10

STIMULUS CONTROL AND COGNITION

To behave adaptively is to behave differently in one situation than in another. As we move up from paramecia to pigeons and people, the number of different modes of possible behavior increases enormously, and with it the number of different situations for which a unique behavior is appropriate. How do animals organize this knowledge? And how are situations recognized? The two questions are not really separate, since some kinds of organization make recognition quick and accurate, while others make it slow and unreliable. The ability to recognize when particular adaptive behaviors are appropriate has been taken for granted in previous chapters. In this chapter I look at one aspect of recognition and the organization of individual knowledge.

What does it mean to *recognize* something? In a formal sense the answer is simple: it means to be in a unique state (as defined in Chapter 4), so that in the presence of object *w*, the animal is always in state *W*, and state *W* never occurs at any other time. This is also a necessary condition for the animal to *discriminate w* from things that are not *w*. But the formal answer conceals a great deal. For example, it isn't much help in constructing machines that can recognize, for which much more specific information is required: We need to know how to process particular visual (or auditory, touch, or whatever) inputs; how to direct the visual apparatus on the basis of past information — where should the machine look next? How to distinguish objects from their backgrounds; how to identify the same object from different points of view; and how to encode all this information so that it provides a useful basis for action.

Most of these questions are about *perception*, and I can't do justice to them here. In order to get on with the study of learning and motivation we must take for granted the processes that translate a particular physical environment into some internal representation that allows the animal to recognize the environment on subsequent occasions (or, more cautiously phrased, the processes that allow the animal to behave in the same way — or ways that are the same in essential aspects — every time he is in the same environment). Perceptual processes are not trivial; in many respects they are much more complicated than the things I do deal with. I pass them by not because they are negligible, but because they are little understood — and because my main interests are elsewhere. This chapter is about the last step in the process: the encoding of information in ways useful for action. I am concerned not with how the animal "sees" a Skinner box or a colored light, but with how these things resemble or differ from other things in his world.

It is not at all clear that the best way to answer this question is to begin with the concept of *stimulus*, where a stimulus is some very specific, physically defined event. Nevertheless, because the study of learning in animals grew up under the influence of reflex-oriented behaviorists, the relevant terms, experimental methods, and concepts have all evolved from "stimulus-response" psychology. Hence, "stimulus" is the natural place to start. The term's behaviorist ancestry at least means that we will know what we are talking about, even if we are uncertain about how useful it will prove to be.

I first discuss the concept of stimulus and stimulus element, then describe how control by simple stimuli is measured in transfer tests. The results of transfer tests can often be summarized by rules describing the effects of stimulus compounds in terms of the effects of their constituent elements. Tests are also useful for assessing the effects of reinforcement on stimulus control. The chapter ends with a discussion of similarity and the invariant relations among simple and complex stimuli that may underlie performance in stimulus-control experiments, and in animals' natural environments.

THE DEFINITION OF STIMULUS

Discriminative and Eliciting Stimuli

The etymology of "stimulus" implies the existence of a response: a stimulus is a stimulus *for* something. As I noted in Chapter 4, this definition is too narrow: a stimulus may change the organism's internal state without immediately eliciting a response. But some stimuli are clearly more important for the actions they produce, others for the change of state. Some stimuli are goads to action; others function more as signals. The stimulus for an "ideal" reflex is a goad: the response follows the stimulus invariably and at once. A "pure" eliciting stimulus is one that produces a reaction but has no other effect on the animal. Obviously, few, perhaps no, stimuli fit this description exactly, but stimuli for simple protective reflexes of decerebrate organisms (light for pupillary contraction, air puff for blinking, touch for the scratch reflex) come close. In reality, of course, the phenomena of temporal and spatial summation, refractory period and so on (see Chapter 2) show that even eliciting stimuli have effects on the animal that go beyond the response they produce.[1] And normal (not decerebrate) animals can remember past stimuli, which can therefore affect future behavior after long delays.

The distinctive property of *discriminative* or *controlling* stimuli (I use these terms synonymously) is that they define a certain *state* of the organism, defined as a set of stimulus-response and stimulus-state-change relations different from those associated with other states. Discriminative stimuli serve as signals defining situations in which a particular course of action is appropriate. For example, young bluejays will attack and eat most butterflies. Monarch butterflies are mildly poisonous and make birds sick. After some unfortunate experiences with monarchs bluejays learn to avoid them. The distinctive red and black pattern of the monarch signals (*controls* in the conventional terminology) a pattern of behavior different from that normally elicited by butterflies. Thus, in the presence of a discriminative stimulus, the animal behaves according to a set of rules different from those applying in the presence of other discriminative stimuli.

The signal defining the situation need not be the thing attacked or avoided; the signal stimulus may be different from the stimulus responded to. For example, on rainy days birds may look for worms brought to the surface by water-logging of their burrows, whereas on dry days they may look for other prey: the weather is the signal, but the worm is attacked. A hungry dog may rush to its food bowl at the sight of its master if this usually means a feeding. The number of caterpillars a female digger wasp (*Ammophila campestris*) brings back to feed each of her separate larvae is determined during a daily checkout visit: the food store for each larva is the signal controlling subsequent foraging. Even primitive animals can show contextual reactions of this sort; we saw in Chapter 3 that the protozoan *Paramecium* under suitable conditions is geopositive in the light and geonegative in darkness: I termed this kind of flexibility systematic variation. For mammals, birds and a few other animals, systematic variation — sensitivity to context, control by discriminative stimuli — is often acquired during the life of the individual: This is *learned* systematic variation.

This view of stimulus effects is not too different from the commonsense idea that animals perceive the world as a set of situations, each with its own regularities.

This is an approach, not a finished theory. I present it at the outset because the experimental methods I discuss in a moment, and the history of this field, foster the deceptively simple alternative that operant behavior can be understood solely in terms of stimulus-response relations, where both stimuli and responses are defined as physical events. Skinner[2] enlarged the definitions of stimulus and response to embrace classes of physical events linked by a common consequence: a stimulus was the class of all physical stimuli that signaled an operant contingency; a response was the class of all acts (physical topographies) that satisfied the contingency.

But he was silent on the relations among these classes and their structural properties. The stimulus-response view, in either its simple or enlarged form, is experimentally convenient, but it provides few clues to *process* — to understanding how animals work. Because it never goes beyond single stimuli, or stimulus classes, to the relations among them, it is little help in understanding how animals get about in natural environments, why some complex stimuli appear similar while others look different, how pigeons differ from people, or why stimuli that are psychologically simple to identify (faces and other natural objects, for example) are often physically complex.

There are two main approaches to these questions. One is perceptual and physiological: to identify the transformations imposed on the physical stimulus by the animal's nervous system that allow it to detect complex invariances, such as the fixed shape of a three-dimensional object perceived from different angles or the constant size of the same thing seen at different distances. Perception is hard to study with animals, and as I suggested earlier, our knowledge even of human perceptual processes falls far short of answering these questions. The second approach is functional: The evolutionary function of knowledge must be as a guide to action; hence, the animal's task is always to organize the welter of physical stimulation in ways that enable it to deploy its behavioral resources with maximum efficiency. Physical stimuli that signal the same set of regularities (in the life of the individual, or in the lives of his ancestors) should be treated as equivalent, and come to control the same pattern of adaptive behavior: the class of such physical stimuli then constitutes a stimulus in the functional sense. The relations of similarity and difference between stimuli so defined constitute the animal's knowledge about his world. The functional approach to stimulus control (which turns out to be also a cognitive approach) therefore begins with the study of similarities and differences, and hopes to end with some representation of knowledge.

Stimulus equivalence and data structure

The experimental study of stimulus effects therefore boils down to two questions — about stimulus equivalence, and about what may be termed (borrowing from computer science) data structure: (a) *Stimulus equivalence* — what stimuli are equivalent in their effects to a given stimulus? (b) *Data structure* — How many different equivalence *classes* are there, and how are they related to one another? Stimulus equivalence is usually studied under the rubric of *stimulus generalization*. The question of data structure has rarely been studied directly, although it is related to the traditional problem of *discrimination*.[3] I return to data structure in the section on similarity.

Species differ greatly in complexity of stimulus classes and data structures. Simple invertebrates are often guided by signals that can be identified with relatively simple physical properties. They classify the world in terms of relatively crude categories. For example, ticks locate their prey by seeking a certain height above the ground. There the tick waits, until an animal passes close enough to provide the necessary chemical stimulus — whereupon the pest releases its hold and drops on its unwitting host. For a tick, evidently, the world is divided into different heights (how measured?) — most bad, a narrow range good — and the presence vs. absence of butyric acid. No doubt the animal can also identify acceptable ranges of temperature, illumination and humidity. But other features, obvious to us, are ignored. The music of Mozart, the beauty of a summer evening, even the difference between one grass and another, all pass it by. The tick asks rather few questions of the world and is content with simple answers. The more intelligent the animal, the less will this be true. Mammals and birds can identify dozens or hundreds of different situations and obviously react to complex properties of their environments. Occasional demonstrations that birds and mammals can be tricked by simple *sign stimuli* (such as the male robin's foolish attacking of a red piece of fluff, or the stickleback's attack on primitive models) attract experimental attention in part because they are easily studied exceptions to the prevailing complexity.

