairings of a light and weak shock. The third group (*T*→*0*) just heard the tone. In the second phase of the experiment, all three groups received pairings of the tone with a strong shock. The question is: Which group will learn most rapidly to suppress lever pressing in the presence of the tone?

The second phase is surprising for all groups, because of the increase in shock intensity, so all will learn something. But the tone is novel only for the light group (*L*→*s*); hence, these animals should learn fastest to suppress to the tone. The tone has some prior significance for both the tone-alone (*T*→*0*) and tone-weak shock (*T*→*s*) groups. Hence, one might expect less rapid conditioning to them. Moreover, the tone-alone group has had an opportunity to learn that the tone signifies nothing, whereas the tone-weak-shock group has at least learned that the tone predicts something about shock. Hence the expected order of conditioning is *L*→*s*>*T*→*s*>*t*→*0*, which is what occurred. This result is puzzling at first, because the group with prior experience of tone-weak-shock pairings learned about the contingency between the tone and the strong shock more slowly than a group with no prior tone-shock experience. Yet it follows directly

from Bayesian inference: A change in the US is more plausibly attributed to a novel stimulus (about which nothing is known) than to one with a prior history, even a prior history of association with (weak) shock. In the next chapter I describe a way of representing prior probabilities that makes these predictions in a more cut-and-dried way than this purely verbal argument.

Our techniques for answering subtle questions about when particular things are learned in the course of classical-conditioning experiments obviously leave something to be desired. Even quite simple questions, such as whether the effect of a surprising event (e.g., added shock in the Mackintosh et al. study) is immediate or delayed, lead at once to large experiments with many groups of animals. Yet as large as they are, there never seem to be enough groups to answer all the questions that arise. Moreover, it is not always clear that individual animals all behave as the group does. Nonetheless, the role of surprise, the importance of the first trial of any change, and the occurrence of processing both before and after the presentation of a US, are amply demonstrated.<sup>22</sup>

### *Expectation and Classification*

Learning evidently depends on surprise; and surprise implies a discrepancy between experience and the animal's representation of the world. To understand learning, therefore, we need to understand something about how animals represent things.

I argued in Chapter 10 that animals learn about objects and situations, not disembodied dimensions or features. One way to look at an object is as a stable set of properties or attributes. For example, dog Fido has a certain size, sex, color, temperament, physical location, number of legs, and so on. Some of these attributes are more variable than others (location and size are more variable than number of legs and sex, for example), and some are more *differential* than others: all dogs have four legs, but only Fido has his particular coat color and pattern of likes and dislikes.

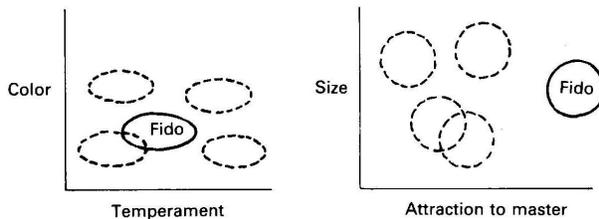
But which of these features *really* defines Fido? This problem has exercised philosophers ever since Plato raised the problem of universals (What makes a table a table? Is a chair being used as a table really a table?). Yet I believe there is a pretty good answer to it, at least within the limited confines of animal behavior and artificial intelligence. It is captured by the epistemological aphorism: "If it walks like a duck, quacks like a duck, and lays eggs like a duck, it must be a duck." In other words, an object is defined by its attribute set. In a moment I will show one way to represent a set of attributes so that objects are easily differentiated from one another.

If an animal can distinguish just ten grades for any attribute (color, size, etc.), and it can assess say 10 different attributes, then it is potentially capable of uniquely identifying  $10^{10}$  different objects. This is a very large number; to store information about so many things would require much memory and be highly inefficient. Most of the potential entries would be blanks (i.e., the individual will never have encountered objects with just that set of properties), and there is no need to discriminate even among all nonblank entries because many items will be functionally equivalent.

The memory load can be substantially reduced if the animal resorts to *classification*: Fido may be unique, but there is no need to distinguish very carefully among the myriad other dogs with which one comes into contact. Certainly little purpose would be served by remembering every one of them individually. Moreover, as we have already seen, not all attributes are equally valid identifiers: Fido's location varies from minute to minute; it's almost useless as an identifier. His temperament varies from day to day over a narrower range; it's a bit better. Fido always has four legs, but so do many other dogs. On the other hand; an object with only two legs is very unlikely to be Fido. Fido answers to the name "Fido"; rather few other dogs do so. And Fido knows and loves his master, a quality shared by few other organisms, perhaps. These two attributes are obviously good identifiers.

All these properties can be captured very simply in a multidimensional Euclidean representation of the sort discussed in Chapter 10. We can get the basic idea just by taking 2-dimensional “slices” like those in Figure 14.5. The left panel shows a slice along the two “dimensions” of color and temperament (of course, color is itself multidimensional, but this is just an illustration). The cigar-shaped *region* corresponds to the representation of “Fido”; the dashed regions correspond to other dogs. The fact that the cigars are narrower in the vertical direction indicates that color is less variable (within individual dogs) than temperament. The fact that dashed regions are equally scattered through the space surrounding Fido indicates that neither Fido’s color nor his temperament is unique.

The right panel in Figure 14.5 shows a slice through the space along the dimensions of size and “attraction to master”. Here Fido is closer to being unique: while there are many other dogs of similar size, none rivals him in affection for Fido’s master. But notice that this method of representation can differentiate even among objects that are not unique on any single dimension.



**Figure 14.5.** Two-dimensional slices through the regions of attribute (semantic) space containing the representation of dog Fido. Left panel: a slice parallel to the color-temperament plane. Right panel: a slice parallel to the size-”attraction to master” plane

The fact that objects are represented as regions (rather than points) takes account both of the variability of attributes and of organisms’ limited storage capacity. You can think of the 2-dimensional slices in Figure 14.5 as being actually divided into discrete cells (like the memory cells of a computer), so that objects falling within the same cell are classified as equivalent. Or imagine the organism viewing the space with a metaphorical “mind’s eye” of limited acuity. The point is that the resolution of memory is limited: no two objects can be closer than some mini-

imum distance in “category space” without being treated as equivalent.

**Learning and classification.** As I pointed out in Chapter 10, the dimensions of semantic spaces such as those in Figure 14.5 do not correspond precisely to physical dimensions, although there is often a loose relation. One dimension, often the most salient, is completely nonphysical: good-bad. When the objects to be represented have any hedonic significance at all (animals or odors, for example), human similarity scaling experiments always yield a space in which *value* is one of the major dimensions.

Value is something that must usually be learned.<sup>23</sup> Classical conditioning can be regarded as a procedure that leads an animal to assign a positive or negative value to an initially neutral stimulus. Unconditioned stimuli such as food, water and electric shock of course differ in more than one way (food has a positive value for a hungry animal, as water does for a thirsty one, but their effects are not identical; cf. Chapter 7), and this difference can be accommodated within a multidimensional semantic space. Whether represented on one dimension or several, value is obviously an important, perhaps the most important, attribute of any object; and it is an attribute generally (but not invariably) acquired through experience.

A spatial representation has several advantages. It provides us with a potentially precise meaning for the term *expectation*, and it suggests that stimulus dimensions should be linked during learning. The bee learning again provides a simple example.

Learning about a particular food source corresponds to locating a *vector* of attributes  $A(c,s,p,l)$ , where  $c,s,p$  and  $l$  represent the color, scent, pattern and location of the object, along the value ( $v$ ) dimension. In other words (this model suggests), each object is represented by the bee as a point,  $A(c,s,p,l,v)$ , in  $N$ -dimensional space, where  $v$  is its value. Each learning trial with a given, constant, food source gives the bee additional confidence that an object with these coor-

dinates exists. One can think of the first learning trial as creating a single point corresponding to the object  $A(c,s,p,l,\dots v)$  in the appropriate region of the animal's memory. Each succeeding trial adds additional points until finally it reaches the level at which the animal is certain that a food source, with such and such properties, has a permanent existence.

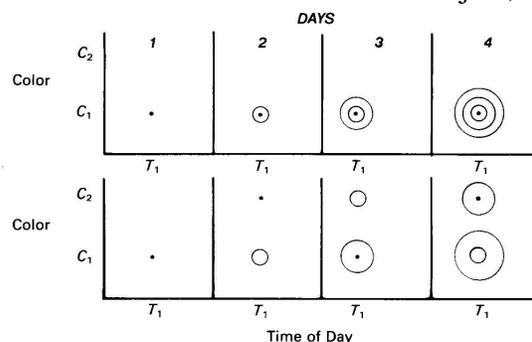
A more concrete model is as follows. Suppose that every valued outcome corresponds to a point in N-dimensional space; thus each such point has four coordinates  $(c,s,p,l)$  — this is not too different from the model already proposed for rat spatial learning in Chapter 12. Old information is less valuable than new information. This fading with time was represented by the perspective metaphor, and trace decay, in the previous chapter. But since I want something easy to visualize, let's just say that the point representing each valued event is like a spot of ink, the more valuable the original event, the denser the original ink spot. The ink diffuses at different rates along the different dimensions. The density of the ink at any time then represents the net value of that object (region of semantic space) at a given time. The spread along the odor axis will be slow, since odor is learned fast; spread along the time-of-day axis will be more rapid, as time-of-day is learned more slowly; spread along the value axis should be slowest of all. Thus, after passage of time, the animal will still be able to recall the odor and value of the object, but its color, time of day and other variable attributes will be less certain.

However it is expressed, the rule that the *rate at which attributes are forgotten is inversely related to their importance* to the animal seems to be a general one. Attributes of high a priori (innate) validity are learned rapidly and forgotten slowly.

After several learning trials, an equilibrium between the rates of diffusion and the rate of replenishment (conditioning) will be reached, and the animal will be said to have learned about the food source.

In a system of this sort, the rate at which learning proceeds will depend upon the variability of the object; learning about one attribute will depend upon the variability of all attributes. For example, if the time of day at which the bees are fed remains constant, but the color varies from day to day, then the learning of *both* color *and* time of day is retarded. This is because a change in *any* dimension moves the point representing the learning to a new region in the space. i.e., changes the object. Unless the new region is adjacent to the old, there will be no cumulative effect.

This dimensional interaction is illustrated in Figure 14.6. The top row of panels shows the appearance of the space after successive trials in which both color and time of day are constant. Each concentric ring represents an increment of strength. Thus on day 1, a point is created in the space at coordinate  $(T_1, C_1)$ . After the second conditioning day, it has diffused to the ring shown in panel 2, and a second dot is added. By the fourth day, there is a dot and three diffusion rings at that point. The bottom panels show the comparable changes when colors alternate from day to day, but time of day remains constant. Clearly, after four days, the concentration at the point  $(T_1, C_1)$  after consistent reinforcement (top panels) is much greater than the concentration at either point after inconsistent reinforcement (bottom panels). Granted that the bee's behavior is determined by the regions of highest concentration (it seeks the most promising sites) the accuracy with which it chooses either the color *or* the *time of day* in the variable case must be less than its accuracy in the constant case.



**Figure 14.6.** Diffusion model of object learning. Each encounter with object  $(T_1, C_1)$  or  $(T_1, C_2)$  adds a diffusing spot at that point in attribute space. Top panel: lines of equal concentration after 1-4 days of training with both attributes constant. Bottom panel: lines of equal concentration after 1-4 days of training with rewarded color alternating each day.

Bees show precisely this kind of linkage between attributes. For example, in an experiment by Bogdany (1978) a colony of honey bees was trained to feed on sugar water at an artificial feeder some 400 m from the hive. At the feeder site, 16 color cards, each with a feeder in the center, were arranged in a square. Only one feeder contained sugar water. Two conditions were compared: Both color and time of day constant; and time constant, color variable.

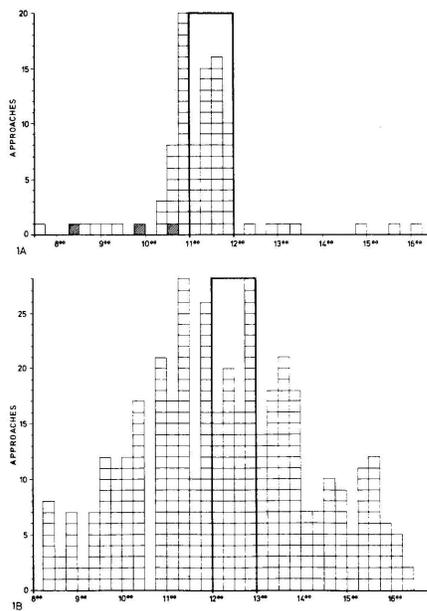
In both conditions, the bees were trained for five days. In the both-constant condition, the bees were rewarded for picking out the color blue (the other 15 cards were different shades of gray) between 11:00 and 12:00 h each day. On day 6, the feeding station was set up and observed until 18 h, but all 16 feeders were empty. Approaches to the blue card only were counted as correct.

The results are shown in the top panel of Figure 14.7. The bees were very accurate: 96% of approaches (white squares) were correct; only 4% (gray squares) were not to the blue square. The animals' timing was also very good: almost all the approaches were within 15 minutes of the 1-hour training "window" (11:00-12:00 h).

In the time-constant, color-variable condition, the full feeder was on a different color card each of the five training days. The bottom panel of Figure 14.7 shows the test results. Color accuracy must obviously be poor here, and is not shown. Time accuracy is also much impaired, however. As the figure shows, the majority of approaches were outside the training window.