Although everyone knows that mammals and birds must have rich world-models, in practice we are usually reduced to studying one or two simple stimuli and a comparable number of responses. Nevertheless, it is worth remembering that from the animal's point of view each stimulus class has meaning only in relation to others — just as each state of the finite-state systems described in Chapter 4 makes sense only in relation to the whole set of states. A state is a *relation*, defined by difference from other states, not a thing. Consequently, to really understand the effects of any particular stimulus, we need to compare its effects with the effects of a wide range of others.

The needs of a tick are simple, its responses few, and its information-processing capacity limited. It functions rather like a guided missile, using an easily measured cue sufficient to identify its intended host. Butyric acid is evidently an adequate signal for a meal for a tick, as heat suffices to guide the sidewinder missile to a plane or a tank. More complex animals can do and recognize more things, and their niches demand more complex relations between what they can recognize and what they can do. But simple or complex, animals need differentiate among states of nature only to the degree that the states make a difference — both in terms of consequences of value, and in terms of the animal's ability to respond appropriately.

"Consequences of value" are of course just the events termed "reinforcers." States of nature that make no difference, now or in the lives of ancestral animals, will not usually be differentiated — will not produce different states of the animal (cf. Chapter 4). I describe how stimulus control depends on the function of stimuli as reward signals in a moment.

"Ability to respond appropriately" is harder to define. Animals are limited in three ways: by the physical properties to which they are sensitive; by the responses they can make; and by the their ability to integrate stimulus information and use it as a guide to action.

For example, bees are sensitive to near-ultra-violet light, but humans are not; hence bees can respond to an aspect of the world that we cannot. Sensitivity to the appropriate physical property is a necessary but not a sufficient condition for adaptive response. Similarly, birds can fly, pigs cannot; some responses are physically possible for an animal, others are not.

But the most interesting constraints are information-processing ones: Octopi are reported to have difficulty in discriminating between a figure and its vertical reflection (between \ and /, for instance). The necessary information is evidently available at the retinal level (the octopus can discriminate between \ and |, for example) but processing limitations prevent the animal from making appropriate use of it. The digger wasp just mentioned is subject to a particularly interesting limitation. A female usually maintains several burrows, each containing a larva. Hence she must keep track of the provisioning requirements at several locations — a considerable feat of memory. Yet she updates her information on each burrow only on the first visit each morning. Classic experiments by the Dutch ethologist G. P. Baerends showed that the information is not updated on subsequent provisioning visits, even if circumstances have changed. If, in between visits, some food is removed, the animal does not make up the loss that day.[4] Again, the information is available, but the animal cannot process it appropriately. The functional explanation for this limitation, of course, is the absence of such interventions in the life of wasp ancestors: a predator that removes the larva's food is likely also to remove the larva, so there can have been no direct selection for rechecking after the first daily visit. (We can confidently predict that the wasp will notice the absence of her larva on later visits, however.)

The most widespread information-processing limitations derive from the imperfections of memory. Most animals are very poor at learning sequences. For example, rats cannot learn to make a sequence of choices such as ABAABBAAABBB to get food. Instead they will make A and B choices in an irregular way so that occasionally by accident the correct sequence occurs and food is delivered. For the rat, the key property — the property it can detect as a predictor of food — may be not the sequential arrangement of As and Bs, but their relative frequency. Perfect responding here requires either learning a *rule* — or exact determination of the nth choice by up to eleven preceding choices, that is, accurate recall of these events in order. People are partly immune from this limitation because symbolic representation — language, numbers, and so on — provides a *digital* method of representing past events. In contrast, "lower" animals appear to encode past events in *analog* form, so that accuracy decreases as sequence length increases. Thus, a rat finds it very difficult to learn sequences of the form NAB (AB, AAB, AAAB, etc.) where *N* is greater than two or three; a person can count to *N* and has no difficulty[5]. Nevertheless, in situations where symbolic coding is difficult or impossible, the temporal resolution even of human memory is far from impressive. For example, when similar events occur in temporal alternation, as in changes of "serve" during a tennis match, after several alternations it may be hard to remember which event occurred most recently.

The "situations" into which animals organize the world are therefore limited by their ability to detect certain kinds of relations and not others. Animals are usually good at detecting the times at which food will occur and at identifying situations where it occurs with highest frequency; they are not good at picking up complex stimulus sequences.

Measurement of stimulus control

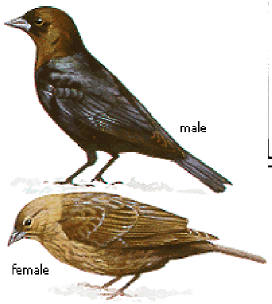
Stimulus equivalence can be studied with simple stimuli by training an animal with one stimulus, and seeing to what extent the trained response will occur to others (cf. Chapter 4). This approach raises two general questions: (a) How is control established? (b) How do the physical properties of stimuli affect stimulus equivalence? The answer to the first question was foreshadowed in Chapter 5: a stimulus comes to control behavior when it predicts something about positive or negative reinforcement. For example, suppose we take a hungry pigeon and place it in a Skinner box in which every minute or so a red key light comes on for 8 sec. Now imagine two possible experiments with this basic procedure. In the first, the animal is given brief access to food 60 times per hour, with the occurrence of food and onset of the red light determined by independent, random processes. In the second experiment, food again occurs 60 times per hour, but now always at the end of the 8-sec light. In the first experiment, the pigeon will probably look at the light the first few times it comes on, but will soon ignore it. The pigeon will spend no more time near the light than near any other feature of the apparatus an equal distance from the feeder (no doubt he will spend more than a chance amount of time near the feeder).

The result of the second experiment is quite different. The bird will attend more and more to the light instead of less and less, and within 30 or 40 food deliveries is likely to peck at the red key. Once it has developed, this *autoshaped* pecking will be maintained indefinitely. Autoshaping is quite a general result.[6] The particular response to be directed at the stimulus will depend on details of the apparatus and the species of animal. Rats, for example, are notoriously unwilling to peck a key, but they may press or chew an illuminated lever. Most species will learn to approach the signal stimulus when it comes on, and many will also learn something about its fixed duration, approaching the food hopper as the end of the 8 sec approaches. Thus, the rule for the development of stimulus control is that there be a stimulus-reinforcer *contingency*, in the sense described in Chapter 5.

The only way to find out which physical properties of a stimulus are important to the behavior it controls is to vary the stimulus and look at the effect. This is termed *transfer testing*, since the idea is to see how much of the original behavior transfers to the control of the stimulus variants during the test. Transfer testing involves two steps. The first is to identify a physical stimulus that controls an identifiable aspect of behavior. Sometimes the control has been established through training, but it may also be a natural relation. The second step is to vary aspects of the stimulus under conditions where any associated change in behavior can be attributed solely to these stimulus variations.

Vocal communication of the American brown-headed cowbird (*Molothrus ater*) provides a neat example of natural stimulus control. The cowbird is a brood parasite: like the European cuckoo, it lays its eggs in other bird's nests. This unlovely habit has made it of great interest to biologists from Darwin onward, but for present purposes the important thing is the vocal repertoire of the male. The male cowbird produces a song consisting of a series of whistle-like sounds that elicits a distinctive "copulatory posture" from a receptive female. This response is rapid and easily recognizable; hence it provides an excellent way to measure the effectiveness of song variants. A picture of a typical song is shown as Figure 10.1. In one experiment, (West, King, Eastzer, & Staddon, 1979), preliminary tests with tape-recorded songs showed that the song can be divided up into three significant units: phrase 1 (P1), phrase 2 (P2) and the inter-phrase unit (IPU). In subsequent tests, songs in which one or two of the three segments had been deleted were played to females.

Table 10.1 shows the results, represented as follows: S = complete song; P1, P2 = phrases 1 and 2; IPU = inter-phrase unit; S - IPU = complete song with IPU deleted (the same as P1 + P2); P1 + IPU = phrase 1 followed by IPU; P2 + IPU = IPU followed by phrase 2. (Data are not shown on the IPU presented in isolation, since it then had no effect.) Each of five receptive females heard about 200 songs, equally divided among these variants. The table shows the percentage of positive responses over the whole group for each song variant.



Brown-headed Cowbirds by Larry McQueen

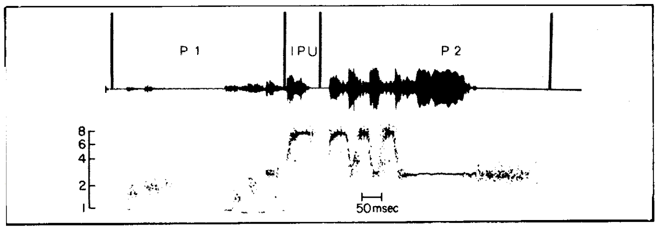


Figure 10.1. Typical song of male cowbirds reared in isolation. The upper display shows the changes in amplitude (intensity) through the song. The lower display shows the associated changes in instantaneous frequency (frequency in kHz on the vertical axis). (From West, King, Eastzer, & Staddon, 1979, Figure 1.)

The results can be summarized by two statements: (a) P1 and P2 contribute to song effectiveness in an independent additive way: if p_1 and p_2 are the probabilities of response to P1 and P2 presented alone, then the probability of response when both are presented (assuming independence) is just

$$1 - (1-p_1)(1-p_2) = p_1 + p_2 - p_1p_2,$$

that is, one minus the probability that the animals respond to neither P1 nor P2. Plugging in the values for p_1 and p_2 from Table 10.1 yields a predicted joint probability of .41, which compares well with the obtained value (for S-IPU = P1 + P2) of .38. (b) Addition of the IPU approximately doubles the effectiveness of a song variant (compare S - IPU with S, P1 + IPU with P1, and P2 + IPU with P2). Because probabilities cannot exceed unity, the proportional increase is somewhat less as the base value increases (e.g., 75/38 < 46/19).

Table 10.1.

Song variant	Percentage Response
S	75
S − IPU	38
P1 + IPU	62
P1	27
P2 + IPU	46
P2	19

This example illustrates the plan of attack when attempting to measure stimulus control. First, you need some idea of the general features of the physical stimulus that are likely to be important. In the cowbird case, for example, preliminary work showed that the amplitude of the signal was less important than changes in its frequency over time (frequency modulation). Then these critical features are varied, either by selective omission, as in this example, or by graded variation, as in *generalization testing*.

In either case, the spacing and frequency of tests must be chosen with care, so that the response does not change for reasons unrelated to the stimulus. These confounding effects are principally *habituation*, for naturally occurring stimulus-response relations (as in the cowbird example), and *reconditioning* or *extinction* for relations established through differential reinforcement. Habituation is the eventual weakening of a response following repeated elicitation (cf. Chapter 4). Female cowbirds habituate during repeated song playbacks unless the playbacks are relatively infrequent, as they were in this experiment. Reconditioning, that is the establishment of control by a test stimulus that would otherwise be ineffective, can occur if reinforcement continues to occur during the generalization test. Conversely, the response may extinguish if reinforcements are omitted during the test. I explain the experimental solution to these conflicting requirements in a moment.

If the critical stimulus features have been correctly identified, the results of the tests should lend themselves to a simple description of the sort just offered for the cowbird song. If no simple pattern can be discerned, then it may be either that no simple pattern exists — or we have failed to define the essential stimulus features.

STIMULUS GENERALIZATION

Cowbirds need no training either to make or respond to calls of the type shown in Figure 10.1. (Indeed, one result of this experiment was to show that isolated male cowbirds have more effective songs than males reared normally in the company of their fellows.) Every animal species shows examples of stereotyped, innate relations of this sort.[7] For many, these reactions, together with primitive orienting mechanisms of the type described in Chapters 2 and 3, constitute the animal's entire behavioral repertoire. In mammals and birds, and a few other species, however, much behavior consists of reactions acquired to relatively arbitrary stimuli.

A simple procedure for studying learned stimulus control of this type is as follows. A hungry pigeon is first trained to peck a key for food reinforcement, delivered on a variable-interval schedule of intermediate value (e.g., VI 60 sec). The response key is illuminated with the stimulus of interest (termed *S*+ or S^D — *D* for discriminative). In early studies *S*+ was usually chosen to be physically simple — light of a single wavelength, a single vertical line, a tone. But physical and psychological simplicity are not the same thing so, as the technology for stimulus presentation has improved, physically complex stimuli — pictures of scenes or animals, for example — have received more attention.[8]

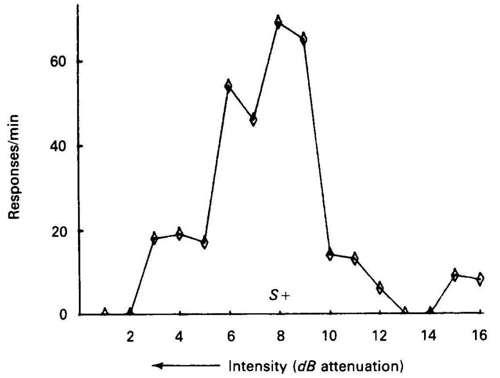


Figure 10.2. Key-peck rate of an individual pigeon as a function of the intensity of a green keylight during a single 60-min generalization test (Staddon, unpublished data). (Averaged data show a smooth peaked distribution centered on *S*+.)

After the pigeon has learned to peck *S*+ for food, variants on the test stimulus are presented for relatively brief periods — for example, if five variants are tried, each may be presented for one minute in an irregular order, for a total of perhaps 60 presentations. Thus, each variant is presented repeatedly, so that any slow changes in the tendency to respond are shared equally by all. No reinforcement occurs during a *generalization test* of this sort, so as to avoid reconditioning. Extinction is prevented, or at least mitigated, by the VI training. On the VI, long periods without food are common; hence, the animals do not soon cease to respond when food is omitted entirely. The *extinction curve* is quite gradual, and over a 60-min period, responding generally shows little decline. Thus, the average rate of response in the presence of each test stimulus is likely to be an accurate measure of its tendency to facilitate or suppress responding relative to *S*+.

Figure 10.2 shows the results of a typical generalization test. The horizontal axis shows the physical values (here intensities of a monochromatic green light) of eight test stimuli. The vertical axis shows the average number of pecks/min made by an individual pigeon to each of these values during a single 60-min test session. The resulting symmetrical curve is typical: responding is highest to *S*+ and decreases more or less smoothly as the physical value of the test stimulus departs from the *S*+ value.[9] Because it is highest at *S*+ and declines on either side, the gradient in Figure 10.2 is termed an *excitatory* or *decremental* generalization gradient.

A great many physical stimulus properties, such as wavelength of light, line tilt, roughness, spatial frequency and others, have been tested like this. The almost universal result, with pigeons, rats, monkeys, goldfish and people, is the kind of gradient shown in Figure 10.2: responding is maximal at (or near) *S*+, and falls off systematically with the physical stimulus difference between *S*+ and the test stimulus.

In some respects this result is unsurprising: why shouldn't behavior bear an orderly relation to the properties of the physical world? Often the physiological basis for the relation seems obvious. Tonal frequencies are spatially represented on the basilar membrane of the cochlea, for example; many neural transducers, such as those for pressure and light, fire at a rate directly related to physical stimulus intensity.[10] But in other cases, the result is puzzling. Color perception, for instance, depends upon central integration of information from three or four types (depending on species) of photoreceptor, each with a different wavelength-absorption peak. Wavelength, as a continuum, has no simple representation in such a system. Yet wavelength generalization gradients are among the most orderly, and show no sudden slope changes, even at color boundaries. In a moment I discuss evidence from human experiments suggesting that generalization gradients probably represent something like stimulus *similarity*, a cognitive rather than a purely sensory property.

Compounding of elements

I have now described two kinds of stimulus control: by a stimulus *element* (the cowbird example), and by a stimulus *dimension*. These two kinds of control reflect the different test operations: When a stimulus dimension is varied, but the stimulus element is always present, then we are measuring dimensional control; when the element is sometimes removed entirely, we are measuring control by the element.

Elements and dimensions can be defined by simple physical properties, such as intensity or wavelength, or in some other way. For example, rather than splitting up the cowbird song into elements, we could have considered "proportion of total song time" as a stimulus dimension: This is a perfectly objective property, but it would it not have been very useful because, as we have seen, the elements (P1, P2, etc.) vary greatly in their effectiveness as elicitors of the response, and two of them combine additively, while the other one seems to act multiplicatively. In other words, the things we choose to vary in a generalization test cannot be arbitrary. We must judge the correctness of our choice by the comprehensibility of the results. The justification for labeling P1, P2 and IPU as elements in the cowbird song is that they behave in intelligible ways when compounded.

Subjective experience suggests two ways that stimulus elements can combine, and these seem to correspond to different algebraic rules. For example, a visual stimulus element such as a triangle must have some color: neither form nor color can exist in isolation. People and animals tend to perceive colored objects as wholes; they don't normally attend to form and color separately. Dimensions treated this way are described as *integral*, and they roughly follow a multiplicative rule — a value of zero on either dimension and the stimulus has no effect: a form with no color cannot be seen. On the other hand, it is easy to imagine visual displays whose elements are not so intimately related: a pigeon might attend to the stars or to the stripes in a picture of Old Glory, for example. Stimulus elements of this sort are termed *separable* and follow an additive rule.[11] In the cowbird experiment, the two phrases P1 and P2 appear to be perceived by the birds as separable elements, since their effects were independent. But the inter-phrase unit, IPU, looks more like an integral element, since it had little effect on its own but greatly enhanced the effect of other elements with which it was compounded.