Other conditions of this experiment showed similar linkages between time and odor and between all three attributes, time, color and odor.

I am not aware of any experiments of precisely this sort with birds or mammals. A few



**Figure 14.7.** Top panel: number of bee approaches to an artificial feeder at different times of day after training with constant color and fixed time (11:00-12:00 hr). Bottom panel: number of bee approaches to the feeder at different times of day after training at a fixed time, but with variable feeder color. (From Bogdany, 1978.)

years ago, there was much interest in a related hypothesis: that attention of pigeons or rats to one stimulus dimension (measured as the slope of the generalization gradient; cf. Chapter 10) might be *inversely* related to attention to another. The results of this work were inconclusive; some people found the expected relation, while others found the opposite (see Mackintosh, 1977, for a review). The present argument suggests that one might well expect a positive relation if the two dimensions could be integrated as a single object (a color and a pattern located on the same response key, for example); whereas the expected negative relation might be found if the two dimensions cannot be integrated in this way (a tone and a spatially separated light, for example).

The effect of pairing a valuable outcome with a complex stimulus obviously depends critically on how the stimulus is represented by the animal. For example, suppose we pair a loud tone and dim light with shock and then later test the tone and light separately for their suppressive effects. You know the result: the loud tone will overshadow the dim light — the light will have little or no effect. But that's because tones and lights do not form integral combinations: "tone-light" isn't in any sense an object for a rat. It is otherwise with tastes and odors: Garcia and his students (1981) have shown that if rats are made sick after experiencing a weak odor and a

strong taste, both odor and taste are later avoided. In another experiment, Garcia's group showed that hawks can't learn to avoid mice of a particular color when poisoned after eating them. But if black mice (say) *taste* different from white, black mice are violently rejected. In terms of my

ink model, a given taste and odor, or taste and mouse color, seem to represent a single region in semantic space, whereas a tone and a light represent two separate regions.

Different species seem to differ in the dimensions that combine integrally or separably; indeed, as I argued earlier, much of what we term “species differences” seems to boil down to differences in the way animals represent their worlds — and these differences, in turn, no doubt reflect their different niches.

**Recognition and expectations.** I am arguing that the sensible attributes of an object, its color, location, odor, and so on, identify a region in the animal’s semantic space. If that region has a known value, the animal can be said to *recognize* the object as a familiar one. This value constitutes the animal’s *expectation* about the object. If, on a particular occasion, the appropriate positive or negative consequences fail to occur, there will be a discrepancy between the stored location of the object on the value axis, and the location indicated by the new experience. It is this discrepancy — violation of expectation — that will tend to change the animal’s representation of the object, that is, produce new learning.

## SUMMARY

Obviously, a great deal remains to be understood about the acquisition of knowledge through conditioning. My purpose in elaborating this speculative scheme is to clarify the kinds of questions that must be answered before we arrive at full understanding. Terms like “memory,” “expectancy” and “concept” have gained currency partly in reaction against the simplistic stimulus-response psychology that once dominated the study of animal learning. However, these terms do not go far enough — and may mislead by encouraging a reliance on empathy with the animal. The terms must be given an exact, mechanical meaning if they are to be of much scientific use. We need to know both the form in which the world is represented by the animal, and the rules by which experience changes its representation. We must also recognize that these two things may not be separate. In simple animals, the rules of action may be the only form of representation the animal has. But in mammals and birds — and bees — there does seem to be a point in separating the two.

Classification based on some kind of semantic space looks like a good working hypothesis in answer to the representation question. The facts of song learning and imprinting fit into this kind of scheme, as does much of what we know about conditioning. In the song learning of white crowns, for example, a region of semantic space appears to be reserved for conspecific song. Anything within the boundaries of that region is accepted, and is stored as the object or template by which later song will be guided. Songs lying outside the region are rejected. The imprinting object is similarly constrained to a broad region within semantic space.

The rules for changing the representation are broadly Bayesian, with priors initially determined genetically. However, since a discrepancy between the current representation and experience is what drives any change, these priors also change with experience.

In this chapter I have dealt largely with learning that takes place on one or a few trials, and I have devoted little space to the action guided by that learning. The discussions of bee learning do not conflict with the emphasis on classical, rather than instrumental, conditioning, because bee foraging seems to be rigidly tied to what they have learned of the value of food sources. For a bee, to know a good flower is to seek it out: the action does not have to be learned, only its object. We will see a few other such rigid links between knowledge and action in the next chapter, as well as cases where behavior is not forced by the nature of valued objects. I shall also there consider in more detail the problem of sampling and the nature of multi-trial learning.

---

## NOTES

1. Lifespan is obviously not the only factor that determines a species' ability to learn; its *niche* is also critical. Galapagos tortoises, for all their charm, are not great intellectual performers. Presumably their combination of large size, longevity and stupidity is well suited to a stable and relatively undemanding environment. Size is also not critical. All ethologists are familiar with the "law" that no matter how clever the trick learned by any given mammalian species, behavior of comparable difficulty can be demonstrated in bees. The ways in which bees cope with their limited computational capacities will be discussed later. Nevertheless, other things being equal, longevity and large body size conduce to the evolution of learning ability.
2. The term *precocial* refers to the relative maturity of the birds on hatching, and distinguishes newly hatched chickens, ducks and the like from newly hatched songbirds, pigeons and doves. The young of *altricial* species like these are utterly helpless at birth. As one might expect, these differences relate to the species' niche: mobile species with primitive nests tend to have precocial young; species that provide a more protected early environment have altricial young. There are comparable differences in mammals: the young of grazing animals such as deer, guinea pigs and antelope "hit the ground running," whereas young rats, dogs and cats are blind and helpless at birth.
3. Much of the work on imprinting has been done by German animal behaviorists (ethologists), although the very earliest report seems to be by Douglas Spalding (1873), tutor to the infant Bertrand Russell, the early twentieth century English philosopher, aristocrat and political activist (see Boakes, 1984, for an interesting, brief account). The original German work is due to Heinroth (1911) and Lorenz (1935; reprinted and translated in Lorenz, 1970). Some of this work is not yet translated into English: for example, the studies on shrews by Zippelius (1972), and the work on ducks by Schutz (1971; see note 5). The studies of auditory imprinting in ducklings are due to Gottlieb (e.g., 1971), and the work on maternal attachment in macaque monkeys is by Harlow (e.g., Harlow & Harlow, 1965). The studies of imprinting in goats are by Klopfer and Gamble (1966). Hess (e.g., 1973) has contributed much of the recent knowledge about imprinting in ducklings. A good general review of imprinting is Bateson (1974); Immelman (1972) has discussed the relation between filial and sexual imprinting, arguing that the two are largely independent.
4. This may not always be true, in the sense that for some species there may be imprinting stimuli more effective than the natural one. Such *supernormal stimuli* (Tinbergen, 1948) are not uncommon even when no learning at all is required. For example, male silver-washed fritillary butterflies (*Argynnis paphia*) are stimulated to court females by the flickering movement of their wings, and this response requires no learning. The natural frequency of flicker is about 8-10 Hz, yet in laboratory experiments Magnus (1958) found that male butterflies would court more vigorously to stimuli flickering at up to 140 Hz — frequencies much higher than anything encountered in the animals' natural environment. There are many other examples: birds preferentially retrieving extra large eggs, gull chicks preferring to peck at artificial bills longer, thinner and more brightly colored than their parents', and so on.