Animals must behave with respect to objects or states of the world, not stimulus dimensions or elements. One use for the elegant technique of generalization testing, therefore, is to shed some light on the way that the animals classify things as a guide to behavior. Since objects differ not in one dimension but in many, the interactions among dimensions have first claim on our attention. Unfortunately, rather little is known about multidimensional stimulus generalization. One reason is technical: it is no simple matter to create and manipulate at will multidimensional stimuli. A second reason is that once invented, techniques take on a life of their own. We know that rewarding an animal for pecking a key illuminated with monochromatic light will cause him to attend to wavelength. Why not look at the effects of reinforcing two or more wavelengths, of alternating reinforced and unreinforced stimuli of different wavelengths, of successive vs. simultaneous comparison, and so on? All these things have been done, usually with orderly results not devoid of interest — the next chapter is largely concerned with them. But the relation between these neat manipulations and the animal's knowledge about its world, its *umwelt*, in von Uexküll's phrase, is not always clear. The technique of generalization testing has generally been used to study the effects of reinforcement rather than either stimulus equivalence or data structure.

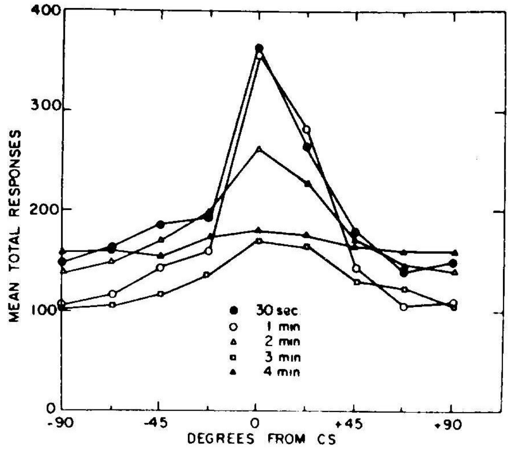


Figure 10.3. The total number of key pecks made by five groups of pigeons, each trained with a different VI schedule (ranging from VI 30-s to VI 4-min), to a line-tilt stimulus during a generalization test. *S*+ was a vertical line (0°). (From Hearst, Koresko, & Poppen, 1964.)

Stimulus control and reinforcement: attention.

Although unidimensional gradients leave us rather far from understanding cognition, they are useful tools for the study of reinforcement mechanisms. The steepness of the gradient is a measure of the degree to which the animal's natural tendency to vary — to respond indiscriminately — is restrained by the schedule of reinforcement. Gradient steepness is also affected by the availability of other sources of control.

For example, Hearst, Koresko and Poppen trained pigeons to peck a key on which a vertical line was projected, and looked at the effect of overall rate of reinforcement on the steepness of the gradient for line tilt. Different groups of birds were trained with different VI schedules. Hearst et al. found that the higher the rate of reinforcement, the steeper the gradient. These results are shown in Figure 10.3. Others have found similar results. Evidently, the better the correlation between a stimulus and food, the tighter the control of a given stimulus dimension over operant behavior, the tighter the limit placed on of the animal's natural tendency to vary. The critical factor is the predictive nature of the stimulus. In general, the stimulus, or stimulus aspect, that best predicts reinforcement will be the one with sharpest control.

A related effect is that control by a highly predictive stimulus aspect tends to weaken control by a less predictive aspect. Thus, in another condition of the Hearst et al. experiment pigeons were trained to peck a vertical line, but the schedule was spaced-responding (food was delivered only for pecks separated by more than 6 sec) rather than variable-interval. In a spaced-responding schedule, *post-response time* is the best predictor of food delivery. The stimulus on the response key guides the actual peck, but has no other significance. On the VI schedule, on the other hand, the stimulus on the key is the best predictor of food. As this difference would lead one to expect, a generalization test showed that control by the line-tilt dimension was much worse for spaced-responding animals than for VI animals. This result is shown in Figure 10.4.

These results lead naturally to the idea of *attention* as a name for the degree of control exerted by a stimulus. The results I have discussed all fit the generalization that animals, sensibly enough, attend preferentially to stimuli, and stimulus properties, that best predict the available goodies (or baddies). The steepness of the generalization gradient provides an objective measure of attention in this sense: a steep gradient indicates that the animal is attending to a given dimension; a flat gradient that he is not (more on attention in Chapter 14).[12]

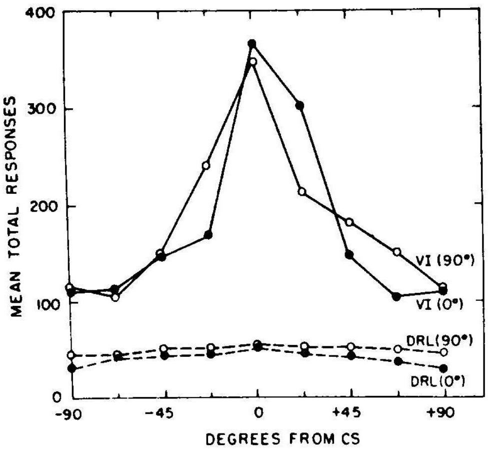


Figure 10.4. Generalization gradients of line tilt following VI or spaced-responding training. There were two groups in each condition, one trained with a vertical line as *S*+, one with a horizontal line as *S*+. (From Hearst et al., 1964.)

Attention to different dimensions

Animals may attend differentially to stimulus dimensions, or stimulus elements. But comparing the control exerted by different dimensions is not always as obvious as in the preceding example, because gradients on different dimensions are not directly comparable. For example, suppose we train a pigeon to respond to a monochromatic light (green, of 550 nm, say) of a certain luminance (say 20 dB above threshold[13]). In a 2-dimensional generalization test we present stimuli differing both in luminance and wavelength from *S*+; five values of each yield total of 25 test stimuli.[14] Suppose that as test wavelengths vary from 400 to 550 nm (at a constant 20 dB luminance), response rates vary from 20 to 60 per minute. As luminances vary from 0 to 20 dB (at a fixed 550 nm wavelength), response rates vary from 30 to 60. Are we justified in saying that since rates vary over a wider range when wavelength is varied and intensity held constant than in the converse case, the animal attends more to wavelength than intensity? Not really, because we could have chosen a different range of intensities — from -10 to 20 dB, for example — over which the range of response variation might have been greater.

There is, of course, a natural scale imposed by the range over which the animal is sensitive to each dimension. If these ranges are known, then the range of experimental variation on the two dimensions can be made comparable in terms of them. For example, if the animal is sensitive to a range of sound intensities of 1 to 10^{10} (i.e., 10 log units -- 100 dB), but only to a range of 1 to 10^2 tonal frequencies (2 log units), then a change of 10 to 1 in intensity (1 log unit) is roughly equivalent to a change of only 1.58 to 1 (0.2 log units) in frequency.[15]

The results of a hypothetical 2-dimensional generalization experiment are illustrated in Figure 10.5, which shows the two independent stimulus dimensions of luminance and wavelength and the response rates to each luminance-wavelength combination, represented as contours of equal response rate. The 2-dimensional gradient is a “hill” with its peak at *S*+. If the two dimensions are appropriately scaled, “equal attention” to both dimensions is represented by circular contours — a symmetrical hill. Of course, we cannot always be sure that axes are in fact scaled appropriately, but in any case *changes* in the contours of Figure 10.5 can always be assessed.

Figure 10.5 represents how two physical dimensions, such as luminance and wavelength, may combine to affect response rate when one particular wavelength-luminance pair is associated with reinforcement. Yet this representation is probably not an accurate model of how the animal sees things, for two reasons: First, for people, at least, the similarity relations among colored objects are not consistent with a space in which luminance and wavelength are at right angles (I describe the evidence for this in a moment). Second, Figure 10.5 does not separate how things look (i.e., how they are represented internally) from how the animal should act with respect to them. I just suggested that "attention", defined as the effect of reinforcement on stimulus control, corresponds to the shape of the contours in Figure 10.5, but there is no reason to suppose that the way colors look, and how similar they are to one another, is much changed by the association of some colors with food (although the idea of search image, discussed in a later chapter, is in partial contradiction to this — but search images are probably restricted to stimuli that are intrinsically hard to discriminate, so they may represent a special case). The notion that animals can separate the physical properties of events from their hedonic consequences seems to be a prudent working hypothesis. I suggest a different way to represent attention in a moment.

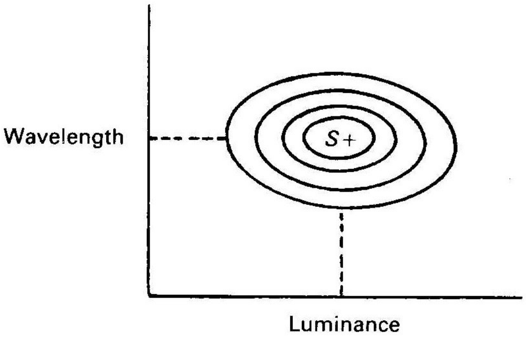


Figure 10.5. A hypothetical two-dimensional generalization gradient. Contours are lines of equal response rate.

Attention is usually guided by reinforcement predictiveness (contingency) but when several stimulus aspects are equally predictive, individuals may differ in their attention to different features. The contours in Figure 10.5 are likely to differ somewhat from animal to animal, and even at different stages of training. Individual differences in attention can be particularly striking when different elements are involved. For example, in one experiment,[16] pigeons were trained on a *successive discrimination*: Two color-form compound stimuli alternated at 60-sec intervals. In the presence of one compound stimulus (*S*+, e.g., a white triangle on a red background) food reinforcement occurred on a VI 60-sec schedule; food was never delivered in the presence of the other color-form compound (*S*-, e.g., a white cross on a green background). This arrangement, in which successive stimuli signal different conditions of reinforcement, is termed a multiple schedule; this particular version is a multiple VI EXT. After the birds had learned to peck at *S*+ but not at *S*-, they were given a generalization test with the four individual stimulus elements (two colors, and two forms projected on a black background), and the four possible color-form compounds.

Typical results from three pigeons are shown as Table 10.2. C and F denote the color and form elements of *S*+, c and f the elements of *S*-; CF, cf and so on are the test compounds. The entries in the table are the proportions of responses during the entire generalization test made to each of the eight test stimuli. The different results from the three pigeons are typical: one bird, #65, apparently attended almost entirely to color; the bird responded not at all to individual elements other than C. A second bird, #66, attended primarily to form, responding negligibly to C. A third animal, #71, clearly attended to both color and form, allocating substantial proportions of responding to both C and F.

Table 10.2 Proportion of responses made to test stimuli made up of color (c or C) and form (f or F) elements of S+ (CF) and S- (cf) in an attention experiment.

	Bird no.		
Test Stimulus	**65**	**66**	**71**
C	.39	.07	.24
F	0	.22	.10
CF	.51	.24	.62
c	0	.01	0
f	0	.21	0
cf	0	0	0
cF	0	.01	0
Cf	.10	.23	.05

More careful scrutiny of Table 10.2 shows that these simple characterizations are approximate at best. For example, bird #65 responded only to stimuli containing C, the *S*+ color, but it also responded much more to CF, *S*+ itself, than to C alone. Clearly, F, the *S*+ form, has an effect in the compound, even if it elicits no responding by itself. Bird #71 shows a similar super-additivity: the proportion of responding to CF is considerably greater than the sum of response proportions to C and F. Bird #66 responded much more to form than color, but it also responded to f, the *S*- form. The response rule for this animal looks like, "respond to any form, except when the background is the *S*- color."

These results illustrate three general principles. One is familiar: that animals will learn to respond for food in the presence of a stimulus[17] that signals responding will be effective, and will learn not to respond in the presence of a stimulus signaling that responding will be ineffective. That is, they can learn a successive discrimination. In any discrimination task, *S*+ and *S*- are likely to differ in several ways. Hence, discrimination *performance* — responding to *S*+ but not to *S*- — is consistent with several different patterns of control by the elements and dimensions of S+ and S-. The second principle is that different animals are likely to pick different patterns of control: in learning, what is not explicitly constrained (by the contingencies of reinforcement), is very likely to vary. Third, test results show that stimulus control involves both *excitatory* and *inhibitory* effects. Excitatory stimulus control is familiar, but as the results for bird #66 show, a positive result can also be achieved by a mixture of excitatory and inhibitory effects. In fact the weight of evidence reviewed in the next chapter shows that all stimulus control is excitatory, but inhibitory effects are produced by excitatory control of activities antagonistic (in the sense that they compete for the available time) to the measured response.

Since different animals can reach the same objective — respond to *S*+, but not to *S*- — in different ways, tests are required to pinpoint the particular rule followed by a given animal. The rule postulated for bird #66 in Table 10.2 — respond to any form, unless the background color is c — implies two things not explicitly tested: that bird #66 would respond even to a novel form, and that c, the *S*- color, is an inhibitory stimulus. To test the first implication, a new test stimulus, say a circle, would have to be presented; the prediction is that the bird would continue to respond. To test the implication that c is inhibitory, c must be compounded with a new form known to produce responding when projected on a "neutral" background; if c is indeed inhibitory, the result should be a decrease in the level of response. This is the standard test for inhibitory control by a stimulus element. Inhibitory control by a stimulus *dimension* is also associated with the production of an inhibitory or incremental generalization gradient. All these effects reflect interactions among competing activities, each controlled in an excitatory way by different stimulus aspects; I return to these topics in the next chapter.

We are obviously much farther along in understanding the ways in which particular physical stimulus elements control the overt behavior of animals than in understanding how they put these things together to represent the world. Yet intuition suggests that underlying the malleability of behavior in response to shifting contingencies of reinforcement must be some invariant structure corresponding to the unchanging aspects of the physical world. Time, three-dimensional space, the properties of solid objects in relation to one another, all are independent of the animal. Although by judicious manipulation of rewards and punishments we can cause animals to do different things with respect to different physical features, still the features themselves are not changed. In a properly designed organism, one feels, there should be some representation of these invariances, an *umwelt*, that remains even as behavior shifts in response to the exigencies of reinforcement. Some promising beginnings have been made towards measuring how people represent objects and some of these ideas have implications for animal studies. I conclude this chapter with a discussion of this work.

SIMILARITY

People can readily group things on the basis of similarity. Asked to classify birds (say) into three or four groups, most would place *hawk* and *eagle* in one class and *robin* and *blackbird* together in another. Experiment has shown that people can give a number to the *dis*similarity between pairs of things: *hawk* and *robin* would get a high number, *hawk* and *eagle* a low one. These numbers can be used to define a *similarity* space, in which the distances between objects correspond to their dissimilarity: *hawk* and *eagle* would be close together in such a space, *hawk* and *robin* would be far apart. Other experiments have shown that distances in similarity space predict the time people take to switch attention from one object to another: if the two objects are similar (close together in the space), the time is short; if they are dissimilar (far apart in the space), it is long.[18]

There are two kinds of similarity space: the simpler kind takes physically defined objects and transforms their physical dimensions so that the transformed values reflect similarity. The second kind is derived by multidimensional scaling techniques from empirical data; the dimensions of such a space need not be simply related to physical dimensions. I give some examples of the first approach now. Data on similarity relations among colors, and among Morse-code characters, discussed later, provide examples of the second type.

The essence of the first kind of space can be illustrated by a simple example. The left panel of Figure 10.6 shows the results of a hypothetical experiment in which generalization gradients were successively obtained with three different S+s: S_1, S_2 and S_3. The figure shows that the gradient spread is directly related to the value of S+: the same stimulus difference produces a large decrease in response rate in the test with S_1 as S+, a much smaller decrease when S_3 is S+. This difference implies that in terms of psychological effects of the same physical difference is substantial at the low end of the scale, but less impressive at the high end. This dependence of differential sensitivity on base value, the Weber-Fechner relation, is characteristic of many physical dimensions, notably sound and light intensity, and time: for example, a second is a lot when one is judging intervals on the order of three or four seconds, it is not so much when judging intervals on the order of minutes.

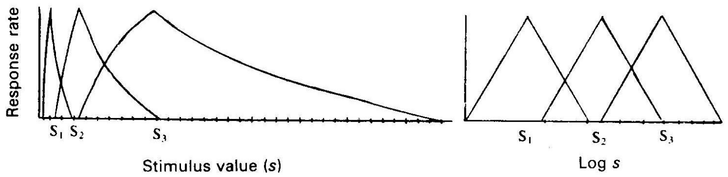


Figure 10.6. *Left panel*: three hypothetical generalization gradients in which range is proportional to S+ value. *Right panel*: gradients as they appear when the stimulus axis is logarithmic.

The right panel of Figure 10.6 shows the transformation of the stimulus axis that reduces these three gradients to the same form, namely, $s' = \log s$, where s' is the psychological (transformed) stimulus value, and s the physical value. Thus, in this hypothetical case (which is not too different from many actual cases that obey the Weber-Fechner relation), equal psychological *differences* are represented by equal physical *ratios*. Note that the transformation here applied to the physical stimulus axis is unrelated to the actual form of the generalization gradient. The essential property of this transformation is just that it renders equal deviations from S+ equally effective, in terms of the measured change in responding.

Figure 10.6 illustrates a simple transformation that relates the unidimensional world of physical intensities to its psychological representation. The same strategy can be applied to more complex aspects, such as color. For example, numerous experiments have obtained judgments from people of the subjective similarity of color samples. In an experiment by Ekman (1954), subjects were asked to rate the similarity of all possible pairwise combinations of 14 spectral (single-wavelength) colors. The results can be plotted as a series of 14 generalization gradients: for each color (wavelength), the height of the gradient is just inversely related to the judged similarity between it and any of the other 13 wavelengths. Shepard (1965) found that these gradients could not be made similar by any transformation of the wavelength axis that preserves the straight-line form with red and blue at opposite ends. He showed that the gradients can be made similar by a transformation that allows the wavelength axis to curl around so that the shortest wavelengths are adjacent to the longest. The transformed gradients are shown at the top of Figure 10.7; the middle panel shows the 14 transformed gradients superimposed on each other, and the bottom panel shows the transformation used: it is just the familiar *color circle*, in which the blue end joins up with the red end, with purple in between.