Many if not all of these instances seem to be explicable in terms of *asymmetrical selection pressure* (see Staddon, 1975b). Thus, the wings of predatory birds flicker at a rate lower than the flicker rate of the female fritillary, so that an error in the downward direction may cost a male butterfly dearly. Higher frequencies are never encountered, however, so that an error in

this direction exerts no selection pressure. There is an obvious resemblance, therefore, between the conditions that produce peak shift in an operant conditioning experiment (cf. Chapter 10) and those that produce supernormality in nature.

There is every reason to expect that the selection pressures that favor certain kinds of imprinting object are of exactly the same sort as those that produce innate reactions to stimuli. Hence we may expect to see instances of supernormality in both cases.

5. The mallard experiment was by Schutz (1971), described in Fantino & Logan (1979, p. 366). It is interesting that the selection by females of their proper mates was apparently *independent* of rearing condition: the females always chose males of their own species. There are two kinds of explanation that can be offered for this difference. One relates to the sexual dimorphism of ducks, and the bright plumage of the males: perhaps the females simply have an easier job because the males of different species look so different. Since the task is easier, perhaps it makes more sense for it to be “pre-wired.” I’m not so sure, given the earlier arguments (Chapter 1) against “simplicity” as a criterion for innateness: some pretty complicated behavior patterns are innate, and some pretty simple ones depend heavily on learning.

A second possibility was foreshadowed in Chapter 1: The argument rests on two premises: (a) That the selectivity of all imprinting implies that species recognition, even in these cases, was at one time innate and independent of particular experience. And (b) that the reproductive cost to a male of making a species identification “error” (in its choice of target for copulation) is much lower than the cost for a female. The greater cost to the female is a consequence of the greater reproductive cost of egg-production as compared to sperm production, and the lesser parental investment of males (this point is a cornerstone of *evolutionary psychology*, the approach to human behavior that has evolved out of the *sociobiology* of the 1970s. A good popular exposition is Pinker, 2002). If hybrids have lowered fitness, then a female duck inseminated by a male of another species is committed to producing a lower-fitness offspring and suffers a delay (an *opportunity cost* in the language of economics) in her ability to produce additional high-fitness (non-hybrid) offspring. The greater fitness cost of error for females suggests that their mechanism for mate choice should leave less to chance. If innateness of response is the ancestral condition, strong selection for female selectivity would tend to preserve the innate process in females, but might permit the degeneration represented by imprinting in the case of the males of species where the male plays little or no parental role.

This premium on selectivity by females may also account for the distinctive plumage of the males, since it pays a male to identify itself as strongly as possible as *not* belonging to a species different from the female with which it is attempting to copulate. Perhaps the variety and brilliance of plumage and other accoutrements in males of many bird species is a consequence of the selection pressure on them to look as different as possible. Recent work also suggests that bright plumage in the male may also advertise his freedom from disease and malnutrition — proxies for evolutionary fitness. The need for camouflage (*crypsis*) tends to produce convergence in the drab appearance of females of different species sharing the same habitat.

Several other mechanisms have been suggested in recent years to show that differential reproductive investment provides a basis for Darwin’s (1871) suggestion that *sexual selection* is responsible for most cases of sexual dimorphism (see reviews in Krebs & Davies, 1978; Shettleworth, 1998).

Not all birds are sexually dimorphic: Male and female geese, doves and pigeons, for example, are essentially identical, and unspectacular, in appearance. Cross-fostering experiments have shown that in doves, females as well as males learn their species identity through early experience, but the altricial nature of these birds and the long rearing period make this learning less prone to error than the imprinting of precocial ducks and chickens. The symmetry here goes along with the symmetry of parental investment: Doves form permanent pair bonds and both

parents are equally involved in the incubation and feeding of young.

6. Perception of flicker and perception of movement are closely related at the neural level; hence the behavioral equivalence of flicker and movement in imprinting studies.

7. Mulligan (1966). Kroodsma (1977) has discussed more recent evidence on the development of song-sparrow song; see also papers in Konishi & Hauser (1999).

8. The now-classical work by Marler and his associates on song development in white-crowns, and the earlier, related work by Thorpe on chaffinches, is widely reviewed. Good summaries appear in Hinde (1969), Thielcke (1976), Kroodsma (1978) and Catchpole (1979).

9. The term *passerine* refers to the *Passeriformes* or “perching birds”, the order that contains most of the songbirds. Finches, swallows, tits, sparrows, and thrushes are all passerines.

10. A number of evolutionary hypotheses have been offered for the enormous variety of bird-song. Several identify song variety as a weapon in territorial defense. The most colorful of these is the “Beau Geste” hypothesis, which likens the defending bird to Beau Geste, of French Foreign Legion fame. Legionnaire Geste single-handedly defended a fort, in which all but he had perished, by propping up the corpses in the battlements, and running from one to another firing each man’s weapon himself. By singing different songs, from different places, a defending bird may convey the impression that there are many residents in a given area, thus deterring potential intruders more effectively than he could if he appeared as one. Other functions for birdsong include (obviously) species identification, courtship and, as an ancillary to courtship, a signal for evolutionary fitness. The evolutionary reasons for the form and diversity of birdsong are still a matter for active debate. See reviews in Catchpole (1979), Konishi & Hauser (1999) and Krebs and Davies (1978).

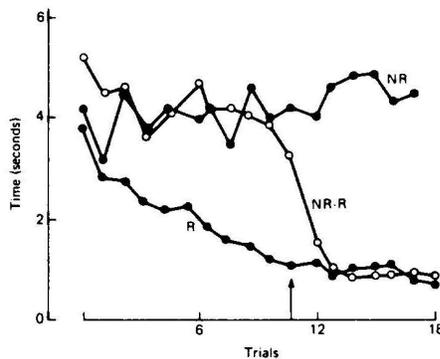
11. The literature on bee behavior is large and much of it is in German. For a clear, readable summary of what is known of their mental life see Gould and Gould (1999), which also contains a number of other references.