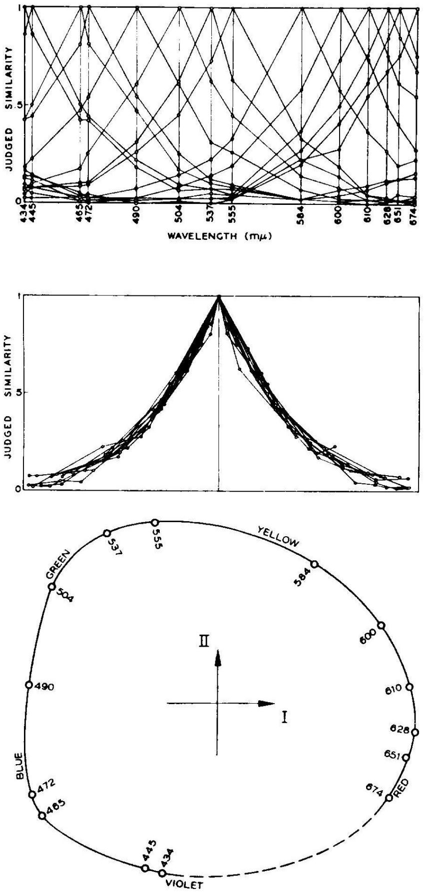


Figure 10.7. *Top panel*: generalization gradients derived from similarity judgments of color samples, with the stimulus axis transformed as shown in the bottom panel. *Middle panel*: transformed gradients from the top panel superimposed to show similar form. *Bottom panel*: stimulus transformation derived by multidimensional scaling techniques — a close approximation to the color circle. (From Shepard, 1965.)

Two things about the color circle are worth noting. First, although there seems to be only one physical dimension involved, wavelength, the space is 2-dimensional. Second, although distances (dissimilarities) are measured directly from one wavelength to another, the region in the center of the circle does not contain wavelengths. In fact, of course, regions inside the circle correspond to *desaturated* colors (single wavelengths diluted with white light). Thus, the 2-dimensional similarity space does correspond to two psychological dimensions: colors, around the rim of the circle, and saturation, varying radically, from white in the center to highly saturated at the rim.

The color space is rather unusual as spaces go. It is 2-dimensional and Euclidean (in the sense that distances conform to the Pythagorean theorem), but the physical dimensions of the stimuli in it don't follow a simple pattern: "north-south" does not correspond to one physical dimension and "east-west" to another. Physical dimensions need not follow any particular pattern in a similarity space. Nor is it necessary that psychological dimensions such as color or saturation be located in any particular position within it. Its essential property is that it should represent accurately the invariances in a set of behavioral data.

In both these examples the "objects" dealt with have been simple physical quantities. This reflects the usual bias of the experimentalist, but it is not necessary — and it may not even be the best way to understand how stimuli are represented by animals. The power of the similarity-space approach is that one can begin with essentially *any* object, even (perhaps especially) "natural" objects such as color pictures of actual scenes.

Practical difficulties have meant that with one exception, rather little work of this sort has been done with animals. The exception is of course work on *orientation*, how animals find their way about their natural environments. There is now ample proof that rats, for example, do so by means of a map that represents the distances and directions of objects in a more or less Euclidean way. The function of particular stimuli is not so much to directly elicit approach or avoidance as to tell the animal where he is in relation to his map, that is, to function as landmarks. We saw a primitive example of this in the light-compass reaction discussed in Chapter 3. Many animals use this reaction (modulated by an internal clock) to define north-south by the position of the sun; given a single landmark or a cue for latitude, they can then locate themselves perfectly.

In the laboratory, experiments with mazes show that visual cues outside the maze usually act as landmarks rather than simple push-pull stimuli. For example, in a recent series of experiments, rats have been presented with an 8-arm radial maze with food boxes at the end of each arm. Rats very quickly learn to pick up each piece of food without revisiting an arm. To do this, they need to know where each arm is in relation to their current position, and they seem to do so by forming a map of the maze with extramaze cues as landmarks. Thus, if, after picking up four pieces of food (i.e., visiting four arms), a rat is covered up in the center of the maze and the arms are rotated, it will revisit arms. However, these revisited arms will be in the spatial position (in relation to the rest of the room) of arms the animal had not visited before the maze was rotated. Well-trained rats do not explore the maze in any systematic way, but rather know the spatial location of each arm, and can remember the locations they have not visited.

In one ingenious experiment, a radial maze was covered with a black cloth dome onto which stimuli could be fixed. In a series of tests, this stimulus array was altered in various ways to see if individual stimulus elements were guiding the animals. The results showed that the animals were using the whole array to orient their map with respect to the maze. Small changes in the array produced no effect; large changes caused the animals to behave as if in a new maze. No individual element had a dominant effect.[19]

So far as we know, cognitive maps are pretty accurate Euclidean representations. When dimensions other than spatial are involved, however, physical and psychological dimensions are rarely identical. For example, in an early study by Rothkopf[20] a large group of human subjects was asked to identify Morse code signals, a set of 36 different patterns. Rothkopf obtained a 36 x 36 *confusion matrix*, in which the entry in cell (*i,j*) (row *i*, column *j*) indicated the percentage of trials that *i* was identified as *j* — an obvious measure of the similarity of signals *i* and *j*. Shepard applied multidimensional scaling (a technique that automatically finds the space with the smallest number of dimensions necessary to accommodate a given set of similarities[21]) to these data and came up with the 2-dimensional Euclidean space shown in Figure 10.8. As in the color example, simple physical stimulus properties are regularly (but not orthogonally) arranged in the space. There is a gradient corresponding to proportion of dots to dashes going from left to right; and another gradient corresponding to number of dots or dashes going from top to bottom.

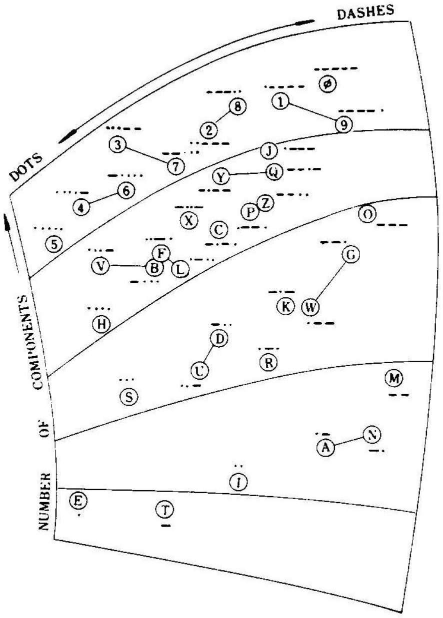


Figure 10.8. Two-dimensional similarity space for the Morse code, derived by non-metric multidimensional scaling by Shepard (1963) from the complete confusion matrix giving percentage of "same" judgments, obtained by Rothkopf (1957).

Spatial representation as a data structure

Spatial representation is a very general idea — we saw it applied with some success to motivational questions in Chapter 7, for example. The world of simple animals, such as the tick referred to earlier, lends itself easily to a spatial description. The animal is sensitive to relatively simple physical properties such as temperature, illumination, altitude above ground (although this may be computed indirectly, from things like illumination, temperature gradient and crawling time), humidity, time of day, and various chemical stimuli. These together define a multidimensional space, and any environment a point in such a space. The representation itself immediately solves the problem of recognition: since the space is defined by just those physical dimensions that are important to the animal, the location of the representative point constitutes "recognition" of a situation. The beast comes equipped by its heredity with a set of preferences: indifference contours in this "world space". By moving around, the animal changes its location not only in the real world, but also in its world space. Simple orientation mechanisms of the type discussed in earlier chapters can then be used to hill-climb up the value gradient. In this way the tick finds the right height, temperature, humidity and other things his genes tell him will promote their survival.

Spatial representation is not just a theoretical trick; it is an efficient data structure. It represents environmental features vital to ticks in a form that allows the animal to compare any set of dimensional values and make a decision about where to go next. In essence, it provides a *relational data-base-management system* (to borrow another computer-science term); that is, a system for representing data that allows open-ended questions to be asked of it: no matter what environment the tick finds itself in, it is never at a loss to evaluate it or decide what to do about it. Spatial representations are efficient, but that doesn't mean evolution has built organisms to use them. Nevertheless, something like this is probably our best current guess about how animals represent their worlds.

The similarity-space idea can be adapted to represent the findings of the 2-dimensional attention experiment described earlier (Table 10.2). Figure 10.9 shows a set of isosimilarity contours in the vicinity of *S*+ in a two-dimensional color-form space. I have made the simplest possible assumption about the form of these contours, namely, that "color" and "form" are equally weighted, so that the contours are circular. The vertical axis corresponds to "similarity to *S*+". As in the Morse-code example, we can expect that colors and forms will correspond to gradients in this space. For simplicity I assume these gradients are at right angles, but they need not be.