12. In many early works on learning, it was customary to express wonder at the faculty of *memory*, and at the fact that animals as lowly as insects could possess it. The idea that learning is a continuous process, to be measured by smooth curves, may derive from this impression that the act of storage itself is a demanding one. Now that electronic gadgets such as tape recorders and computers are commonplace, it is obvious that the processes involved in reversible storage of information need not be complex and are probably ubiquitous. The task for behavioral analysis is not to explain memory as such, but to understand the organizational and computational properties of organisms. It is not learning (in the sense of a mere effect of past experience) that is the problem (see Revusky, 1977), but rather the elementary rules and codes by which animals master complex tasks. I return to this theme later in connection with theories of classical conditioning.

13. *Latent Learning*. The controversy about latent learning arose from the assertions of the early learning theorist Edward Tolman about maze learning. Tolman believed, quite reasonably, that rats learn about mazes as an automatic consequence of exploring them, which they will do without explicit reward. This learning was termed *latent*, but could be demonstrated by comparing groups of animals given different kinds of maze experience.

For example, in one famous study (Tolman & Honzik, 1930) three groups of rats were trained daily in a complicated maze with a single goal box. For one group, there was always

food in the goal box; these were animals rewarded for learning the maze. For a second group, there was never any food in the goal box. For the third group, there was no food for the first 10 days, but food was introduced on the 11th day and thereafter. Obviously the no-food group would show no progressive improvement: with no food in the goal-box, there was nothing to differentiate it from the terminations of other alleys, and no reason the animals should enter it sooner or with fewer “errors” (blind-alley entries) than any other place in the maze. The critical comparison was between the food group and the group fed only after day 10: Would the group fed after the 10th day learn at the same rate as the food group (measured from day 1), or more rapidly, indicating some (latent) learning on the preceding days without food?



**Figure 14.8.** An early demonstration of latent learning in a maze. The curves show the average time to reach the goal box each day (trial) for three groups of rats. Control group NR was never fed in the maze; control group R always had food in the goal box. The experimental group, NR-R, had food in the goal box only after day 10. (From Tolman & Honzik, 1930.)

The results are shown in Figure 14.8. As expected, the no-reward group (NR) shows no improvement over days in time to reach the goal, and the rewarded group (R) shows a progressive improvement. The group rewarded only after day 10 (NR-R) shows an improvement after food is introduced that is much more rapid than the original improvement of the food group, which seemed to confirm Tolman’s latent-learning hypothesis.

This result might seem to settle the matter, but it did not. As often happens when terms are poorly defined, counter-explanations are offered, numerous experiments are carried out, and finally the issue either dies of boredom, or becomes changed into a question that can be answered experimentally. Opponents of Tolman’s position pointed out that rats take time to *habituate* to a new maze. Initially they may be scared and the behavior associated with this fear may interfere with maze learning. Although (these critics continued) the NR-R group was learning nothing about the maze before food was introduced, it was getting over these initial, interfering response tendencies. Once reward was introduced the animals appeared to learn rapidly because their learning was *unimpeded*, not because they had already learned anything. These response tendencies could not be accurately measured, nor could their supposed interfering effect be objectively assessed. Nevertheless, all agreed that rats do behave differently on first exposure to a new maze, so this objection could not easily be ruled out.

Other critics, whose errors were of a more subtle sort, conceded the learning of the NR-R group, but pointed out that food could not be considered the only reward in the maze. Perhaps exploration has its own rewards (“seeing a new alley” may be intrinsically rewarding), so the apparently unrewarded latent learning was not unrewarded at all. Thus free to invoke new rewards at will, Tolman’s opponents rested their case, secure in a theoretical position incapable of disproof.

Later experiments, with much simpler mazes, showed that indeed rats do learn something about the configuration of a maze in advance of food reward. For example, in an experiment very similar in design to the bee-learning study, Seward (1949) allowed a hungry rat to explore a T-maze with two distinctive empty goal boxes. Then he placed the rat in one goal box with some food and allowed it to eat. When returned to the beginning of the T, rats so trained almost invariably selected the arm leading to the rewarded goal box.

For reviews of the literature on latent learning see Osgood (1953) and Mackintosh (1974).

14. This problem has been tackled in connection with the so-called “two-arm bandit” problem by Krebs, Kacelnik and Taylor (1978). In their experiment great tits (*Parus major*) chose between two variable-ratio schedules with different mean values. The problem for the birds was to decide which of the two schedules paid off with the higher probability. This setup is formally similar to two Las Vegas-style one-arm bandit machines, hence the label. The birds sampled both choices for a while and then settled almost exclusively for one choice. Moreover, the longer the experimental session, the longer the birds sampled before settling down, just as theory suggests they should: the longer the session, the more time available to exploit the best choice, hence the greater the return on time invested in sampling. Even in this simple situation, theoretical analysis is quite tricky.

It is not yet clear exactly *how* the birds were able to devise their optimal strategy, but it seems likely that interactions between experimental sessions, according to the memory principles discussed in Chapter 12, may play a large role (see also Staddon, 1988 and Staddon & Horner, 1989, more mechanistic analyses of this problem).

The optimal-sampling problem also turns up in connection with the problem of resistance to *extinction*: How long should an animal persist in a learned response when reward is no longer forthcoming? The so-called *partial reinforcement effect* — greater persistence following intermittent reward — obviously reflects the optimal policy of persisting longer when reward is uncertain (cf. Chapter 14). McNamara and Houston (1980) have tackled this difficult theoretical problem.

15. *Bayes’ rule*. The term *Bayesian* refers to the Reverend Bayes, an 18th century English clergyman and inventor of a rule which allows the estimation of a *posterior probability* for any hypothesis, given its *prior probability* and some new data. The rule can be derived simply by rewriting basic probability identities, as follows: Let  $H$  be the hypothesis that some event is the cause of an observed outcome  $E$ . Let the probability that  $H$  is true be  $p(H)$ , and the probability that event  $E$  occur under some defined set of circumstances be  $P(E)$ . Then if  $H$  and  $E$  are independent the probability of the joint event  $EH$  (i.e., both  $H$  is true and  $E$  occurs) can be written in two ways, in terms either of  $P(E)$  and the conditional probability of  $E$  given  $H$ ,  $p(E|H)$ , or in terms of  $P(H)$  and the conditional probability of  $H$  given  $E$ ,  $p(H|E)$ :

$$p(EH) = p(E)p(H|E)$$

or

$$P(EH) = p(H)P(E|H).$$

Eliminating  $p(EH)$  and rearranging yields

$$p(H|E) = p(H)p(E|H)/p(E),$$

which is Bayes’ rule. So much is tautology; things get tricky when we try to use Bayes’ rule to infer the true value of  $p(H|E)$ , the probability  $H$  is true given  $E$ , from some estimate of  $p(H)$  in the absence of  $E$ , an estimate of  $p(E)$ , and knowledge of the probability of  $E$  given that  $H$  is true,  $p(E|H)$ . The problem is that we cannot really estimate  $p(H)$ , the prior probability of  $H$ ; nor do we really know how to compute  $p(E)$  in the abstract. The rule works perfectly when we are dealing with a well-defined procedure (see, for example, Staddon, Hinson & Kram, 1981, for a Bayesian analysis of some choice procedures). It is essentially unusable as a way of checking out real hypotheses in open-ended situations. I believe that the rule may be more valid as applied to animal’s judgments, however, because the experience of an animal’s ancestors potentially provides it with estimates for  $p(H)$ ,  $p(E|H)$  and  $p(E)$ .