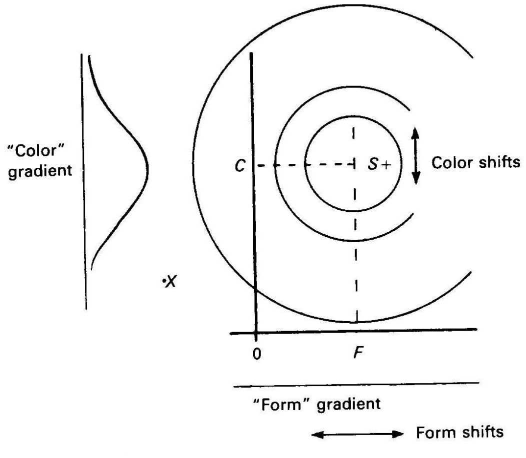


Figure 10.9. Location of the representative point, 0, in a hypothetical two-dimensional color-form similarity space sufficient to represent attention to color (a peaked gradient) but not to form (flat gradient). Point X is the converse case.

Now recall the experimental result (for two pigeons) that in single-element tests, only one element of *S*+ (color, C, for one bird, form, F, for the other) sustained responding. Nevertheless, in the 2-element tests, responding to *S*+ was always greater than responding to the single element: CF > C, for the color animal, CF > F for the form animal. Suppose we represent the color-dominant bird by point 0 in Figure 10.9. Varying color then corresponds to movement along the vertical line through 0. Since this line cuts several isosimilarity contours, similarity to *S*+ varies as color varies and this animal will show control by color: response rate at point C (*S*+ color) will be higher than response rate at point 0 (no color, or at least a color different from the *S*+ color). This control by color is indicated by the gradient on the left. But this same animal will show no control by form, which corresponds to movement along the horizontal line through 0 -- because this line cuts no isosimilarity contours.

The other bird showed opposite results: control by form, not by color. Nevertheless, he can be represented in the same similarity space — by point X. As before, change in color corresponds to vertical movement, but since a vertical line through X cuts no isosimilarity contours, the form-dominant animal will show no control by color. Conversely, horizontal movement from X does cut isosimilarity contours, and this animal will therefore show control by form. Both animals will respond more to *S*+ (CF) than to either C or F alone, however.

Each animal in Table 10.2 corresponds to a particular location of the representative point in similarity space. The location of the point corresponds to "attention": when the point is closer to *S*+ on the "form" axis than the "color" axis, the animal appears to attend more to color (in the sense that color variations cause larger response-rate changes than form variations); when it is closer to *S*+ on the "color" axis, it attends more to form. Thus, the representation in Figure 10.9 separates attention (i.e., response-rate changes) from stimulus similarity; it provides what we are looking for: an invariant representation of the world that underlies diversity of *performance.*

The representation in Figure 10.9 can also describe results showing that generalization gradients are steeper when reinforcement rate is higher, and shallower when other things predict food better than the stimulus being varied. If the "hill" in Figure 10.9 has the bell shape indicated on the left, then the further away from *S*+ the representative point is, the shallower will be the gradient when any physical dimension is varied. If "attention" to a stimulus corresponds to how close the representative point is to the stimulus in similarity space; if alternative sources of reinforcement tend to move the point towards the stimuli that predict each source; and given that movement towards one stimulus will generally mean movement away from another — then the point will tend to be relatively far away from stimuli that are relatively poor predictors of reinforcement, such as stimuli associated with low rates of reinforcement. Consequently, control by such stimuli will be poor.

Figure 10.9 is perhaps best thought of as a small section of similarity space. In reality, other stimuli, such as *S*- and the context in which *S*+ and *S*- occur, must also be represented. Unfortunately, in the absence of a detailed, experimentally based map, Figure 10.9 is just a suggestion on the possible similarity relations that may underlie stimulus control. The figure contains many arbitrary features: the orthogonal axes, the circular isosimilarity contours, the assumed proportionality between response rate and similarity to *S*+, and, most arbitrary of all, the assumption that the space is 2-dimensional. Nevertheless, such a scheme is sufficient to bring together a number of experimental results, and is open to verification — although the experimental effort required to map out similarity spaces is obviously substantial. Practical limitations on the number of stimuli that can be presented in a generalization test, and limits to the reliability of response-rate measures, may mean that better methods are needed before we can assess animals' cognitive representations of simple objects as well as we can those of people.

SUMMARY

This chapter began with the problem of how animals recognize situations. I took recognition for granted in earlier chapters, where we just looked at how animals adapt to various procedures without asking how they know when a particular strategy is appropriate. Traditionally this problem was handled by the concept of *stimulus control.* Particular physical stimuli, or classes of physical stimuli, were said to control particular responses, or classes of responses. This worked well, as far as it went, but it cannot explain relations among complex stimuli. To answer these and other questions, some notion of an *internal representation* seems to be required. A simple and powerful one is to represent things in a multidimensional space. The problem of recognition is solved at once, because any given environment defines a particular point in such a space. If, either innately or as a consequence of past training, different regions of the space have different values, then this representation provides a guide for action, in the sense that it lets the animal know whether his actions have made things better or worse. Thus some kind of spatial representation is our best current guess for the way that animals represent much of their world.

But only a guess: traditional experimental methods of transfer testing with varied elements or physical stimulus dimensions are directed at very different questions and so far tell as rather little about the problem of representation. They do tell us something about the rules describing how physical stimulus elements combine, and about the effects of reinforcement on stimulus control. Stimulus elements can combine additively, multiplicatively, and according to other quantitative principles, all of which can be represented spatially. Stimulus control, measured by the steepness of generalization gradients, is directly related to how well a stimulus feature predicts a valued outcome.

Human experiments on the judgment of similarity provide useful clues to the way that nonverbal objects are represented internally. Some of the results of generalization experiments can also be represented in this way. It is possible that extensions of this approach to animals can tell us more about what they know.

This chapter has been a most unsatisfactory one to write, because the techniques that have been developed to study stimulus effects seem so inadequate to the real task. Generalization testing with single wavelengths was fine so long as we all believed that everything we needed to know about animals could be built up in a routine fashion from atomic stimulus-response units. Research in cognition and machine intelligence has shown, I believe conclusively, that behaving as intelligently as even simple animals behave, requires much, much more. To recognize hundreds or even thousands of different situations and respond rapidly, precisely, and adaptively to them demands a sophisticated world model, as well as a substantial repertoire of stimulus-response routines. We have made much progress in understanding what animals do, but we are still groping to understand what they know.

NOTES

1. The effects of eliciting stimuli such as those for simple reflexes are usually thought of as invariable: a patellar tap always yields a knee jerk. This may well be true of many protective reflexes, but there is no reason to restrict the term *elicitation* just to these. The normal human reaction to stimuli such a fire alarm or a red stoplight is no less elicited because it is also dependent on contextual cues.

2. The term *discriminative stimulus* was coined by B. F. Skinner (1935, 1938) to refer to the special relationship between an “arbitrary” stimulus and a response reinforced in its presence. Such stimuli were said to “set the occasion” for the response, even though in many cases the relation is at least as immediate and powerful as the relation between a reflex (eliciting) stimulus and its response (cf. Staddon, 1967, 1973). The definition I am proposing here does not force one to draw theoretical distinctions between stimulus-response relations that are empirically indistinguishable. It nevertheless accommodates actual usage of the term discriminative stimulus. (See also Chapter 4.)

3. *Discrimination and psychophysics*. I am here passing over in silence several issues important in the history of stimulus control. One such is the relatively unhelpful influence of *psychophysics,* the psychology of sensory processes. In the study of human vision, hearing and touch, it has often been useful to look at the limits of performance of particular sensory systems, attempting to answer such questions as, What is the dimmest light that can be seen, or the quietest sound that can be heard? This is the problem of *absolute threshold*, briefly mentioned in Chapter 2. A related question is the discriminability of differences: What is the smallest change in light or sound

intensity that can be detected? This is the problem of *difference threshold*. It is natural to suppose that generalization, the problem of equivalence, might have something to do with difference thresholds. For example, over regions of the wavelength spectrum where the difference threshold is small, a given physical stimulus difference will encompass many *just noticeable differences* (JNDs), as the difference threshold is termed; the same physical difference over a region where discriminability is poor will span fewer JNDs. Hence, the same physical difference should appear larger in the first case than in the second.

This argument led early students to the hypothesis that the steepness of the generalization gradient at different parts of the wavelength spectrum should be related to the size of the JND. This turned out to be false under most circumstances. Unless special steps are taken to push sensory systems to their limits, generalization effects are not simply related to psychophysical measures. What is left (I will argue shortly) is the notion of *similarity*, a cognitive rather than psychophysical idea. Ultimately stimulus equivalence is determined by the demands the environment makes on the animal. Situations that make similar demands — either now, or in the history of the species — will be treated as, and will appear, similar.