Much of probability theory is mathematically trivial, but conceptually very difficult,

however, so it is hard to say anything very conclusive on these points as yet.

The term *Bayesian inference* is broadly applied to any method of inference that takes account of prior information in its evaluation of new information. It is possible to show that Bayesian inference will usually lead to better decisions than any method which completely ignores prior information. Animals certainly seem to use prior information, their own, and what can infer of the genetically coded experience of their ancestors, in this way.

16. There will always be a tradeoff between misses and false alarms for any given level of accuracy (see the discussion of ROC analysis in Chapter 9). For any given level of accuracy ( $d'$ ), the animal can pick an optimal criterion, which will yield a certain net benefit. However, the terms of the tradeoff can obviously be improved by additional sampling: the larger the number of trials, the more accurately the probabilities in Table 14.1 can be estimated. Natural selection will tend to produce an equality between the marginal costs of additional sampling and the marginal benefit of the increase in accuracy so obtained.

17. I make these assertions with a wary confidence based on general knowledge of insect behavior, but in fact the necessary experiments seem not to have been done.

18. The original experiments on long-delay taste aversion were done by John Garcia (e.g., Garcia, Clarke & Hankins, 1973; Garcia, Kimeldorf & Koelling, 1955). As with many novel observations, the phenomenon was discovered by accident, in this case as a by-product of applied research on the effects of gamma-radiation (which causes nausea, among other disagreeable effects) on rats. The phenomenon had been previously known to rat catchers in the form of *bait-shyness* — the avoidance of novel baits, and the consequent ineffectiveness of poisons associated with them. For a review of taste-aversion learning see Domjan (1980).

19. *Delay-of reward gradient*. The causal status of events is reflected in the *delay of reward gradient*, that is, the function relating probability of conditioning to the delay between CS and US (see Staddon, 1973). For the standard classical conditioning procedures this function peaks at about 0.5 sec, and declines thereafter (see Landauer, 1969), depending on the CR and other procedural features. Presumably the peak reflects the fact that causes take some time to produce their effects, so that the optimal delay is greater than zero. This need for delay is particularly clear in the case of poison-avoidance, since there will always be a substantial delay between ingestion of poison and the onset of its effects. Correspondingly the allowable CS-US interval is longer than usual. One would also expect that there should be a peak in the delay-of-reward function at a time considerably longer than half a second, but as far as I am aware this has not been firmly established.

20. Wagner, Rudy & Whitlow (1973); see also Kamin (1969). Rabbits, cats and many other mammals have a nictitating membrane behind their eyelids. The US for movement of the membrane is a localized electric shock applied at the periphery of the eye. The response of this membrane is readily conditioned to auditory, visual and vibrotactile stimuli. The preparation is widely used in the study of classical conditioning.

21. Note that *some* learning will certainly place on the first blocking trial, but it need not take the form of an association between the US and the novel stimulus. For example, it may be possible to show experimentally that after the first such trial, the animal is *less* likely to associate the new stimulus with either shock or its absence than if the stimulus had not been presented. The effect of the new stimulus will probably be different when presented as part of a compound than when presented alone. In fact Mackintosh, Bygrave and Picton (1977), have shown that the novel ele-

ment does acquire some suppressive power on the first blocking trial, but no further increments are added on subsequent trials.

In this section I am treating very briefly a large and highly technical literature on classical conditioning, in which the phenomena of blocking and overshadowing play a large part. The concepts of *surprise* (violation of expectation) and *expectancy* are also central to theoretical accounts of this work. For good reviews see Dickinson (1980), Dickinson & Mackintosh (1978), Rescorla (1975, 1988) and Mackintosh (1974). I return to these topics in the section on *Mathematical models for conditioning* in Note 22.

**22. Mathematical models for conditioning.** A number of mathematical models attempt to capture in some formalism the joint effects of *surprise* and *Bayesian inference* on conditioning. The simplest way to represent the joint dependency is to assume that a *change* in the power of a stimulus to affect behavior — usually termed a change in *associative value* and equated to learning — is proportional to the *product* of two functions, one representing the surprise aspect, the other the inference aspect:

$$\Delta V = f(\text{surprise}).g(\text{inference}),$$

where  $\Delta V$  is the change in associative value ( $V$  is almost invariably measured in terms of the *suppression ratio* in the CER experiment; see Chapter 5). If  $f$  is a monotonic function passing through the origin, then when surprise = 0,  $\Delta V$  must be zero, which conforms to the axiom that surprise is necessary for learning.

Perhaps the simplest assumption we can make about surprise is that it corresponds to a discrepancy between the asymptotic associative value,  $\lambda$ , a CS would have after consistent pairing with a US and its actual value,  $V$ :

$$f(\text{surprise}) = \lambda - V. \tag{N14.1}$$

If the CS is a compound one, parsimony suggests that  $V$  is just the sum of the associative values of the elements:  $V = \sum V_i$ . For two elements, this yields

$$f(\text{surprise}) = \lambda - V_1 - V_2. \tag{N14.2}$$

The simplest inference assumption is that all stimuli are equally likely candidates for conditioning, except that some are more *salient* (more intense, more noticeable) than others, and so condition faster. This implies that the inference function is a constant which is different for different stimuli; whence

$$g(\text{inference}) = a_i, \tag{N14.3}$$

for the  $i$ th CS.

Combining equations N14.2 and N14.3 yields

$$\Delta V_i = a_i(\lambda - V_1 - V_2), \tag{N14.4}$$

where  $\Delta V_i$  is now the increment in associative strength of *one* of the elements of the compound stimulus. Equation N14.4 is the well-known classical-conditioning model proposed by Rescorla and Wagner (1972). The effects of extinction (omission of the US) and US intensity are handled by setting  $\lambda$  equal to zero or different positive values; variations in stimulus properties are handled by the subscripted parameter  $a$ ; and the different rates of learning associated with different

levels of reinforcement are handled by means of a multiplicative parameter  $b$ , so that the final equation is

$$\Delta V_i = a_i b_j (\lambda_j - \sum V_i), \quad (\text{N14.5})$$

where  $\Delta V_i$  is the sum of the associative strengths of all stimulus elements.

Equation N14.5 is a direct descendant of the linear-operator learning model of Estes, discussed briefly in Chapter 9. The essential difference is the assumption that learning ( $\Delta V$ ) about each stimulus element is proportional to the difference between the asymptotic level ( $\lambda$ ) and the *sum* of associative values of the stimulus elements. A less essential difference is the change from probability to associative value as the dependent measure; this change allows  $V$  to take on negative values (corresponding to inhibitory strength) and values greater than 1.

This clever model accommodates both blocking and overshadowing very well. In a compound made up of elements with high and low salience, when learning is complete ( $\Delta V = 0$ )  $V_1 + V_2$  will equal  $\lambda$  but most of the associative strength will have gone to the more salient stimulus; in fact, it is easy to show that at equilibrium,  $V_1/(V_1 + V_2) = a_1/(a_1 + a_2)$ .

Blocking is accommodated by the fact that after the first phase, in which  $S_1$  alone is paired with the US,  $V_1$  must approximately equal  $\lambda$ . Since  $V_2$  is initially zero,  $\lambda - V_1 - V_2$  is close to zero when  $S_1$  and  $S_2$  are presented together at the beginning of the second phase. Consequently, little or no further change can take place in either  $V_1$  or  $V_2$  in the second phase. Thus, the conditioning of  $S_2$  is blocked by the prior conditioning of  $S_1$ .

The model also accounts for the results of Rescorla's "truly random control" procedure, discussed in Chapter 5. Single-stimulus conditioning can be considered as compound conditioning in which *contextual cues* (the physical environment of the test apparatus) constitute the other elements. In "true discrimination" where the US is paired only with the CS (stimulus 1), the CS and the contextual cues (stimulus 2) receive the same number of increments in associative value during conditioning, but the context also receives a substantial number of decrements (cf. Figure 5.8). For example, if the CS is present half the time, and a brief shock occurs on the average once a minute, then in an hour there will be 30 increments in the strength of both CS and background stimuli. However, there will be 30 missed opportunities for shock in the absence of the CS but in the presence of the context, which will cause reductions in the strength of the context. It is easy to show that this process will result in eventual exclusive control by the CS.

In the truly random control condition (sometimes called *pseudodiscrimination*), however, where the US is equally likely in the presence or absence of the CS, the CS continues to receive 30 increments in every hour, but the context now receives twice as many. Moreover, pairings of the US and context in the absence of the CS lead to larger increments in context strength than pairings of the US with the CS-context compound (because the surprise term is  $\lambda - V_1 - V_2$  in the compound case, but  $\lambda - V_2$  in the context-only case). This leads eventually to predominant control by the context and no conditioning (zero associative value) to the CS.

The Rescorla-Wagner equations therefore constitute a mechanism by which animals might detect correlations between conditioned and unconditioned stimuli. The process is computationally simple because the animal need keep track only of the values of  $V_i$ ; it needs no memory for the specific conditioning history associated with each stimulus element.

Some otherwise-surprising experimental results are predicted by this model. For example, if  $S_2$  has been established as a safety signal, its value must be negative so as to counteract the positive value of contextual cues. At the beginning of compound conditioning, therefore, the net value of the term  $\lambda - V_1 - V_2$  is positive, rather than close to zero, as in the usual blocking arrangement. Consequently, positive increments will at first be added to both  $V_1$  (which is already positive) and  $V_2$ . Since  $V_1$  is already close to  $\lambda$  in value, this procedure leads at first to greater-than-

maximal conditioning of  $S_1$  — the *superconditioning* discussed in the text. This outcome depends on the assumption that “surprise” is given by the discrepancy between asymptotic associative value,  $\lambda$ , and the sum of associative values of all stimulus elements — the last term in equation N14.5.

The assumption that surprise depends on the difference between  $\lambda$  and the sum of associative values leads to other counter-intuitive predictions. For example, consider an experiment with two compound CSs sharing a common element:  $S_{13}$  is a compound of elements 1 and 3;  $S_{23}$  is a compound of elements 2 and 3. Suppose that  $S_{13}$  and  $S_{23}$  are now presented equally often and  $S_{13}$  is consistently reinforced but  $S_{23}$  is not. This situation can be compared with a pseudodiscrimination in which  $S_{13}$  and  $S_{23}$  are both reinforced half the time. The reinforcement schedule for 3, the common element, is identical in both cases, yet the model correctly predicts that it will acquire much less strength in the true discrimination than the pseudodiscrimination.

In a related experiment, Rescorla (1976) used training with two overlapping compounds as a way to enhance the strength of one compound above its asymptotic level. In phase 1, compound  $S_{13}$  was conditioned in the normal way. The model predicts that at asymptote, the two elements, 1 and 3, will have strengths related to their saliences. In the second phase, element 3 is combined with a new element 2 and  $S_{23}$  is conditioned in the usual way. Since the associative value of 2 is initially zero,  $\sum V (=V_2 + V_3)$  will be less than  $\lambda$  at the beginning of the second phase, allowing element 3 to gain strength. The model correctly predicts that in a subsequent test of the 1-3 compound,  $S_{13}$ , the added strength of element 3 should cause the total strength of  $S_{13}$  to rise above its normal asymptotic level.

The surprising feature of this experiment is that Rescorla did not use *explicit* compounds, like a tone and a light, but an *implicit* compound: he assumed that a high- and a low-frequency tone can be considered as compounds differing in frequency, but sharing other properties. Hence he was able to explain why asymptotic suppression to a high tone could be enhanced by subsequent training with a low tone. (A similar prediction follows from the discussion of integral stimuli at the end of the chapter.)

The Rescorla-Wagner model does a great deal with a very simple formalism. Nevertheless, it fails to account for some basic facts about conditioning. For example, blocking can be eliminated by a *decrease*, as well as an increase in the expected level of shock; and the new CS element gains positive associative value in both cases (Dickinson, Hall & Mackintosh, 1976). The model predicts that the new CS element should gain positive value only if the change is an increase. The model predicts that a safety signal will have negative associative strength. Consequently, continued presentation of a safety signal in the absence of any US (a form of extinction) should lead eventually to a loss in its potency as its associative strength increases to zero, the asymptote for extinction. Yet safety signals retain their potency when presented without the US.

Most damaging is the finding that simple pre-exposure (presenting a stimulus without any consequence) to a CS can retard conditioning (*latent inhibition*). Since pre-exposure cannot change the associative value from zero (according to the model) there is no way to accommodate this result within the model.

The experiment by Hall and Pearce (1979), discussed in the text, in which pretraining with weak shock *retarded* conditioning to a strong shock, is also impossible to explain by the Rescorla-Wagner model.

These effects are comprehensible in Bayesian terms. For example, prior experience with a stimulus in the absence of shock should make it less likely that the stimulus has anything to do with shock. Hence subsequent conditioning should be retarded (latent inhibition). A safety signal presented without shock violates no expectation; hence no learning (no change in its properties) should occur.