The study of stimulus effects on operant behavior has traditionally been divided into the two areas of stimulus generalization and stimulus discrimination. Stimulus generalization is not too different from the more general question of stimulus equivalence, although both practical difficulties and the influence of psychophysics have tended to concentrate experimental effort on simple physical dimensions such as sound intensity and wavelength of light. But stimulus discrimination has been almost exclusively studied as a psychophysical problem, hence one of more relevance to sensory physiology than to the behavior of the whole organism. For this, the properties of the set of stimulus-equivalence classes — the set of all those things perceived as different by an animal — are of much more relevance than the limits of sensory transduction.

4. The work on the digger wasp was largely done by G. P. Baerends and his associates in Holland. Excellent accounts in English are by Tinbergen (1951, 1958).

5. The limitations on analogue coding deriving from the psychophysical fact that error is often proportional to the absolute value of the coded quantity (the Weber-Fechner relation). Thus, in a counting schedule, error is proportional to N. In digital coding, the intrinsic error rate is lower and not proportional to N. Of course, we have no real idea how things are coded in the brain, and use of the terms *analog* and *digital* here is to some degree metaphorical — although the digital nature of symbolic communication, at least, is self-evident.

Despite this evidence for analog coding of times and stimulus values in animals, there is no reason to suppose that the fundamental operations of brains are anything but digital, although the evidence is indirect and not conclusive: No man-made analogue machine can perform with anything like the intelligence of a suitably programmed digital computer. There are human "lightning calculators" who can perform extraordinary computational feats, apparently often in a largely unconscious fashion — yet accurate numerical computation demands digital processing. It is hard to see how any complex analogue process can avoid the cumulation of error. Every highly developed information process that we know of, from the genetic code to the integrated circuit, is digital.

6. Brown & Jenkins (1968). See Schwartz & Gamzu (1977) for a review. I discuss autoshaping again in a later chapter.

7. In the older ethological literature, unlearned reactions of this sort to well-defined, and often simple, stimuli, termed *releasers*, are referred to as *innate releasing mechanisms* (IRMs); the stereotyped reactions are termed *fixed action patterns* (FAPs). As so often happens, additional

research has blurred this simple picture; these reactions are neither so stereotyped nor so independent of experience as was at one time believed. See Hinde (1970) for a review. I return to this topic later.

8. See for example the now-classic experiments by Herrnstein and his students on "concept formation" by animals, in which pigeons easily learn to recognize color slides of natural scenes exemplifying categories such as "people", "trees" and "water" (Herrnstein and Loveland, 1964; Herrnstein, Loveland and Cable, 1976; Cerella, 1979).

9. This technique was first used by Guttman and Kalish (1956) to study wavelength generalization. See Rilling (1977) and Honig and Urcuioli (1981) for reviews of the extensive work since then.

10. These physiological facts do not really explain generalization, of course, because the relation between transducer coding and behavior of the whole organism is not understood. The fact that neural firing is proportional to sound intensity doesn't force the animal to *behave* in any particular way with respect to intensity differences.

11. Integrality and separability correspond to different kinds of rule defining the transformation necessary to derive similarity relations among compound stimuli from their physical properties. The effect of integral stimuli can be represented by a multiplicative rule: $E = S_1 . S_2$, whereas the effect of two separable elements can be represented by an additive rule: $E = S_1 + S_2$. There are obviously many possible combination rules so that the terms integral and separable do not define all the possibilities.

12. *Attention*. The term *attention* is used rather differently by animal and human experimental psychologists. For both it refers to control of behavior by one stimulus aspect rather than another. But in studies of human attention, individuals typically shift back and forth between attending to this or that several times within a single experimental session, usually in response to verbal instructions. A comparable animal experiment might be to require a pigeon to attend to the color difference between two stimuli in the presence of a tone, but to their size difference in the absence of the tone. This would be termed a *conditional discrimination* experiment in animal discrimination learning. Pigeons don't find such discriminations easy, and attention in animals is studied in other ways. When steady-state behavior is the main interest (as in the study by Reynolds, 1961a, referred to later) the term *attention* is used to refer to selective stimulus control resulting from a history of differential reinforcement. When attention is studied in the context of stochastic learning models, attention usually refers to particular parameters of the models (see Chapter 13). (Mackintosh, 1975, is a good review of some of this work; see the cyberbook edited by Robert Cook, 2001, for an interesting summary of some modern work on bird cognition: http://www.pigeon.psy.tufts.edu/avc/toc.htm).

13. The decibel is a logarithmic unit that expresses the *ratio* of two physical energies. Thus 20 dB means that $10 \log_{10}(L/L_0) = 20$, where L_0 is the threshold luminance and L is the luminance of $S+$. A 3-dB increment represents an approximate doubling of energy. Decibels are more commonly used to represent sound than light intensities. For sound, $\text{dB} = 20 \log_{10}(P/P_0)$, where P_0 is the threshold sound-pressure reference level.

14. This example illustrates another practical difficulty in studying multidimensional generalization gradients: the large number of test stimuli required. For N stimulus values on M dimensions, N^M test stimuli are needed. The practical exigencies of the transfer-test procedure (steady-state

generalization gradients, discussed later, are partially exempt from these limitations) limit the number of different test stimuli to a dozen or two, effectively precluding serious study of more than two stimulus dimensions at a time.

15. The hypothesis that the range of sensitivity predicts the psychological effect of a given physical change has been applied with some success to human psychophysical data (cf. Poulton, 1968; Teghtsoonian, 1971).

16. These results are from an unpublished experiment in my laboratory which was a modest extension of an earlier study by Reynolds (1961a).

17. This description introduces yet another meaning for the term *stimulus*, namely a physically defined *object* presented to an animal by an experimenter. This is different from defining a stimulus in terms of its *effect* on the organism, or as a physically defined event, which are other common usages. It is usually clear from the context which meaning is intended, but use of the same term for so many different concepts sometimes leads to confusion.

18. The idea that stimuli can be represented in a multidimensional Euclidean space seems to have been first proposed by Richardson (1938). Subsequent developments are due to Torgerson (1952), Attneave (1950), Shepard (e.g., 1964, 1980) and Lockhead (1970, 1972).

19. For reviews of the radial-maze work, see Olton (1978, Olton & Samuelson, 1976; Roitblat, 1982; and Dale & Staddon, unpublished). The dome experiment is reported in Suzuki, Augerinos and Black (1980). In a most ingenious series of human experiments, Shepard and Metzler (1971; see also Shepard, 1975; see Cooper, 1982, and Shepard & Cooper, 1982, for a review) have measured the rotation of mental images. Subjects were presented with an asymmetrical target object (a set of cubes arranged to form a figure, but a picture of a left hand would have done as well). They were then shown pictures either of the same object rotated through a variable number of degrees (i.e., the hand turned trough some angle), or of its mirror image (a right hand). The subjects had to respond one way if the object was the same, another way if it was different. Shepard and Metzler found that the *time* taken to respond to the "same" object was proportional to the angular difference in orientation between it and the reference object. Evidently, the subjects had to rotate (at a constant speed!) their mental image of the reference object to bring it into registry with the projected figure before they could decide whether or not the two were the same. In this case, the images were of three-dimensional objects projected on a screen, but the alignment of cognitive maps by means of landmarks is a closely similar process. Indeed, it is something many people become aware of in the course of normal life: for example, many will have noticed the perceptible time it takes to reorient oneself after emerging from an unfamiliar subway entrance, or from a novel department-store exit.

20. Rothkopf (1957). The application of multidimensional scaling techniques to Rothkopf's data is discussed in Shepard (1980) and earlier papers there referred to.

21. *Multidimensional Scaling*. Multidimensional Scaling (MDS) refers to a whole library of techniques representing extensions and elaborations of linear regression. Linear regression finds the straight line that best fits a set of points. MDS finds the plane or higher-dimensional surface that best fits the distances (dissimilarities) between a set of entities.

For example, suppose we have just three items, A, B and C and the dissimilarity relations are A-C: 2 (A and C are quite dissimilar), A-B: 1 (quite similar), and B-C: 1 (quite similar — note it is assumed that the dissimilarity of A vs. B is the same as the dissimilarity of B to A. This is not a trivial assumption, but it makes the math easier). The simplest space that will accommodate these dissimilarities is obviously one-dimensional: A, B and C just lie on a line. But if the dissimilarity of A-C had been 1 rather than 2, the three would have had to be arranged as points of a triangle in a 2-dimensional space.

The similarity relations among *N* objects can always be accommodated by a space of *N*-1 dimensions. The space need not be Euclidean, however. For example, suppose we had the following similarity relations:

A – B: 1,
B – C: 2,
A – C: 4.

There is no way these can be represented in a Euclidean space, but they can be accommodated in a two-dimensional non-Euclidean one.

Obviously, multidimensional scaling becomes useful only when the similarity space is Euclidean or has some other relatively simple form, and the number of dimensions is significantly fewer than the number of entities.