These limitations have prompted a number of suggestions for improving the original model to allow it to deal with these exceptions without losing its ability to handle blocking and

related effects. Since the “surprise” term in equation N14.5 is essential for the prediction of blocking, most modifications concern the “inference” term.

An obvious possibility is to allow the stimulus saliences,  $a_i$ , to vary as a function of experience; this can accommodate the unique lability of novel stimuli (e.g., the special ease with which they become associated with changed conditions of reinforcement, and the retarded conditioning caused by simple pre-exposure, i.e., latent inhibition). For example, Mackintosh (1975) and Wagner (1978) have elaborated on a suggestion made by Rescorla and Wagner (1972) that stimulus salience may decrease with continued exposure to a stimulus. Wagner proposes that continued presentation of a CS in given context leads to associations between CS and context, and that the associability of the CS with a US is inversely related to the strength of these contextual associations: the more predictable the CS, the less the value of  $a$ .

Mackintosh proposed that the value of  $a_i$  decreases as stimulus  $i$  becomes a better predictor of the US. He also includes an explicit competition assumption (which is implicit in the surprise term of the Rescorla-Wagner model), proposing that increments in  $a$  occur only if stimulus  $i$  is the least predictive stimulus in the situation. This explicit competition assumption permits Mackintosh to give up the Rescorla-Wagner surprise term in favor of the original Estes version:  $\Delta V_i = a_i(\lambda - V_i)$ .

Most recently, Pearce and Hall (1980) have proposed that the predictive value of each stimulus is assessed independently (rather than all together, as in the Rescorla-Wagner model). Nothing further is learned about stimuli that predict their consequences accurately. Their model retains the surprise x inference form of equation N14.1 yielding

$$\Delta V_i(n) = S_i \lambda(n) |\lambda(n-1) - V_i(n-1)| \quad (\text{N14.6})$$

where  $V_i(n)$  is the associative value of stimulus  $i$  on trial  $n$ ,  $\lambda(n)$  is the intensity of reinforcement on trial  $n$  and  $S_i$  is intensity of CS-element  $i$ . In words, this model says that learning is proportional to the absolute value (the term in  $||$ ) of the discrepancy between the conditioning asymptote on the preceding trial and the associative value of the stimulus element on that trial (the surprise term), multiplied by the product of stimulus and US intensity (the inference term). The surprise term embodies the presumption that animals encode each individual CS in terms of the consequences that it predicts. The Rescorla-Wagner model, of course, only encodes the US. The Pearce-Hall model asks: Are the consequences of this CS accurately predicted? whereas the Rescorla-Wagner model asks: Was this US accurately predicted?

These modifications of the Rescorla-Wagner model are able to handle most of the discrepant facts I have described. The most effective so far seems to be the Pearce-Hall approach. For example, latent inhibition is explained as follows: When a novel stimulus is presented without a reinforcer,  $\lambda$  will be 0; since the associative value of the stimulus is also 0, the term in  $||$  is zero and the future associability of the stimulus will soon be very low. When the stimulus is first paired with a US, no learning will occur on the first conditioning trial. Hence learning will be retarded relative to conditioning where the novel stimulus appears for the first time on the first conditioning trial. The same argument accounts for the retardation of conditioning to a strong shock by pretraining with a weak shock: on the first strong-shock trial, the term in  $||$  will be zero, so that no learning will occur on that trial. In addition, the model makes the counter-intuitive, but correct, prediction that some interpolated *extinction* after weak-shock training will facilitate subsequent training with the strong shock — because the extinction causes a discrepancy between  $\lambda(n-1)$  and  $V(n-1)$  in equation N14.6, which allows learning to occur on the first strong-shock trial. (In colloquial terms, occasional extinctions teach the animal that it cannot trust to its earlier assessment of the significance of the CS.)

None of these modifications of the Rescorla-Wagner equations has the mathematical

simplicity of the original — both inequalities (Mackintosh) and absolute-value expressions (Pearce-Hall) prevent simple solution of the finite-difference equations. And none deals with special relations between CS and US, as in taste-aversion learning, nor do they deal adequately with trial-by-trial changes or the different effects to be expected when stimuli are integral or separable. All assume a smooth learning process for which the evidence (in individual animals) almost entirely negative. They account well for the effects of transitions between conditions, and for asymptotic levels of conditioning. They do not provide an accurate description of the real dynamics of learning. Nevertheless, unlike a Bayesian analysis, which is just one kind of optimality analysis — functional not a causal explanation — these imperfect models are a stab at a causal, dynamic explanations for Pavlovian effects.

In all the experiments dealt with by these models, the temporal relation between CS and US is standardized. For example, in the CER paradigm, shocks are paired (i.e., concurrent with) CS presentations. Yet we have already seen that some “conditioning” can take place even when the CS occurs after the US (Keith-Lucas & Guttman, 1975). If the CS and US are properly chosen, moreover, conditioning can occur over long delays, and some CSs can become preferentially conditioned over others that are (in terms of relative frequency of pairing) better predictors of the US. These CER-based models deal only with *relative frequency* of pairing between CS and US and *salience* (of CS and US) as the procedural features that determine conditioning, even though other experimental arrangements show that USs and CSs have other attributes that make them more or less associable with different USs. It seems clear that animals weigh several factors in deciding which of a number of candidate stimuli is the likely cause of a hedonic event such as food or electric shock: the type of stimulus in relation to the type of US, other properties of the CS such as its intensity, the time relation between CS and US, and the animal’s prior history with both CS and US. Models of the Rescorla-Wagner type provide elegant summaries of the effects of frequency of pairing. But we need to know much more about how the physical events in conditioning experiments are represented by animals before all these other factors can be accurately incorporated.

**23.** It is generally assumed that unconditioned stimuli such as food, water and electric shock have innately determined values, but this is not always true. For example, chicks need experience to learn that seeds are food (have positive value). Hogan (1973) exposed young chicks to piles of sand and seeds. They pecked equally often at both. Then one group was offered just sand. Given a choice an hour later, they still showed no preference for seeds. However, another group, offered just seeds after the initial experience with both, when given the choice after an hour strongly preferred the seeds. The birds evidently could not learn that sand is *not* food, but given appropriate experience could learn that seeds *are* food. A further experiment in which the chicks were intubated with high-calorie diet after eating sand showed that the preference for seeds in the first experiment was due to the nourishment derived from them, not to any special stimulus properties.

Since the development of this preference depends on the subsequent beneficial effect, a delay between the seeds-only exposure and the subsequent test is essential — presumably to allow time for some digestion and absorption to occur.

It makes great adaptive sense that chicks must learn the value of seeds, a motionless, inanimate stimulus similar to many nonnutritive objects in the environment. As one might expect, they require no learning to take live mealworms (although at first they show some conflicting fear responses to large and active worms). They almost immediately show a preference for mealworms over either seeds or sand.